Cybercrime and the Cross-Section of Equity Returns*

Jiatao Liu† Ian W. Marsh‡ Yajun Xiao §

January 15, 2023

Abstract

We use public news coverage about cybercrime to form a cybercrime news attention measure. This measure is consistent with the criteria for a state variable in ICAPM that is expected to forecasts economic conditions, thereby possessing the ability to predict cross-sectional equity returns. We estimate stock-level exposures to a tradeable cybercrime tracking factor. Stocks with the most positive sensitivities generate close to 10% lower annualized risk-adjusted returns than stocks with the most negative sensitivity. Our results indicate that risk-averse investors demand extra compensation to hold stocks with negative cybercrime beta and are willing to pay high prices for stocks with positive beta that hedge exposure to cybercrime. Though our main results are derived from a proprietary cybercrime series, we show that a particularly simple publicly-available alternative based on Google search trends yields very similar conclusions.

JEL Classification: G11, G12, C13, E20

Keywords: Equity returns, cybercrime, state variable, ICAPM, risk premia

*We are grateful to Richard Peterson, managing director of Thomson Reuters MarketPsych, for access to the dataset of news stories used in this paper, and to Aneel Keswani, Richard Payne and Xiao Xiao for discussions and helpful comments. All errors are our own.

†Department of Finance, Xi'an Jiaotong-Liverpool University; Email: Jiatao.Liu@xjtlu.edu.cn

‡Faculty of Finance, Bayes Business School, Email address: i.marsh@city.ac.uk

§Department of Finance, Xi'an Jiaotong-Liverpool University; Email: Yajun.Xiao@xjtlu.edu.cn

1 Introduction

Given the size of the factor zoo, it is natural that there is reduced appetite for research identifying additional risk factors, especially when it is not clear that they add value beyond combinations of other known risk factors. However, new species of risk factors, largely unrelated to better-known financial and economic risks, are more welcome. As awareness of environmental issues has risen, it is natural to expect that markets increasingly price such risks (Choi et al., 2020; Krueger et al., 2020; Bolton and Kacperczyk, 2021; Hsu et al., 2022). Similarly, as the economic environment in which businesses operate evolves, it is likely that new risk factors will emerge. Cybercrime is one such evolution. The U.S. Council of Economics Advisers has reported that malicious cyber activity cost the U.S. economy between \$57 billion and \$109 billion in 2016 (CEA, 2018). The Center for Strategic and International Studies (2018), claim that almost 1% of global GDP, close to \$600 billion, is lost to cybercrime each year.

We test the impact of cybercrime on the cross-section of expected stock returns, seeking to determine whether exposure to cybercrime is priced and to provide an estimate of the market price of cybercrime. We form a cybercrime news attention measure by using a novel news dataset at the daily frequency, distributed as the Thomson-Reuters (now Refinitiv) MarketPsych Index (TRMI) of cybercrime for the United States. This is a natural language processing engine's score derived from articles published in media outlets including both news and social media.

Our investigation is guided by Merton (1973) intertemporal capital asset pricing model (ICAPM). This implies that expected excess returns should vary in the cross-section according to conditional exposures to innovations in state variables that predict future investment opportunities. Bybee et al. (2022) show both theoretically and empirically how news narratives are associated with the idea of systematic risk, in particular, the "state variables" in the ICAPM. They argue that the news narrative captures investors' concerns about future investment opportunities and hence drives the marginal value of consumption and stochastic discount factor. Inspired by Bybee et al. (2022), we shape economic insight with an example of a cybercrime news narrative. Alongside reading an increase in the "cybercrime" narrative in the news media, investors will expect a worsening of investment opportunities, perhaps in the wake of exacerbated economic conditions or increased uncertainty. Essentially indicated in ICAPM, investors' marginal utility and the stochastic discount factor increase as the lessening of concurrent consumption in a state deteriorates "future investment opportunities". Consequently, risky assets that ex-ante more positively covary with the "cybercrime" narrative in the news should have a lower risk premium for its virtue in hedging consumption risk, ceteris paribus, and vice versa. In our context, if cybercrime, proxied by the news attention, adversely (favorably) affects such opportunities (or equivalently, the distribution of future returns from the aggregate wealth portfolio), then we should observe a negative (positive) relation between the average excess return of the asset and that asset's sensitivity to innovations in cybercrime news attention.

We validate that increasing cybercrime news attention negatively forecasts future economic activities and aggregate stock market returns and, in particular, positively forecasts market volatility. As a consequence, the price of risk on the cybercrime news measure is negative, which directs our exploration of its implications in asset pricing. We estimate monthly cybercrime betas (β_{CCA}) using 12-month rolling regressions of daily excess returns on the innovations in the TRMI cybercrime index for stocks traded on the New York Stock Exchange (NYSE), American Stock Exchange, and Nasdaq. We then examine the performance of the β_{CCA} in predicting the cross-sectional dispersion in future stock returns. Stocks in the highest cybercrime beta portfolio generate monthly returns about 0.60% per month lower than the returns from stocks in the lowest cybercrime beta portfolio. Controlling for well-known factors, we find the difference between the returns on the portfolios with the highest and lowest cybercrime beta remains negative and highly significant.

The higher demand for assets that hedge cybercrime risks (high β_{CCA}) increases their price and lowers their average return. We find strong empirical support for this. The Fama-French three-factor alpha on this portfolio is -0.41% per month and is highly statistically significant. Similarly, we might expect that assets exposed to cybercrime should offer higher returns as compensation. Indeed, the portfolio with stocks that are subject to cybercrime risks (low β_{CCA}) offers Fama-French three-factor alphas of +0.23% per month.

Recognising the issues arising from using sensitivities to a non-traded economic series in asset pricing tests, we follow the approach of Lamont (2001) and construct a cybercrime tracking portfolio. Repeating the above exercise and forming portfolios based on sensitivities to the tracking series generates similar results. Stocks with positive sensitivity to the cybercrime tracking portfolio generate significantly lower next-month returns than those from negative sensitivity stocks. The high-minus-low (HL) portfolio generates an average excess return of -0.95% per month, is highly statistically significant, and survives after we control for other factors. The dominant share of the negative HL portfolio return comes from the high cybercrime beta portfolio. This portfolio offers a highly statistically significant alpha of -0.5% per month, compared to the low beta portfolio's statistically significant alpha of 0.4% per month. That is, stocks that offer positive returns when there are positive shocks to cybercrime command a high price and hence offer low expected returns. Stocks that vary negatively with cybercrime shocks, conversely, earn a marginally significant positive risk premium. We push this one step further by splitting the sample into stocks that demand significant amounts of cybersecurity and stocks that do not, based on the industry classifications of the report by World Economic Forum (2016) and Cybersecurity Guide.[1] There is only weakly significant evidence of relationship between cybercrime sensitivities and subsequent returns/alphas for the cybersecruity demanding stocks. Our findings are primarily driven by the stocks that do not demand high levels of cybersecurity

[1] In 2016, WEF issued a report to propose a view of systematic risk caused by cyber-related issues. See the specific information about industry classification based on the importance of cybersecurity in the WEF 2016 report for cybersecurityguide.org.

and their consequent abilities to hedge cybercrime exposure.

To ensure that the return differences are driven by cybercrime exposures rather than well-known stock characteristics or risk factors, we perform bivariate portfolio sorts and re-examine the raw return and alpha differences. We control for size and book-to-market (Fama and French, 1992;1993), profitability and investment (Fama and French, 2015; Hou et al., 2015), betas with the market factor, with market volatility (Ang et al., 2006; Campbell et al., 2018) and with economic policy uncertainty (Brogaard and Detzel, 2015), momentum (Jegadeesh and Titman, 1993), short-term reversal (Jegadeesh, 1990), illiquidity (Amihud, 2002), idiosyncratic volatility (Ang et al., 2006), and the dispersion of analyst earnings forecasts (Diether et al., 2002). The negative relation between the cybercrime beta and future returns remains economically significant after controlling for each of these stock return predictors. We also examine stock-level cross-sectional relations using Fama and MacBeth (1973) regressions. Controlling for all return predictors simultaneously, the cross-sectional regressions provide evidence for an economically and statistically significant negative relation between the cybercrime beta and future stock returns.

Our third major result is that a pricing factor based on the tracking portfolio is capable of pricing a set of 349 portfolios; 300 formed of 10 by 10 bivariate portfolio sorts on (i) size and book-to-market, (ii) size and investment and (iii) size and profitability, plus 49 industry portfolios. In Fama-MacBeth regressions, the cybercrime factor betas prove highly statistically significant and suggest that a two standard deviation increase in sensitivity to the tracking factor results in a 0.65% per month decrease in portfolio return. In their entirety, our results suggest that the cybercrime factor commands an economically significant risk premium and helps to price assets beyond current models.

Related Literature As befits an emerging and increasingly important risk in the economy, cybercrime is an active research topic in the finance literature. Studies by Jiang et al. (2020), Jamilov et al. (2021), and Florackis et al. (2022) are closely related to ours.

Florackis et al. (2022) apply textual analysis tools to 10-K risk factor disclosures by firms to generate a firm-level measure of cybersecurity risk for all US-listed firms. The "Item 1A. Risk Factors" section of 10-Ks require firms to describe the most significant risks they feel managers are are exposed to. Jiang et al. (2020) apply several machine learning techniques to a broader set of information - though including 10-K filings - to estimate the ex-ante probability that a firm will face a cyberattack. Both papers then show that their cyber-security measure is related to stock returns. Related to this strand of research, Jamilov et al. (2021) build text-based measures of cybersecurity risk from quarterly earnings calls, finding that exposure to cybersecurity risk affects stock returns both directly and via contagion effects.

All three papers build firm-specific measures of exposure to cybersecurity risk. As such, they naturally conflate idiosyncratic and systematic cybersecurity risks. Our analysis differs since we estimate firms' sensitivities to a common measure of cybercrime reported in news articles. While we acknowledge that there are a variety of measures of cybercrime risk from which to choose, having made this selection we follow standard and transparent techniques to derive firm-specific sensitivities to cybercrime risk. The simplicity of our approach has several advantages, not least the ability to be applied in jurisdictions with less rigorous regulatory filing requirements and/or less informative disclosures through analyst calls than in the United States. We provide reassurance that the choice of cybercrime measure is not critical by generating very similar findings using a particularly simple measure based on publicly-available Google search data.

Our approach has some other advantages. The paper by Florackis et al. (2022) uses data beginning in 2007, at which time fewer than 30% of U.S. firms make cybersecurity-risk disclosures in their 10-K filings. This proportion jumps from 39% in 2010 to over 60% in 2012 following specific guidance from the SEC in 2011. Whether firms' managers chose to disclose the true cyber-risks they faced appears to have been at least partially driven by regulatory requirements for a large part of their sample. Of course, there is also the question of whether firms' managers can accurately assess the cyber-risks that they face given the novelty of this particular risk and the fast rate of change of both vulnerabilities to cybercrime and the scale of the activities of cyber-criminals. We rely on the market's assessments of cybercrime risks as evidenced by stock returns.

Perhaps most importantly, our paper differs in terms of the emphasis we place on firms that are not seen as being at risk from cybercrime. Much of the important asset pricing implications we draw are driven by the underperformance of firms that offer a hedge to cybercrime, rather than any positive risk premium demanded by investors to hold positions in stock vulnerable to cyber attack. The key measure in Jiang et al. (2020) is an ex ante estimated cyber attack probability for the following year, naturally bounded at zero. Florackis et al. (2022) compute a cybersecurity risk index for each firm and this takes the value zero up to at least the 25th percentile of their sample. In these two applications, firms cannot act as hedges for cybercrime risks, they can only be said to be at less risk than other firms. In our paper, the distribution of estimated exposures to cybercrime risk (after controlling for market risks) is distributed reasonably symmetrically around zero. While we find a significant positive Fama-French five factor alpha for the portfolio of firms with the most negative sensitivities to innovations in cybercrime news (approximatelty the most at risk 20% of firms as we sort into five portfolios), we find large and very statistically significant negative alphas for the portfolios of firms with the most positive sensitivities to cybercrime innovations. Similarly, when we split our sample into high and low cybersecurity demanding sub-samples, the most significant findings are concentrated in the low cybersecurity demanding firms that act as effective hedges against cybercrime risk, not the companies that are exposed to the risk. We note that this is also true - though less emphasised - in the existing literature. Jiang et al. (2020) note that most of their alpha "comes from the bottom decile portfolio, where stocks with lowest cyber risk reside."[2]

[2] Florackis et al. (2022) consistently find significant underperformance of their low cyber-risk portfolio with a cybersecurity risk index of zero. Their high cyber-risk portfolio (index = 0.46) only commands a positive alpha in

Jamilov et al. (2021) estimate cyber-risk betas in a manner similar to ours but based on a non-traded cyber-risk measure derived from earnings calls. They too find average portfolio returns are decreasing in cyber-risk betas, though their slope is shallow and the long-short portfolio pays an average return of -3.3% per year (considerably less than the -11% per year that we find).[3] Further, their Table 5 puzzlingly suggests that portfolios of positive cyber-risk beta stocks offer significantly positive alpha.

The paper is organized as follows. Section 2 describes the key data used in the paper and demonstrates that our measure of cybercrime derived from news reports has characteristics that suggest it may be a candidate state variable in the ICAPM setting. Section 3 presents the cross-sectional asset pricing results that demonstrate significant risk premia consistent with the results in Section 2. Section 4 concludes the paper.

2 Data and Initial Analysis

2.1 TRMI cybercrime News Index

We obtain the cybercrime news index data from Thomson Reuter MarketPsych (TRMI). TRMI derives news feeds of newly published content from approximately 40,000 internet news sites. More specifically, the news or social media content of information is assembled via TRMI crawls through hundreds of financial news sites, including, for example, *The New York Times, The Wall Street Journal, The Financial Times, Seeking Alpha* and many other sources that financial professionals widely read. In contrast to the traditional method of lexical analysis used in textual study, the technology used to create TRMI overcomes several shortcomings of the conventional approach broadly used in extant finance and economics studies (detailed information can be found in Peterson (2016)).

The TRMI cybercrime news data used in our empirical study calculates the intensity of cybercrime-related news reported to the public. The index is unipolar and ranges from 0 to 1.[4] The higher the number, the more references to cybercrime stories in news articles. Specifically, the TRMI cyberCrime score is calculated by counting all cybercrime-related news references scaled by the total news references, named *Buzz* within TRMI. Hence, the TRMI cyberCrime is a fraction of total news references and scrutinizes only news stories related to cybercrime events. Based on the TRMI data user guide we create the measure of cybercrime news coverage

a value-weighted setting.

[3] Jamilov et al.'s (2021) estimated betas also appear much smaller in magnitude than the ones we estimate. Average absolute betas in their two extreme portfolios are around 0.035 compared to 0.25 in ours.

[4] There are rare cases of negative values due to *Buzz* being low, and the references to those indexes are "Negated". For example, references about "anti-cybercrime measures" or "fighting cybercrime" will cause a negative value. The words "anti" and "fighting" allow companies whose business prevents or stops cybercrime not to have positive scores. However, only 30 out of 8306 days have negative values, less than 0.4% sample size. Therefore, our results are not affected by excluding days with negative values.

(CCB) by multiplying the cybercrime score and *Buzz*.

$$CCB_t = cybercrime_t * Buzz_t \quad (1)$$

We calculate the CCB_t at a daily frequency.[5] Figure 1 displays the two daily series about cybercrime and CCB from 1998 to 2021. As shown in Figure 1, the cybercrime news score varies at daily frequency and steadily increases from 1998 to 2021, consistent with the advancement of the internet and data implementations in the economy. The CCB_t displays the same pattern as the cybercrime measure, and significantly increases after 2005 from the increased amount of news sources.[6]

2.2 State variable validation of cybercrime

Cybercrime has become a major global concern due to the steady increase in cyber attacks since the internet was born in 1990.[7] Therefore, we propose cybercrime as a new species of risk within a state variable argument in the ICAPM of Merton (1973). In the Appendix, we present a continuous economy with a representative agent whose marginal value of wealth is affected by cybercrime and describe the conditions under which cybercrime can price asset returns cross-sectionally consistent with ICAPM. To be a legitimate pricing factor, it must forecast asset returns ("changes in the investment opportunity set") or macroeconomic variables (Cochrane, 2009). We empirically verify the candidacy of cybercrime measured by the TRMI news coverage (CCB) as a state variable before using it to discipline our empirical cross-sectional asset pricing.

We follow Maio and Santa-Clara (2012), Boons (2016), and Cooper and Maio (2019), to test whether cybercrime contains characteristics of a state variable in line with the core concept of ICAPM (Merton, 1973). Specifically, we investigate the correlations between the change in cybercrime news coverage and economic activity and components of the investment opportunity set. We conjecture that increasing cybercrime coverage in the news media ought to predict a worsening of future aggregate economic activity and investment opportunity set (and hence, in our subsequent asset pricing tests, stocks with more positive exposure to cybercrime news innovations should offer a negative risk premium). We use growth in industrial production (IPD) and Chicago FED National Activity Index ($CFNAI$) as the measure of future aggregate economic activity.[8] We use future market return (R_{MKT}) and realized variance ($SVAR$) as prox-

[5]The daily measures from TRMI are calculated from newsfeeds before 3:30 PM ET each day.

[6]Please refer to detailed information about the TRMI news sources in the book by Peterson (2016). Additionally, this is a common feature of using news or social media data in textual analysis. Please see the survey studies about textual analysis in finance by Tetlock (2015), Loughran and McDonald (2016), and Loughran and McDonald (2020) .

[7]Studies and governmental reports that propose cybercrime is a "new-born" systematic risk Sommer and Brown (2011), WFE(2013), WEF (2016), TCEA (2018), BOE (2018), ESRB (2020) (Sommer and Brown, 2011; Tendulkar, 2013; ESRB, 2020; Brando et al., 2022)

[8]We obtain the data on both indexes from St. Louis FED database (FRED).

ies for components of the investment opportunity set.[9] First, we calculate the cumulative sum from monthly values:

$$y_{t+1,t+q} = \sum_{i=1}^{q} y_{t+i}, \text{ where } y_{t+1,t+q} \in \{IPD, CFNAI, R_{MKT}, SVAR\} \text{ and } q \in \{1,3,12,24,36,48\}$$

We test $y_{t+1,t+q}$ for each proxy of investment opportunity set or economic activity in the next 1, 3, 12, 24, 36, and 48 months. Second, we calculate the daily ΔCCB_t as the log difference and sum the daily ΔCCB_j in the past 5 years for each month t.

$$\begin{aligned} \Delta CCB_j &= \ln CCB_j - \ln CCB_{j-1} \\ \Delta CCB_t &= \textstyle\sum_{j-365*5}^{j} \Delta CCB_j \end{aligned}$$

First, we investigate the univariate correlation between ΔCCB_t and $y_{t+1,t+q}$. Mainly, we are interested in whether ΔCCB_t has the correct signs as suggested by the ICAPM, viewing cybercrime as a potential state variable indicating a "bad time". Panel A in Table 1 presents the Pearson correlations between the ΔCCB and the measures of future economic activity and market return and volatility. On the one hand, the correlations are significantly negative at most long-horizons for IPG, $CFNAI$, and R_{MKT}. The only exceptions are IPG and R_{MKT} at 1 and 3 horizons. Additionally, Panels A, B, and C in Figure 2 display the series for the past five-year cumulative ΔCCB is negatively related to the next 3, 12, and 24 months cumulative IPG and $CFNAI$. Essentially, there are counter-cyclical between the series of ΔCCB and the other series of measures to show peaks and troughs.[10] Notably, as the horizon increases, there will be more negatively correlated between ΔCCB measures for economic activity and market return. On the other hand, increasing news coverage strongly and positively correlates with future market volatility at all horizons. The magnitude of the average correlation coefficients between ΔCCB and $SVAR$ is relatively more prominent than the other measures. Panel D in Figure 2 shows the strong comovement between the past five-year cumulative ΔCCB and the future market volatility at 3, 12, and 24 months horizon.

Second, we conduct multiple predictive regressions to validate the predictive power of ΔCCB on future economic activity and investment opportunities as follows:

$$y_{t+1,t+q} = a_q + b_q \Delta CCB_t + c_q X + u_{t,t+q} \text{ where } y_{t+1,t+q} \in \{IPD, CFNAI, R_{MKT}, SVAR\} \quad (2)$$

X is a vector of control variables that are well-known state variables in the literature (Maio and Santa-Clara, 2012; Boons, 2016). For example, the term spread is denoted by TS and measured by the difference between the ten-year bond yield and the risk-free rate; default spread is denoted by DS and measured by the difference between BAA and AAA corporate

[9] We download the data for market return and realized variance from Amit Goyal's webpage. $SVAR$ is the logarithm of realized stock variance.

[10] We only tabulate the figures for 3, 12, and 24 months for better visualization.

bond yield; the dividend-to-price ratio is denoted by DP that is calculated by taking the log difference between dividend and price from the market portfolio, and $infl$ is the inflation that is the Consumer Price Index.[11]

Panel B in Table 1 shows the results for multiple predictive regressions. While controlling benchmark state variables, ΔCCB remains significantly negative forecasting power for IDG, $CFNAI$, and R_{MKT} at 12 and longer horizons. Also, the results for testing ΔCCB forecasting power on the future market volatility is consistent with the aforementioned evidence in the coefficients correlation. The five-year daily cumulative ΔCCB has compelling positive forecastability for future market volatility even after controlling for well-known state variables.

In sum, backward-looking cybercrime news coverage predicts a worsening of future aggregate economic activity (negative forecastability on IPG and $CFNAI$) and investment opportunities (negative/positive forecastability on R_{MKT}/$SVAR$). Hence, as cybercrime issues become more critical garnering increased coverage in the news, investors' future investment opportunities decrease, primarily through higher levels of volatility but in addition, though to a limited extent, worse future stock market returns.[12] We conclude that the cybercrime news index contains legitimate characteristics of an ICAPM state variable. This has the important implication that since cybercrime forecasts negative changes in both economic activity and investment opportunities in a time-series sense, its innovations should earn a negative risk premium in the cross-section of stocks. The rest of the paper addresses this implication in detail.

2.3 Cybercrime news shock

Following the standard procedure of calculating news-based risk shocks in the literature,[13] we apply an AR (6) model to extract the cybercrime news innovations as the measure of cybercrime shock (CCA). The first column in Table 2 shows that the CCB is somewhat persistent, with AR(1) coefficient being 0.68. Nevertheless, we reject at 1% level the null hypothesis that CCB has a unit root. To further investigate the potential correlation between the cybercrime news measure and other benchmark economic risk variables, the second column in the Table shows the results by adding ΔVIX and ΔEPU as additional controls in the AR(6) model. Indeed, our cybercrime news measure is not related to ΔVIX and ΔEPU. That is, firstly, we estimate

$$CCB_t = a + b_i \sum_{i=1}^{6} CCB_{t-i} + \varepsilon_t^{CCB} \tag{3}$$

where ε_t^{CCB} is the cybercrime news attention innovations. We select one year as backward rolling window with approximately 365 daily CCB data. Secondly, for each rolling regression

[11] We download these state variables from Amit Goyal's webpage.

[12] In online appendix, for robustness, we also conduct a validation test for ΔCCB_j, it shows strong positive forecasting on the daily realized variance of S&P 500 that is estimated by using tick data from TAQ.

[13] See related studies by Brogaard and Detzel (2015) and Engle et al. (2020).

sample data, we standardized ε_t^{CCB} as follows:

$$CCA_t = \frac{\varepsilon_t^{CCB} - \bar{\varepsilon}_t^{CCB}}{\sigma_{\varepsilon_t^{CCB}}} \tag{4}$$

where $\bar{\varepsilon}_t^{CCB}$ is the mean of cybercrime news innovation in each rolling regression sample and the $\sigma_{\varepsilon_t^{CCB}}$ is the standard deviation. We define CCA_t as the measure of cybercrime news attention shock. Throughout the paper, we estimate the CCA_t based on monthly updated rolling regressions rather than the whole sample to ensure that we do not include any future information in the following tests.[14]

We examine all common stocks (share codes 10 and 11) traded on the NYSE, Amex, and Nasdaq exchanges from January 1998 through December 2021. The daily and monthly return data are from the CRSP, and we adjust stock returns for delisting effects following Shumway (1997). Following Amihud (2002) and many other studies, we eliminate stocks with a price per share less than \$5. Financial variables are obtained from the merged CRSP-Compustat database. Analysts' earnings forecasts come from the Institutional Brokers' Estimate System (I/B/E/S) data set. Google search variable (SVI) is downloaded from the Google open API. Benchmark pricing factors and testing portfolios are downloaded from related data libraries.[15]

3 Cross-Sectional Asset Pricing Results

We conduct both parametric and non-parametric tests to assess the predictive power of cybercrime beta for future stock returns. We begin with univariate portfolio-level analyses before moving on to bivariate portfolio-level analyses to confirm the power of the cybercrime beta after controlling for well-known risk factors. We then present cross-sectional regression results. Finally, we replicate our main findings using cross-sections of equity portfolios as test assets.

3.1 Cybercrime news attention shock exposures

The exposures to cybercrime of individual stocks are obtained from one-year rolling regressions of daily excess stock returns on innovations in the TRMI cybercrime news index. Our specific procedure is as follows:

First, we compute normalised innovations in the cybercrime series as equations (3) and (4). We consider innovations to this series as it exhibits a strong first-order correlation (of

[14]This standardization also benefits the following beta estimation will be more comparable since the cybercrime news coverage is different as technological progress changes dramatically in the sample period (from 1998 to 2021). For example, Figure 1 shows that samples after 2005 have relatively more cybercrime news stories. Therefore, standardization is helpful to make each rolling regression sample has the same scale, particularly zero mean and unit standard deviation.

[15]We thank Kenneth R. French, Robert F. Stambaugh, and Lu Zhang for providing open sources of pricing factor data.

around +0.8 over the full sample). The sample starts on January 1st, 1998, so our first rolling window runs from this date until December 31st, 1998. [16]

Second, for each stock, we regress daily returns on the market return and the cybercrime news innovations series over the same one-year window:

$$R_{i,t} = \alpha_i + \beta_{MKT,i} R_{MKT,t} + \beta_{CCA,i} CCA_t + \varepsilon_{i,t} \tag{5}$$

where β_{CCA} is the estimated cybercrime beta. We require at least 60 trading days for a stock to be included. We only control for the market factor since Liu et al. (2021) argue that the CAPM outperforms more complicated models in testing zero alphas from a market efficiency perspective.[17] Nevertheless, we control for many benchmark factors in subsequent portfolio sorting analyses.

Stocks with a negative $\beta_{CCA,i}$ suffer poor returns when there are positive innovations in the cybercrime news coverage (*CCB*). Stocks with positive sensitivities to cybercrime news attention are hedging stocks that offer insurance against positive shocks in cybercrime news. When the public's awareness of cybercrime increases from the news media report, these stocks offer positive returns. Such hedging stocks should command a high price and hence offer a low-risk premium if investors are concerned about cybercrime negatively shifting future economic conditions or investment opportunity set. Conversely, stocks with negative $\beta_{CCA,i}$ are more risky and exposed to cyber-related crimes, so those should command a positive risk premium.

The third step is to allocate each stock to one of five portfolios and compute the returns on these cybercrime exposures sorted portfolios in January 1999, the month following portfolio formation. All portfolios are value-weighted and we use NYSE stock break-points throughout.

We then move the sample forward one calendar month, re-estimate both the cybercrime innovations series and the firm-level exposures to cybercrime, and compute next-month portfolio returns. We continue until the sample is exhausted in December 2021. This rolling-window approach means that all estimates are based on information available to investors in real-time with no look-ahead bias. It also accounts for the increasing level and volatility of the TRMI cybercrime index over our sample. Standardising the innovations using full-sample data as done by Jamilov et al. (2021) is potentially problematic given that the daily change in the TRMI index has a standard deviation of XXX in 1998 but YYY in 2021.[18]

The left-hand panel of Table 3 presents the univariate portfolio results. For each month, we form five portfolios by sorting individual stocks based on cybercrime beta (β_{CCA}). Portfolio 1 contains stocks with the lowest (most negative) β_{CCA} during the past month, and portfolio

[16] We use an AR(6) model as this reduces the autocorrelation issue, controls for potential day-of-the-week effect, and results in standardised innovations that pass both ADF and KPSS stationarity tests in each rolling window.

[17] The backward-looking rolling regression with the CAPM model is also suggested by Barroso et al. (2021) to capture the conditional relationship between the state variable and the tested variables. We also estimate the $\beta_{CCA,i}$ by using FF3 and CARHART Model, and the results are barely changed.

[18] Standardization with full sample information contains look-ahead bias.

5 contains stocks with the highest (most positive) β_{CCA}. The first column reports the average β_{CCA} for each portfolio using full-sample breakpoints and an equal number of stocks in each portfolio. The second column presents the average value-weighted excess return and associated t-statistics for each portfolio. The last row reports these for the high-minus-low portfolio (P5-P1).

The first column of Table 3 shows that there is significant cross-sectional variation in the average values of cybercrime betas as we move from portfolio 1 to portfolio 5. The average beta rises monotonically from -0.23 to +0.23. Stocks in the central portfolio, P3, have mean cybercrime betas very close to zero. The symmetry of the mean cybercrime betas across the portfolios is striking (and is not greatly affected by using more or fewer portfolios).

The second row shows that next-month average excess returns decrease monotonically from 0.94% to 0.34% per month when moving from the lowest to the highest β_{CCA} portfolios. The average return difference between highest and lowest beta portfolios is -0.60% per month with a Newey and West (1987) t-statistic of -2.66.

The remainder columns in the left panel present the magnitude and statistical significance of risk-adjusted returns (alphas) from three four factor models: (1) the Fama-French three factor model (α_3), (2) the Fama-French five factor (α_5) (3) the Fama-French five factor plus the momentum factor (α_6), and (4) the Fama-French five factor model plus momentum, short-term and long-term reversal factors (α_8). Results are quite insensitive to the specific factor model used. Mean value-weighted alphas decline monotonically from around 0.3% for Portfolio 1 (low cybercrime beta) to around -0.4% for Portfolio 5 (high cybercrime beta). The alphas earned by Portfolio 1 are marginally statistically significant with Newey-West t-statistics of around 2.0. The alphas of high cyber-risk beta stocks (Portfolio 5) are statistically significant with t-statistics in excess of 2.4. The high-minus-low portfolio mean alpha is also stable at around -0.60% and is strongly statistically significant for all four alternative factor models.

3.2 Cybercrime news-driven hedging demand

The results of univariate portfolio analysis for β_{CCA} are consistent with the theoretical prediction in ICAPM, in which investors pay a higher price for stocks having high β_{CCA} to hedge future bad states. In this subsection, we investigate investors' trading behaviors subject to the aforementioned statement about the news-driven state variable. As investors read an increasing number of stories related to cybercrime in the news, their trading decisions need to incorporate the implied information from the cybercrime narratives. Therefore, in line with the investor demand mechanism in ICAPM, high cybercrime beta stocks exert more buying orders placed by investors. Inspired by studies that explore news-driven trades (Huang et al., 2020; Fisher et al., 2022; Jeon et al., 2022), we use tick-trading data from TAQ to create proxies to verify the cybercrime news-driven trading behaviors, particularly investors place more buying orders as hedging demand on high cybercrime beta stocks.

To examine investors' buy-initiated orders, we use signed dollar-trading volume from TAQ.[19] Specifically, we create three measures as proxies of investors' buy behaviors. First, we calculate the stock buy fraction (BF) as the total buy trading volume over the total volume. Second, we create a relative measure which is the difference between buy trading volume and sell trading volume (BS). Additionally, we measure investors' buy orders by taking the logarithm of the total buy trading volume (BV).[20]

On the one hand, we identify cybercrime news-driven orders by running univariate regressions for each stock in each cybercrime beta estimation rolling sample.

$$y_{i,t} = \gamma_0 + \gamma_{i,CCA} CCA_t + u_{i,t} \ , \ \ y_{i,t} \in \{BF, BS, BV\} \tag{6}$$

where $\gamma_{i,CCA}$ is the stock sensitivity of buying orders to cybercrime news attention shock. The higher the value of γ_{CCA}, the more buying orders are placed by investors who read an increase in cybercrime news narratives.

On the other hand, to investigate the relationship between cybercrime news-driven buy orders and cybercrime beta stocks, we first sort quintile portfolios based on β_{CCA} in each month and calculate the contemporaneous average γ_{CCA} in each cybercrime beta portfolio. Panel A in Table 4 presents the non-parametric results. There are strongly monotonic increasing patterns for the sensitivity of buy orders to cybercrime news attention shock (γ_{CCA})from P1 to P5 in all three measures of investors' buy-initiated orders. The results in HL portfolio for γ_{CCA} are also positive and statistically significant for BS and BV.[21] Therefore, consistent with the investors' hedging demand statement in ICAPM, investors place more cybercrime news-driven buy orders as hedging demand on high cybercrime beta stocks contemporaneously.

Second, to further disentangle potential effects that may confound the positive relationship between β_{CCA} and γ_{CCA}, we conduct Fama-Macbeth cross-sectional regressions as parametric analysis.[22] Panel B in Table 4 displays the results consistent with the model-free results in Panel A. Columns (1), (3), and (5) are univariate regression results. All three measures positively relate to β_{CCA}. The results of adding control variables in columns (2), (4), and (6) do not comprise the strong implication that stocks with high cybercrime beta are bought more by investors. The high demand for high cybercrime beta stocks is induced by the high cybercrime news attention.

All in all, stocks' sensitivity of buy orders to cybercrime news attention accords with the stocks' cybercrime beta. Stocks with high cybercrime beta are contemporaneously demanded more by investors who place more buy orders on these stocks for hedging purposes stated in ICAPM.

[19]The algorithm used to sign trading orders follows the method by Lee and Ready (1991).

[20]For stock dollar-trading volume, we divide it by \$1000,000.

[21]The HL portfolio is calculated by taking the difference between P5 and P1 for the average of γ_{CCA} in each portfolio.

[22]In untabulated tables, we also conduct pooled or fixed effect regressions, and the results are barely changed at all.

3.3 Cybercrime tracking portfolio construction

Having demonstrated that cybercrime betas are capable of predicting the cross-sectional variation in future stock returns, we now construct a tracking factor that captures the returns associated with cybercrime.

Non-traded factors - including cybercrime news - which capture fundamental risks in the economy ought to explain the cross-section of expected returns. However, measured changes in these factors contain measurement errors. To reduce factor noise, factor-mimicking or tracking portfolios containing traded assets that represent the underlying non-traded factors are widely used (Huberman et al., 1987; Breeden et al., 1989; Giglio and Xiu, 2021). We follow the time-series approach of Lamont (2001) and regress the non-traded cybercrime series on contemporaneous returns of traded assets (Z_t), using the fitted values from this regression as a traded asset-based proxy for cybercrime for each one-year long rolling regression window:

$$CCA_t = c + b'Z_t + u_t \tag{7}$$

The traded assets we use are the five portfolios sorted according to sensitivities to the cybercrime news attention described in the previous section. Essentially, we apply the technique of economic tracking factor construction developed by Lamont (2001) with ex-ante tracking rather than studies using ex-post constructions (Ang et al., 2006; Engle et al., 2020). Specifically, by the end of each portfolio formation month, we know which stocks are in portfolio 5 for hedging and which stocks are sorted into other portfolios (1 - 4). Hence, from the first day of each rolling regression sample, we sort stocks into five portfolios by using the portfolio number based on β_{CCA} obtained in the portfolio formation month (the last day in each rolling regression sample). In other words, the base assets Z_t - five portfolios only contain information up to the last day of each rolling regression without any future information.[23] Because the traded assets are excess returns, the coefficients in the vector b can be interpreted as weights in the zero-cost portfolio. For each regression window, we construct the daily tracking factor return factor, $b'Z_t$, which we denote $TCCA_t$.

$$TCCA_t = b'Z_t \tag{8}$$

The tracking portfolio $TCCA$ contains the portfolio of asset returns maximally correlated with realized innovations in cybercrime news coverage using a set of basis assets with different exposures to cybercrime news attention (β_{CCA}). Accordingly, we are able to examine the contemporaneous relationship between average returns and factor loadings of cybercrime news attention in a manner of portfolio return. By virtue of this mimicking factor, the primary advan-

[23] For example, the first rolling regressions use data from 1/1/1998 to 12/31/1998. Therefore, on 12/31/1998, we have β_{CCA} for all stocks and sort them into five portfolios. Next, we take each stock with related portfolio numbers to form daily portfolios from 1/1/1998 to 12/31/1998. Consequently, we have portfolios that mimic CCA.

tage of using $TCCA$ in the following analysis to measure the aggregate cybercrime risk is that we have a good approximation of innovations in cybercrime news, and allows us to alleviate the issue caused by hidden noises in the news data.

Panel A in Figure 3 presents the average daily Pearson correlation between CCA_t and the cybercrime news attention tracking portfolio returns is around 0.78 from the 264 regression windows. The correlation ranges from 0.52 to 0.92. Figure B displays time-varying weights (b) for the five portfolios. On average, the weights of portfolio 5 are always positive and the average weight on portfolio 5 is close to +0.95. In the meantime, the weights of portfolio 1 are always negative and the average weight on portfolio 1 is close to -0.86. Additionally, the average weights on each portfolio are also monotonically increasing. Together, these results suggest that the tracking portfolio indeed efficiently tracks innovations in cybercrime news (CCA) in a manner consistent with expectations.

3.3.1 Tracking factor performance

We next repeat the procedure of the previous section but now estimate time-varying stock-level sensitivities to the cybercrime news attention tracking factor ($TCCA_t$), rather than to the non-traded cybercrime news attention proxy itself.

$$R_{i,t} = \alpha_i + \beta_{MKT,i} R_{MKT,t} + \beta_{TCCA,i} TCCA_t + \varepsilon_{i,t} \tag{9}$$

Rolling one-year regressions of returns on the market excess return and the $TCCA$ generate cybercrime tracking betas (denoted $\beta_{TCCA,i}$) that are used to allocate stocks to one of five portfolios. We then examine the returns on these portfolios in the following month. The results, reported in the right-hand panel of Table 3, are very similar to those in the left-hand panel. As we move from Portfolio 1 (smallest betas) to Portfolio 5 (largest betas) value-weighted average portfolio betas rise from -0.66 to +0.81, again in a relatively symmetric pattern. Similarly, next-month average excess returns decrease monotonically from 1.08% to 0.14% per month. The returns of individual stocks in Portfolio 1 correlate negatively with shocks to cybercrime news mimicked by $TCCA$ and so risk-averse investors require extra higher expected returns to hold these stocks. Conversely, as the stocks in Portfolio 5 correlate positively with increases shocks in cybercrime they are viewed as hedge stocks that perform well in times of increased risk related to cybercrime. Hence investors pay higher prices for these stocks and willingly accept lower returns.

The average return difference between the highest and lowest beta portfolios is -0.95% per month with a Newey and West (1987) t-statistic of -2.93. Alpha analysis reported in columns 3-6 of Panel B also replicates the conclusions based on cybercrime betas (β_{CCA}). Irrespective of the factor model used, monthly alphas from Portfolio 1 are around 0.4% and statistically significant around 1%, falling to around -0.60% (with very large t-statistics) for Portfolio 5. The high-minus-low portfolio return alphas are only slightly below the raw return and are statistically significant.

This significantly negative cybercrime premium is much as predicted by the intertemporal capital asset pricing model of Merton (1973). An unexpected increase in cybercrime adversely affects future investment and consumption opportunities. Investors prefer to hold stocks whose returns increase upon such unfavorable events and thus hedge their exposures to cybercrime. That is, they compensate for reduced consumption and future investment opportunities by holding stocks that positively correlate with cybercrime. This intertemporal hedging demand implies that investors are willing to pay higher prices and accept lower returns for stocks with higher cybercrime betas.

It is noticeable from Table 3 that irrespective of the pricing factor model used, the majority of the negative alpha in the high-minus-low portfolio comes from the high cybercrime beta leg (Portfolio 5). This proportion is never below 60% and in the case of the Fama-French five-factor model is as high as 64%. The alpha from the high beta leg is always statistically significant while the low beta leg (Portfolio 1) provides an alpha that is always marginally significant. Stocks in portfolio 5 are the hedges, that tend to offer high payoffs when cybercrime news increases and these stocks offer typically low expected returns as a result of this hedge characteristic.

3.4 The effect of cybersecurity demand across industries

This section examines the significance of cybercrime risk premium for stocks across different industries. We group stocks into two categories that depend on the importance of cyberbersecruity in different industries based on the guidelines provided by WEF (2016) and Cyberseurityguide.org. This is an open-source website and provides cybersecurity educational information. Based on the industry analysis in Cyberseurityguide.org and Fama-French industry SIC codes, firms in Consumer nondurables, Energy, Hightech, Telecommunication, and Healthcare sectors are classified as heavily demanding cybersecurity. The other remaining sectors are classified as less demanding of cybersecurity. We now conduct dependent bivariate sorting to control for the effect of demand for cybersecurity. Within each cybersecurity group, we sort stocks in to five portfolios based on β_{TCCA}.

Table 4 reports the alphas and associated t-statistics for each portfolio using the usual factor models. The right panel reports results for stocks in the industries where the demand for cybersecurity is high, there is evidence that these firms contribute to significantly positive risk premium in portfolio 1. While the portfolio 5 with the highest cybercrime tracking beta offer negative alpha, it is only marginal significant at 10% level.

Conversely, the left panel presents evidence for stocks in the industries where cybersecurity is less critical. These firms generate alphas more consistent with our broader findings to-date. Alphas decline (non-monotonically) as betas rise, and these are significantly negative for the portfolios comprised of the stocks with the highest sensitivities. A high-minus-low portfolio based on stocks demanding less cybersecurity yields a monthly alpha close to -1% and again, over 60% of this alpha comes from Portfolio 5.

3.5 Average stock characteristics and tracking factor betas

In this section, we determine the average characteristics of stocks with high and low cybercrime tracking factor betas (β_{TCCA}). We use Fama-MacBeth (1973) cross-sectional regressions of cybercrime tracking factor betas for each firm in the sample on stock-level characteristics and risk factors:

$$\beta_{TCCA,i,t} = c + \delta X_{i,t} + \varepsilon_{i,t} \tag{10}$$

where X contains market betas (β_{MKT}), market volatility betas (β_{VIX}) estimated by following the study by Ang et al. (2006) , economic policy uncertainty betas (β_{EPU}) estimated by using the economic policy uncertainty index (EPU, Baker et al. (2016)) augmented from the CAPM , the log of market capitalisation (SIZE), book-to-market ratio (BM), operating profit (OP), investment (I/A) (Fama and French, 2015), short-term reversal (LRET) (Jegadeesh, 1990), Amihud illiquidity (ILLIQ), idiosyncratic risk (Ang et al., 2006),[24] and momentum (MOM) (Jegadeesh and Titman, 1993). Additionally, we include analyst forecast variables such as the number of analysts following a stock and their forecast dispersion (Diether et al., 2002).

Column 1 of Table Table 5 shows that the average slope coefficient on the market beta is positive and significant, implying that stocks with high cybercrime beta have a high market beta. This finding is consistent with Freazzini and Pedersen's (2014) finding, evidence that stocks have lower one-month ahead return. Column (4) and (5) report that the average slope coefficients on size and BM is significantly negative and positive, respectively. Therefore, stocks with high cybercrime beta are large and value stocks. The size result is consistent with Fama-French (1992,1993) that small stocks have higher expected returns than big stocks. However, the BM result shares a different view of value stocks having higher expected returns than growth stocks.

As presented in Column (9), the average slope on illiquidity is negative and significant, indicating that stocks with high cybercrime beta are liquid and the low cybercrime beta stocks are illiquid. This result is consistent with the study by Amihud (2002) that explores illiquid stocks generate higher one-month ahead returns as the liquidity risk premium. Our high β_{TCCA} stocks serve as the hedging assets; therefore, liquidity as an implementable trading purpose is not a major concern in our hedging argument.

Column (11) displays the average slope on momentum is negative, indicating the cybercrime hedging stocks (higher β_{TCCA}) underperform in the past 11 months. We do not find

[24] We follow the same rationale in these studies, however using Fama-French five-factor rather the 3-factor model.

strong supportive evidence for the relevance of other fundamental variables.

Nevertheless, the final column in Table 5 shows that when we include all variables simultaneously, only β_{MKT}, BM, illiquidity, and momentum survive to have previously strong cross-sectional relation with β_{TCCA}. Overall, stocks with high cybercrime beta are liquid, have large market beta, high BM ratio, and lower past 11-month return performance.

3.6 Bivariate portfolio-level analysis

We next examine the relation between cybercrime sensitivities and next-month stock returns controlling for well-known cross-sectional return predictors. We perform bivariate portfolio sorts on the cybercrime tracking beta (β_{TCCA}) in combination with the market capitalization (SIZE), book-to-market ratio (BM), operating profitability (OP), investment (I/A), market beta (β_{MKT}), market volatility beta (β_{VIX}), economic policy uncertainty beta (β_{EPU}), momentum (MOM), short-term reversal (ST), illiquidity ($ILLIQ$), idiosyncratic volatility ($IVOL$), and analyst dispersion ($DISP$). We first form five portfolios based on the predictor variables. Then, within each predictor portfolio, we sort stocks into five portfolios based on the cybercrime beta (β_{TCCA}) so that portfolio 1 (portfolio 5) contains stocks with the lowest (highest) cybercrime beta values. We then average the next month's value-weighted portfolio returns across the five predictor portfolios for each of the five cybercrime beta portfolios. This creates a set of five portfolios with very similar levels of the predictor variable but which differ by cybercrime beta. We report value-weighted portfolio results from these conditional bivariate sorts in Table 6.

The first column of Table 6 shows that after controlling for size, the α^8 controlling FF five-factors, momentum, short-term and long-term reversal factors tends to fall as the cybercrime tracking portfolio beta increases from portfolio 1 to 5. The high-minus-low portfolio alpha is about -0.7% per month with a Newey-West t-statistic in excess of 2.3. Subsequent columns of Table 6 show a very similar pattern and in most cases, the high-minus-low alpha is even greater than that seen when controlling for size. Statistical significance is strong for all predictors. We conclude therefore that none of the predictors explains the high (low) returns on low cybercrime beta stocks.

3.7 Stock level cross-sectional regressions

We have tested the ability of cybercrime tracking betas (β_{TCCA}) to determine the cross-section of future returns at the portfolio level. Non-parametric portfolio-level analysis has the advantage of not imposing a specific functional form on the relation between β_{TCCA} and subsequent returns. However, it also has at least two disadvantages. First, aggregation loses a large amount of information in the cross-section via aggregation, and second, it is difficult to control for more than one other factor simultaneously. Hence, as is standard, we now examine the cross-sectional relation between cybercrime betas and expected returns at the stock level using Fama and MacBeth (1973) regressions of the following form:

$$R_{i,t+1} = \lambda_0 + \lambda_1 \beta_{TCCA_{i,t}} + \lambda' X_{i,t} + \varepsilon_{i,t} \tag{11}$$

where $R_{i,t+1}$ is the realized excess return of stock i in month $t+1$, $\beta_{TCCA_{i,t}}$ is the cybercrime tracking portfolio beta of stock i in month t, and $X_{i,t}$ is a collection of stock-specific control variables observable at time t for stock i. X contains market betas (β_{MKT}), market volatility betas (β_{VIX}), economic policy uncertainty betas (β_{EPU}), the log of market capitalization (SIZE), book-to-market ratio (BM), operating profit (OP), investment (I/A), lagged returns (LRET), illiquidity (ILLIQ), idiosyncratic risk ($IVOL$), momentum (MOM), and the dispersion of analysts forecasts (DISP). The cross-sectional regressions are estimated monthly from January 1999 to December 2021.

The univariate regression results reported in the first column of the left panel in Table 7 indicate a negative and statistically significant relation between the cybercrime beta and the cross-section of future stock returns. The average slope is -0.46 with a Newey-West t-statistic of 5% significance. The second and third columns shows that adding several standard control factors only reduces the magnitude of this coefficient slightly, and it remains statistically significant. The average stock in Portfolio 1 has a cybercrime beta of -0.66, while in Portfolio 5, this rises to 0.0.81 (see Table 3).[25] Were a stock to move from Portfolio 1 to Portfolio 5, all other things equal, the expected return of that stock would decrease by around 0.68% per month (-0.46×(0.66-(-0.81)).

We add industry controls in the right-hand panel of Table 7. The results are barely changed from those without industry controls. In an unreported table, we also test the long-term predictive power of cybercrime tracking beta with 6, 9, and 12 months ahead of returns. In sum, β_{TCCA} negatively predicts the next 6, 9, and up to 11-month future monthly returns and remains marginal predictability with a 10% significant level for the 12-month ahead monthly return. This long-term predictive power of cybercrime tracking factor beta is consistent with the extant study by Florackis et al. (2022) that uses the firm-level measure of cybersecurity risk.

3.8 Power of β_{TCCA} for cross-sectional equity portfolios

To be consistent with earlier findings from individual stocks, we investigate if cybercrime tracking beta has the same predictive power for the cross-section of equity portfolios. We obtain portfolio daily return data from Kenneth French's data library. The portfolios we use in this test include 49 industry portfolios and three sets of 10×10 portfolios that are bivariate sorts based on size and book-to-market, size and investment, and size and profitability such that we consider 349 portfolios in the estimation. These portfolios are widely used in the literature since they generate significant cross-sectional differences in portfolio expected returns.

Following the estimation procedure in sections 3.1 and 3.2.1, we first estimate rolling

[25] Since illiquidity may have a high correlation with size, we control these two variables separately for a clear presentation. However, the results are mostly the same if we control size and illiquidity in the same model.

cybercrime betas from regressions in the form of:

$$r_{p,t} = \alpha_i + \beta_{TCCA,p} TCCA_t + \sum \beta_{p,j} f_{j,t} + \varepsilon_{i,t} \ , \quad j \in \{MKT, SMB, HML, CMA, RMW\} \tag{12}$$

for each test portfolio. We control for Fama-French five factors and $\beta_{TCCA,p}$ is the cybercrime tracking beta for each portfolio.[26]

Table 8 presents the univariate portfolio sorting results. Notably, the quintile portfolios sorted by β_{TCCA} estimated by cross-sectional portfolio returns have consistent results with the ones by using individual stock data. The first column shows how the average cybercrime tracking beta increases from portfolio 1 to 5. The one-month ahead expected returns decrease monotonically when moving from the lowest to the highest β_{TCCA}. The HL portfolio generates -0.3% negative expected returns. The difference in risk-adjusted returns (α) between high β_{TCCA} and low β_{TCCA} portfolios is significantly negative, and the magnitude is about -0.25% and consistent across different pricing models.[27] Again, we note that the alpha is concentrated in the high beta portfolio.

Overall, the results in Table 8 indicate that the cross-sectional return difference induced by high β_{TCCA} and low β_{TCCA} estimated on equity portfolio returns are negative and statistically significant after controlling classical pricing factors. Hence, we conclude that the cybercrime tracking beta is priced not only in the cross-section of individual stocks but also in the cross-section of equity portfolios. One can also imply that hedging against cybercrime can be implemented via cross-sectional portfolios rather than diving into the entire stock universe, which is relatively costly.

3.9 Ex-post cybercrime risk pricing factor

We have presented evidence that the sensitivities of stock returns to a tradeable factor that tracks innovations in cybercrime news coverage predict the cross-sectional variation in future stock returns. We now test the ability of an ex-post cybercrime pricing factor to explain the returns on large numbers of equity portfolios.

We follow the methodological rationale from the study by Lamont (2001) and implement the similar practice used by Ang et al. (2006) and Engle et al. (2020) to construct the ex-post pricing factor as follows. From each rolling window one-year regression of equation (7), we estimate weights, b_t, on returns from each candidate value-weighted traded asset portfolio, Z_t, that track the risk exposure to innovations in the cybercrime news series. Therefore, we have portfolio weights at the end of each portfolio formation period (see Panel B in Figure 3 for time-varying weights of base assets). We then calculate the ex-post pricing factor, $FCCA_{t+1}$, as:

[26] The results are not sensitive to use the other models for our beta estimation.

[27] We also use q-factor to test the cross-sectional predictive power of β_{TCCA} estimated by 349 equity portfolios. The risk-adjusted return based on the q-factor is α_q.

$$FCCA_{t+1} = b_t' \times R_{t+1} \tag{13}$$

where $R_{t+1} = r_{p,t+1}^{\beta_{CCA}}$ is the vector of the five portfolio returns in the month following the estimation window. For example, in step 1, we estimate the weights b' using data from 01/01/1998 to 12/31/1998 by using equation (7). In step 2, we multiply b' by the vector of returns earned by the five value-weighted portfolio returns in January 1999 sorted by β_{CCA} to obtain the cybercrime pricing factor return for January 1999. We then roll forwards one calendar month, estimating weights over the period 2/1/1998 to 1/31/1999 and multiplying these by returns earned in February 1999 to obtain the pricing factor return for February 1999. We continue to roll forwards by one month until December 2021.[28]

The first column of Table 9 shows that the ex-post cybercrime news attention pricing factor earns a -0.44% per month, with an associated t-statistic of 3.51. Subsequent columns show that while this factor's returns are marginally correlated with other benchmark pricing factors, it bears little relation to the returns of other commonly used factors and a large unexplained component remains irrespective of the benchmark factors included in the regressions. Harvey et al. (2016) argue that the usual five percent level is too low a threshold when testing for statistical significance of a new pricing factor because of data mining concerns and the large extant body of research examining the cross-section of expected returns. They argue the case for higher requirements before we can accept empirical results as evidence of true economic phenomena. Specifically, they suggest that a new factor needs an associated t-statistic greater than three. It is then comforting to note that the average monthly return of the tracking factor has an associated Newey-West t-statistic of about -3.20.

3.10 The price of cybercrime risk

To estimate the unconditional factor risk premium λ_{FCCA} on the ex-post mimicking cybercrime factor $FCCA$, we source monthly returns based on the same 349 portfolios used in section 4.2.6.

We then estimate two-step Fama-Macbeth (1973) regressions including the standard Fama-French five factors, augmented by the cybercrime pricing factor:

$$r_{i,t} = \alpha_p + \sum \beta_i \lambda_i + \beta_{FCCA} \lambda_{FCCA} + \varepsilon_{i,t} \ , \quad i \in \{MKT, SMB, HML, CMA, RMW\} \tag{14}$$

where $r_{i,t}$ is the excess return on each of the 349 portfolios. In the first stage, we estimate betas using the full sample from January 1999 to December 2021. In the second stage, we conduct cross-sectional regression to estimate factor premia. Our interest is in λ_{FCCA}, the price of cybercrime risk. The results are presented in Table 10. The first column uses the Fama-French

[28] We also construct the ex-post mimicking pricing factor by following the classical method developed by Fama and French (1993). The result is very similar.

five factors to price the 349 portfolios. All the factors are statistically significant at conventional levels, but as usual, some of the prices are not of the predicted sign. More relevant for this paper, column 2 augments the regression with pricing factor betas. The regression results suggest a price of cybercrime risk of -2.82% per month. The standard deviation of β_{FCCA} is 0.12. Therefore, a two standard deviation increase in β_{FCCA} is associated with a decrease in the portfolio return of a little over 0.68% per month.

Similarly, the right-hand panel in Table 10 presents results based on the q-factor model. The results are consistent with the Fama-French model test in the left panel. On average, the price of cybercrime risk is about -3.95% per month in the 349 portfolios. Furthermore, a two standard deviations increase in β_{FCCA} is associated with a decrease in the portfolio return of about 1.1% (2×0.144×-3.95%=-1.1%). All in all, the cybercrime mimicking factor induce significantly negative risk premia that are attributed to exposure to the cybercrime risk that is quantified by $FCCA$.

3.11 Robustness checks

3.11.1 Google search trend data

The basis of our analysis has been the TRMI cybercrime news series, essentially a measure of the supply of information about cybercrime in the press. This section shows that an alternative, publicly-available measure of the demand for information about cybercrime generates very similar results. Specifically, we use daily Google search trend data on the single keyword“cybercrime” from 01/01/2007 to 12/31/2021, an interval considerably shorter than the one used for TRMI-based analysis but still long enough to provide meaningful results. We take the first log difference of the Google search trend measure as a measure of investors’ demand for cybercrime-related issues.

We estimate the cybercrime Google cybercrime beta as follows:

$$\Delta SVI_t = \log SVI_t - \log SVI_{t-1}$$

$$R_{i,t} = \alpha_i + \beta_{MKT,i} R_{MKT,t} + \beta_{SVI,i} \Delta SVI_t \tag{15}$$

thus, we conduct the portfolio analysis and contract the ex-ante portfolio tracking factor (denoted by $TSVI$)in the Google Search Trend data universe. Table 11 shows the results are very consistent with the ones in Table 3 as we have done by using the TRMI cybercrime news index. Additionally, we also repeat the construction of the ex-post tracking factor. In an untabulated appendix, the results are consistent with the tracking factor created by using news-measured cybercrime risk.[29]

We conclude that our results are insensitive to considering innovations in the supply of

[29] Please find the untabulated tables from the Online Appendix

information on cybercrime from the TRMI data series, or innovations in the demand for cybercrime information from Google search data. In both cases, hedging stocks that offer better returns when such shocks occur command a high price and hence offer lower returns, on average.

3.11.2 Alternative pricing factor construction methods

In this sub-section, we demonstrate the insensitivity of our results to constructing the pricing factor using the Fama-French approach, rather than the regression approach detailed in section 3.9. We also continue to demonstrate the consistency of our results when using Google search trend proxies of cybercrime.

The left panel of Table IV in the Online Appendix reports the results of constructing a Fama-French-style pricing factor using estimates of stocks sensitivities to innovations in the TRMI series (β_{CCA}). Stocks are sorted into intersections of two portfolios according to market capitalization using NYSE breakpoints and three portfolios according to cybercrime news sensitivities. The factor is then constructed as the average return in the large and small cap high sensitivities portfolios minus the average return in the large and small cap low sensitivities portfolios. The average return on this factor is -0.373% per month with an associated t-statistic of 2.81. The other columns in this panel show that this survives the inclusion of alternative commonly-used factors.

The right-hand panel replicates this but uses sensitivities to the Google search trend of cybercrime rather than innovations to TRMI's cybercrime news series. The headline results show that the pricing factor is economically large, statistically significant, and is not much affected by accounting for other factors remains. However, compared with both the results in the left-hand panel and those in Table 9, it is noticeable that several other factors are significantly related to this version of the pricing factor. Nevertheless, including these only serves to increase the magnitude of the pricing factor's alpha.

When we repeat the 349 equity portfolio factor pricing tests, both of these alternative factors prove to be highly significant. The TRMI-based Fama-French-style factor bears a coefficient of -2.99 with an associated t-statistic of 4.03. A one standard deviation increase in $\beta_{C\hat{C}A}$ is associated with a decrease in expected portfolio return of 0.71% per month. We conclude that, if anything, our core findings from on a regression-based pricing factor are conservative.

3.11.3 Robust to only S&P 500 Stocks

We further investigate if our results are driven by small and illiquid stocks, which are not implementable and suffer data mining issues stressed by Harvey et al. (2016). We re-create Table 3 by only testing stocks from S&P 500. Table VI in the Online Appendix shows the results even prevail in the cross-sectional large, liquid, and S&P 500 stocks. Therefore, the results are convincing and implementable in line with the hedging story in ICAPM.

4 Conclusion

This paper investigates the role of cybercrime in the cross-sectional pricing of individual stocks and equity portfolios. Cybercrime is quantified by a news-based index provided by TRMI though this specific choice of proxy series is not crucial, and we obtain similar results based on simple and publicly-available Google search trends data.

More cybercrime-related issues reported in the news media forecast a worsening of future investment opportunities, in particular higher levels of market volatility but also to a lesser extent lower aggregate stock market returns. This suggests that our news-based measure of cybercrime has characteristics of a state variable in the ICAPM of Merton (1973). If this is the case, variations in stock exposures to innovations in cybercrime news should predict cross-sectional variations in stock returns. Given that cybercrime forecasts a worsening of the investment opportunity set and economic activity, cross-sectional returns should be lower as the sensitivity of stock returns to cybercrime shocks increases.

We estimate beta exposures of US stocks to innovations in cybercrime news. Consistent with an ICAPM interpretation, stocks that covary negatively with innovations in cybercrime news (that is, firms that have negative cybercrime betas) on average offer high next period returns. In contrast, stocks with positive betas hedge cybercrime shocks, paying off when cybercrime news innovations are high, but offer poor next period returns on average.

Bivariate portfolio-level analyses and stock-level cross-sectional regressions that control for well-known pricing effects, including size, book-to-market, momentum, short-term reversal, liquidity, idiosyncratic volatility, dispersion in analysts' earnings estimates, market volatility beta, economic policy uncertainty beta, investment, and profitability generate similar results. After controlling for each of these variables one by one and then controlling for all variables simultaneously, we provide evidence of a significantly negative link between cybercrime betas and future stock returns.

Cybercrime betas predict a significant proportion of the cross-sectional dispersion in future stock returns. The economic tracking factor analyses indicate an annualized risk-adjusted return of -5.04%. These results hold when we instead consider equity portfolio exposures to a cybercrime tracking factor. The important role of cybercrime hedging stocks is also demonstrated when considering sub-samples of our data. Controlling for the level of dependence of cybersecurity across 10 industries, the negative risk premium of hedging stocks is mainly contributed by firms that require less cybersecurity.

References

Amihud, Y. (2002). Illiquidity and stock returns: cross-section and time-series effects. *Journal of financial markets*, 5(1):31–56.

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

Baker, S. R., Bloom, N., and Davis, S. J. (2016). Measuring economic policy uncertainty. *The quarterly journal of economics*, 131(4):1593–1636.

Barroso, P., Boons, M., and Karehnke, P. (2021). Time-varying state variable risk premia in the icapm. *Journal of Financial Economics*, 139(2):428–451.

Bolton, P. and Kacperczyk, M. (2021). Do investors care about carbon risk? *Journal of financial economics*, 142(2):517–549.

Boons, M. (2016). State variables, macroeconomic activity, and the cross section of individual stocks. *Journal of Financial Economics*, 119(3):489–511.

Brando, D., Kotidis, A., Kovner, A., Lee, M., and Schreft, S. L. (2022). Implications of cyber risk for financial stability.

Breeden, D. T., Gibbons, M. R., and Litzenberger, R. H. (1989). Empirical tests of the consumption-oriented capm. *The Journal of Finance*, 44(2):231–262.

Brogaard, J. and Detzel, A. (2015). The asset-pricing implications of government economic policy uncertainty. *Management science*, 61(1):3–18.

Bybee, L., Kelly, B. T., and Su, Y. (2022). Narrative asset pricing: Interpretable systematic risk factors from news text. *Johns Hopkins Carey Business School Research Paper*, (21-09).

Campbell, J. Y., Giglio, S., Polk, C., and Turley, R. (2018). An intertemporal capm with stochastic volatility. *Journal of Financial Economics*, 128(2):207–233.

CEA (2018). The cost of malicious cyber activity to the u.s. economy. *The Council of Economic Advisers*.

Choi, D., Gao, Z., and Jiang, W. (2020). Attention to global warming. *The Review of Financial Studies*, 33(3):1112–1145.

Cochrane, J. (2009). *Asset pricing: Revised edition*. Princeton university press.

Cooper, I. and Maio, P. (2019). Asset growth, profitability, and investment opportunities. *Management Science*, 65(9):3988–4010.

CSIS (2018). 2018 election security scorecard. *Center for Strategic and International Studies*.

Diether, K. B., Malloy, C. J., and Scherbina, A. (2002). Differences of opinion and the cross section of stock returns. *The Journal of Finance*, 57(5):2113–2141.

Engle, R. F., Giglio, S., Kelly, B., Lee, H., and Stroebel, J. (2020). Hedging climate change news. *The Review of Financial Studies*, 33(3):1184–1216.

ESRB (2020). Systemic cyber risk. *European Systematic Risk Board*.

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

Fama, E. F. and French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of financial economics*, 33(1):3–56.

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

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

Fisher, A., Martineau, C., and Sheng, J. (2022). Macroeconomic attention and announcement risk premia. *The Review of Financial Studies*, 35(11):5057–5093.

Florackis, C., Louca, C., Michaely, R., and Weber, M. (2022). Cybersecurity risk. *The Review of Financial Studies*.

Frazzini, A. and Pedersen, L. H. (2014). Betting against beta. *Journal of financial economics*, 111(1):1–25.

Giglio, S. and Xiu, D. (2021). Asset pricing with omitted factors. *Journal of Political Economy*, 129(7):1947–1990.

Harvey, C. R., Liu, Y., and Zhu, H. (2016). … and the cross-section of expected returns. *The Review of Financial Studies*, 29(1):5–68.

Hou, K., Xue, C., and Zhang, L. (2015). Digesting anomalies: An investment approach. *The Review of Financial Studies*, 28(3):650–705.

Hsu, P.-H., Li, K., and Tsou, C.-Y. (2022). The pollution premium. *Journal of Finance, Forthcoming*.

Huang, A. G., Tan, H., and Wermers, R. (2020). Institutional trading around corporate news: Evidence from textual analysis. *The Review of Financial Studies*, 33(10):4627–4675.

Huberman, G., Kandel, S., and Stambaugh, R. F. (1987). Mimicking portfolios and exact arbitrage pricing. *The Journal of Finance*, 42(1):1–9.

Jamilov, R., Rey, H., and Tahoun, A. (2021). The anatomy of cyber risk. Technical report, National Bureau of Economic Research.

Jegadeesh, N. (1990). Evidence of predictable behavior of security returns. *The Journal of finance*, 45(3):881–898.

Jegadeesh, N. and Titman, S. (1993). Returns to buying winners and selling losers: Implications for stock market efficiency. *The Journal of finance*, 48(1):65–91.

Jeon, Y., McCurdy, T. H., and Zhao, X. (2022). News as sources of jumps in stock returns: Evidence from 21 million news articles for 9000 companies. *Journal of Financial Economics*, 145(2):1–17.

Jiang, H., Khanna, N., Yang, Q., and Zhou, J. (2020). The cyber risk premium. *Available at SSRN 3637142*.

Krueger, P., Sautner, Z., and Starks, L. T. (2020). The importance of climate risks for institutional investors. *The Review of Financial Studies*, 33(3):1067–1111.

Lamont, O. A. (2001). Economic tracking portfolios. *Journal of Econometrics*, 105(1):161–184.

Lee, C. M. and Ready, M. J. (1991). Inferring trade direction from intraday data. *The Journal of Finance*, 46(2):733–746.

Liu, J., Moskowitz, T. J., and Stambaugh, R. F. (2021). Pricing without mispricing. Technical report, National Bureau of Economic Research.

Loughran, T. and McDonald, B. (2016). Textual analysis in accounting and finance: A survey. *Journal of Accounting Research*, 54(4):1187–1230.

Loughran, T. and McDonald, B. (2020). Textual analysis in finance. *Annual Review of Financial Economics*, 12:357–375.

Maio, P. and Santa-Clara, P. (2012). Multifactor models and their consistency with the icapm. *Journal of Financial Economics*, 106(3):586–613.

Merton, R. C. (1973). An intertemporal capital asset pricing model. *Econometrica: Journal of the Econometric Society*, pages 867–887.

Newey, W. K. and West, K. D. (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica*, 55(3).

Peterson, R. L. (2016). *Trading on sentiment: The power of minds over markets*. John Wiley & Sons.

Shumway, T. (1997). The delisting bias in crsp data. *The Journal of Finance*, 52(1):327–340.

Sommer, P. and Brown, I. (2011). Reducing systemic cybersecurity risk. *Organisation for Economic Cooperation and Development Working Paper No. IFP/WKP/FGS (2011)*, 3.

Tendulkar, R. (2013). Cyber-crime, securities markets and systemic risk. *CFA Digest*, 43(4):35–43.

Tetlock, P. C. (2015). The role of media in finance. *Handbook of media Economics*, 1:701–721.

WEF (2016). Understanding systemic cyber risk. *World Economic Forum*.

Appendix

A ICAPM and Cybercrime

Consider an economy populated by a representative investor facing an optimal consumption and asset allocation problem continuously. There are N risky assets and one risk-free bond. A risky asset i earns an instantaneous risky return and the risk-free bond earns an instantaneous locally risk-free rate

$$\begin{aligned} dR_t^i &= \mu_{i,t}dt + \sigma_{i,t}dw_{i,t} \\ dB_t &= r_t dt \end{aligned}$$

where $(w_1, w_2, \ldots, w_N)'$ is a correlated N-dimensional Wiener. $\mu_{i,t}$ is the expected return, $\sigma_{i,t}$ the instantaneous volatility, r_t the instantaneous risk-free rate. These quantities depend on *CCB*, i.e., $\mu_{i,t} \equiv \mu_i(z,t)$, $\sigma_{i,t} \equiv \sigma_i(z,t)$, and $r_t \equiv r(z,t)$. *CCB* is exogenous and modeled by a diffusion process:

$$dz_t = \mu_{z,t}dt + \sigma_{z,t}dz_t$$

where z_t is a one-dimensional Wiener process. For easy composition, The instantaneous risk return is assumed to be uncorrelated with *CCB* in any instant time, i.e., $\mathbb{E}_t[dw_{i,t} \cdot dz_t] = 0$. The specification suggests that shocks to *CCB* affect asset returns.

The representative investor maximizes her lifetime utility by solving the optimal consumption and investment problem subject to the intertemporal budget constraint:

$$\mathbb{E}_t \left[\int_{s=t}^{\infty} e^{-\rho(s-t)} \frac{c_s^{1-\gamma}}{1-\gamma} ds \right].$$

$\rho > 0$ is the constant discount rate and $\gamma > 1$ the constant relative risk aversion. The investor cares about and takes into account cybercrime news in her maximization. As a result, her consumption and marginal utility of consumption (wealth) responds to the innovation of cycbercrime news coverage (shock of cybercrimew news). It implies that *CCB* describes the conditional distribution of stock returns the investor will face in the future or "shifts in the investment opportunity set."[30]

The value function V is the lifetime utility at optimum. It is a function of wealth W_t and *CCB* z_t at time t. Indeed, in our economy with the homothetic power preference, the value function has the form

$$V(W,z) = \frac{W^{1-\gamma}}{1-\gamma} f(z). \tag{16}$$

$f(\cdot)$ is a positive and differential function in *CCB*. Since the marginal utility of consumption

[30] We are concerned with the implication of cybercrime news attention shock and do not pursue in solving the optimal consumption and portfolios for the representative investor. Given the power preference and the diffusion return processes, the optimal consumption is proportional to wealth, the optimal portfolios are a sum of hedging demand and myopic demand that are characterized by the value function and the investment opportunities (Merton, 1973)

equals the marginal value of wealth, the stochastic discount factor Λ_t is

$$\Lambda_t = e^{-\rho t} V_W(W_t, z_t) = e^{-\rho t} W_t^{-\gamma} f(z_t).$$

Note that *CCB* affects the marginal value of wealth as well as asset returns in equilibrium via $f(\cdot)$. The asset pricing Euler equation implies[31]

$$\mathbb{E}_t[dR_t^i] - r_t dt \quad = \quad \beta_{W,t}^i dt \cdot \lambda_{W,t} + \beta_{z,t}^i dt \cdot \lambda_{z,t}. \tag{17}$$

The term $\lambda_{W,t} = \gamma \sigma_{W,t}^2 > 0$ is the price of risk on aggregate wealth shocks (or "market risks") and $\lambda_{z,t} = -\frac{z_t V_{Wz}(W_t, z_t)}{V_W(W_t, z_t)} \sigma_{z,t}^2$ is the price of risk of *CCB*. $\sigma_{W,t}$ is the volatility of the aggregate wealth process. As usual, the positive λ_W suggests that the risk-averse investor demands the higher expected returns on assets that have more covariances with the shocks to aggregate wealth (market). Whether or not the investor demands the higher expected returns on assets that have more covariances with the innovation to *CCB* depends on the sign of the price of risk of *CCB*, λ_z, which reflects how an increase in *CCB* changes the marginal value of wealth. Maio and Stanta-Clara (2012) and Barroso, Boons, and Karehnke (2021) show the sign of λ_z is essential for *CCB* to price a set assets in the cross section.

It is useful to know that the marginal value of wealth moves in the opposite direction of the investor's value function in response to changes in *CCB*, suggesting that the marginal value of wealth is high when investment opportunities are poor.[32] Suppose an increase in *CCB* indicates "bad times" in the sense of poor investment opportunities (low output/production, high volatilities/uncertainities) in the future. Then, the increase in z_t will be not beneficial for the investor, meaning $V_z < 0$, or equivalently, the increase in z_t raises the marginal utility of

---

[31] The application of the asset pricing Euler equation gives

$$\begin{aligned}
\mathbb{E}_t[dR_t^i] - r_t dt &= -\mathbb{E}_t\left[dR_t^i \cdot \frac{d\Lambda_t}{\Lambda_t}\right] \\
&= -\frac{W_t V_{WW}(W_t, z_t)}{V_W(W_t, z_t)} \cdot \mathbb{E}_t\left[dR_t^i \cdot \frac{dW_t}{W_t}\right] - \frac{z_t V_{Wz}(W_t, z_t)}{V_W(W_t, z_t)} \cdot \mathbb{E}_t\left[dR_t^i \cdot \frac{dz_t}{z_t}\right] \\
&= \beta_{W,t}^i dt \cdot \lambda_{W,t} + \beta_{z,t}^i dt \cdot \lambda_{z,t}
\end{aligned}$$

Since $\mathbb{E}_t\left[\frac{dW_t}{W_t} \cdot \frac{dz_t}{z_t}\right] = 0$, betas in a bi-variate regression setting can be expressed as:

$$\begin{pmatrix} \beta_{W,t}^i \\ \beta_{z,t}^i \end{pmatrix} = \begin{pmatrix} \mathbb{E}_t\left[\frac{dW_t}{W_t}\right]^2 & 0 \\ 0 & \mathbb{E}_t\left[\frac{dz_t}{z_t}\right]^2 \end{pmatrix}^{-1} \cdot \begin{pmatrix} \mathbb{E}_t\left[dR_t^i \cdot \frac{dW_t}{W_t}\right] \\ \mathbb{E}_t\left[dR_t^i \cdot \frac{dz_t}{z_t}\right] \end{pmatrix}.$$

[32] By the value function (16), it is straightforward to show

$$\frac{\partial V(W,z)}{\partial z} = \frac{W^{1-\gamma}}{1-\gamma} f'(z), \quad \frac{\partial^2 V(W,z)}{\partial W \partial z} = W^{-\gamma} f'(z).$$

Given $\gamma > 1$, the marginal value of wealth is positive if the investor's value function in response to changes in *CCB* is negative.

wealth, meaning $\frac{\partial^2 V(W,z)}{\partial W \partial z} = \frac{\partial}{\partial z}\left(\frac{\partial V(W,z)}{\partial W}\right) > 0.$[33] It leads to a negative price of risk of *CCB* with $\lambda_z < 0$ since the margin value of wealth is positive.[34] Therefore, assets that covary highly with *CCB* have high prices driven by high demand and low expected returns. They have high prices, because they have their highest payoff when *CCB* is higher, which are bad times; thus, they are hedging assets. Analogously, an increase in *CCB* indicates "good times" in the sense of good investment opportunities and give rise to a positive price of risk of *CCB*. Assets that have covary highly with *CCB* have low prices and high expected returns; they are not useful for hedging.

[33]For the risk-averse investor ($\gamma > 1$), she values an additional unit of dollar more when investment opportunities are poorer than when they are better. The reason is that it is more difficult for her to increase wealth through investment when investment opportunities are poor.

[34]Alternatively, one can show

$$\begin{aligned} \lambda_{z,t} &= -\frac{z_t V_{Wz}(W_t, z_t)}{V_W(W_t, z_t)} \sigma_{z,t}^2 \\ &= -z_t \frac{f'(z_t)}{f(z_t)} \sigma_{z,t}^2 \\ &< 0 \end{aligned}$$

where $f'(z_t) > 0$ because $V_z(w,z) < 0$.

Figure 1 TRMI cyberCrime News Index and CCB

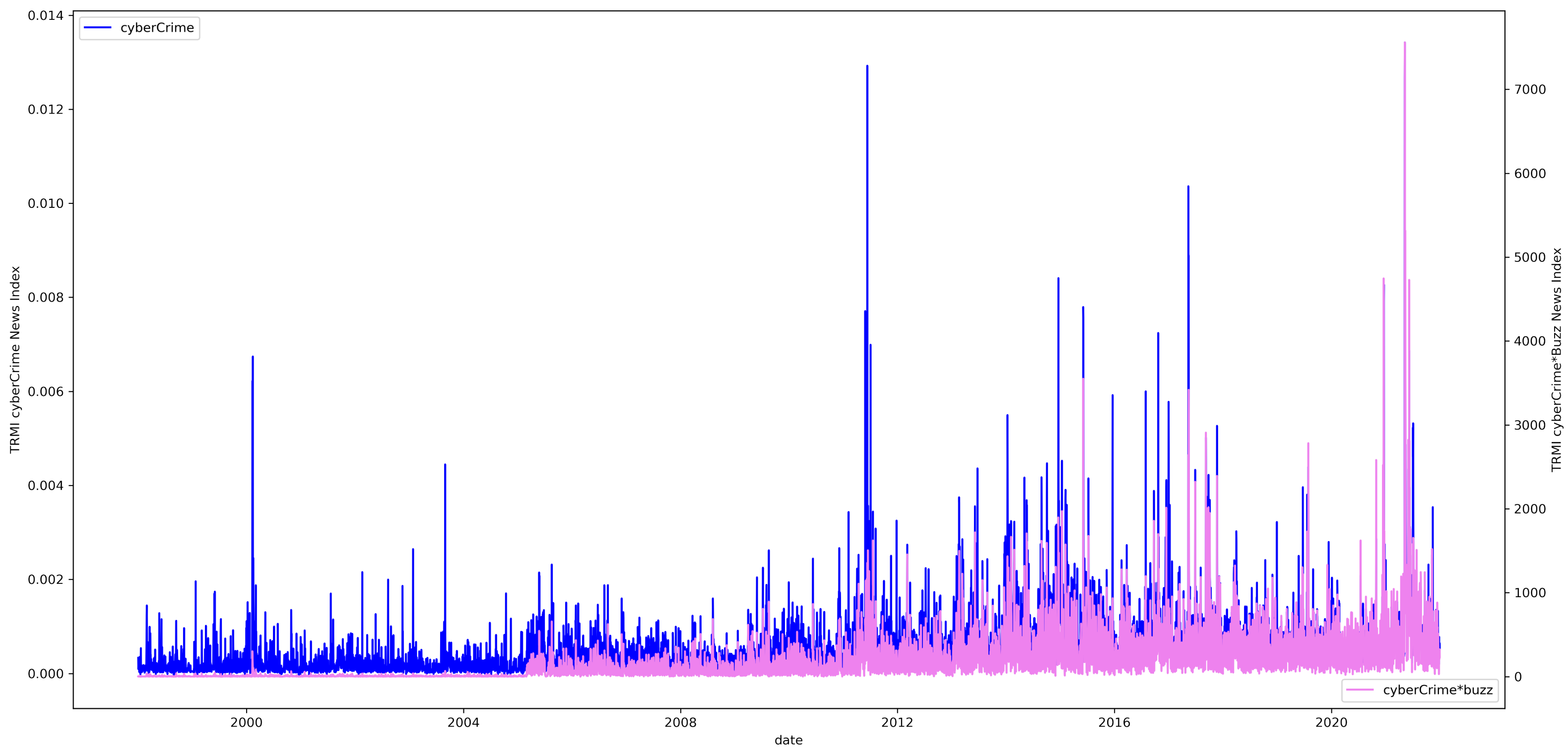

Figure 2-Panel A: Five Years Cumulative Sum of ΔCCB vs. 3, 12, and 24 Months Ahead Growth of Industrial Production Index

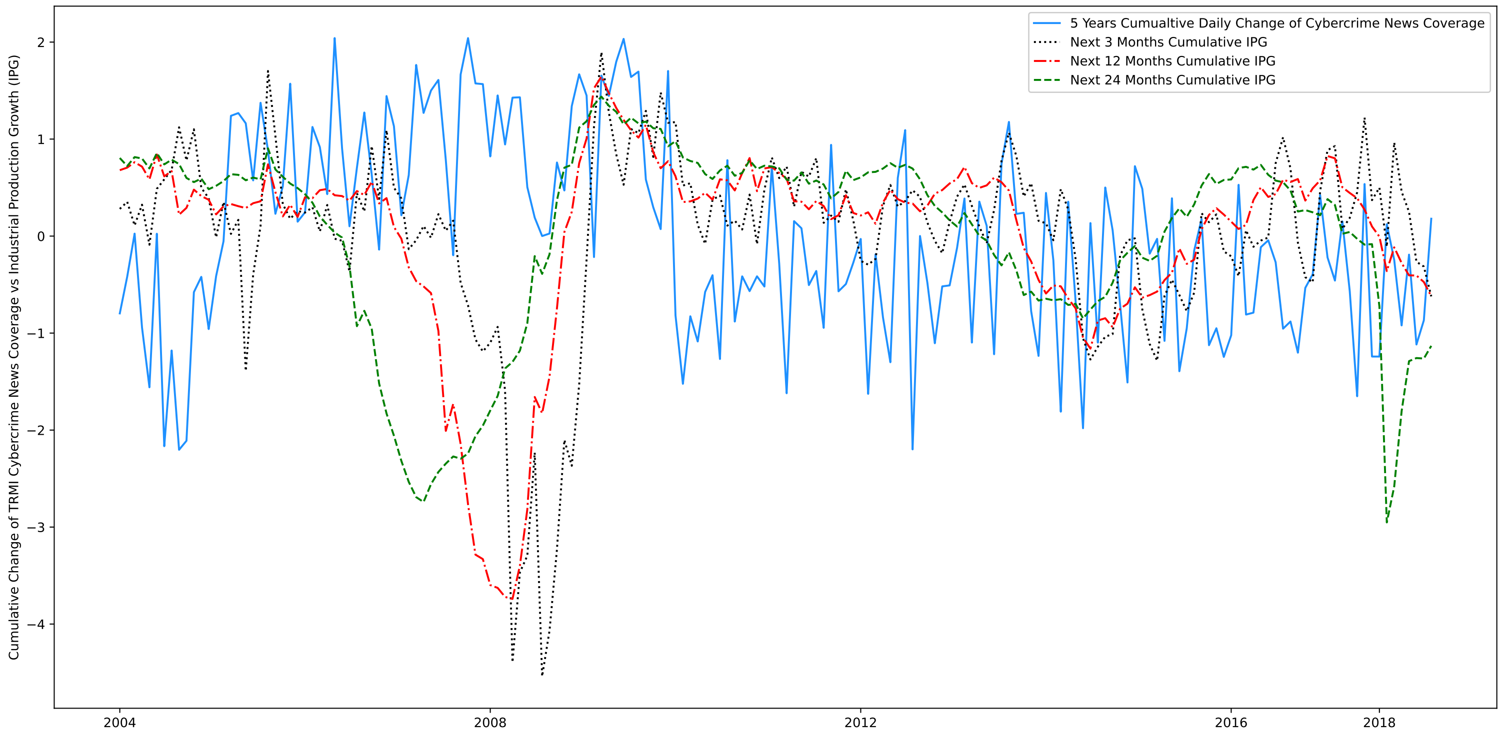

Figure 2-Panel B: Five Years Cumulative Sum of ΔCCB vs. 3, 12, and 24 Months Ahead Change of Chicago Fed Economic Activity

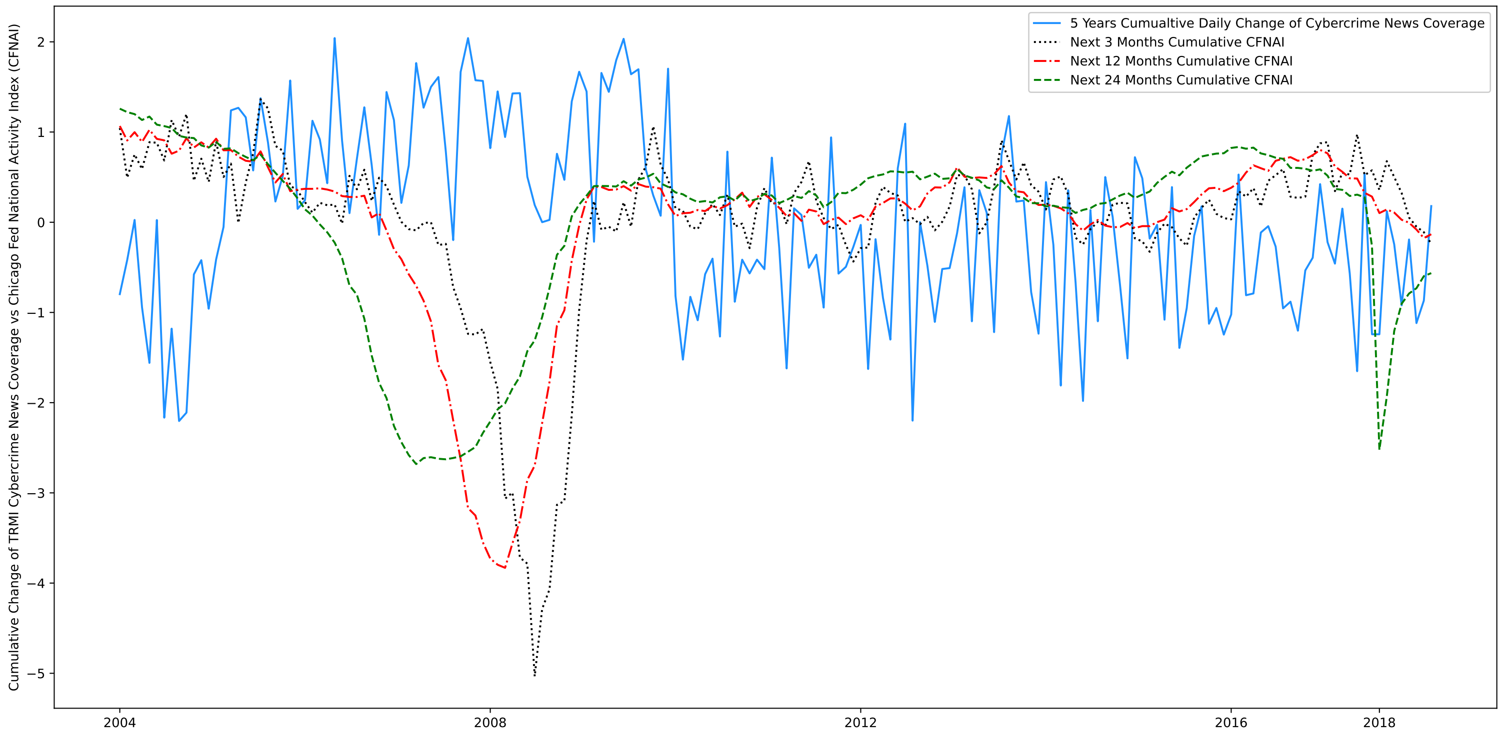

Figure 2-Panel C: Five Years Cumulative Sum of ΔCCB vs. 3, 12, and 24 Months Ahead Excess Market Returns

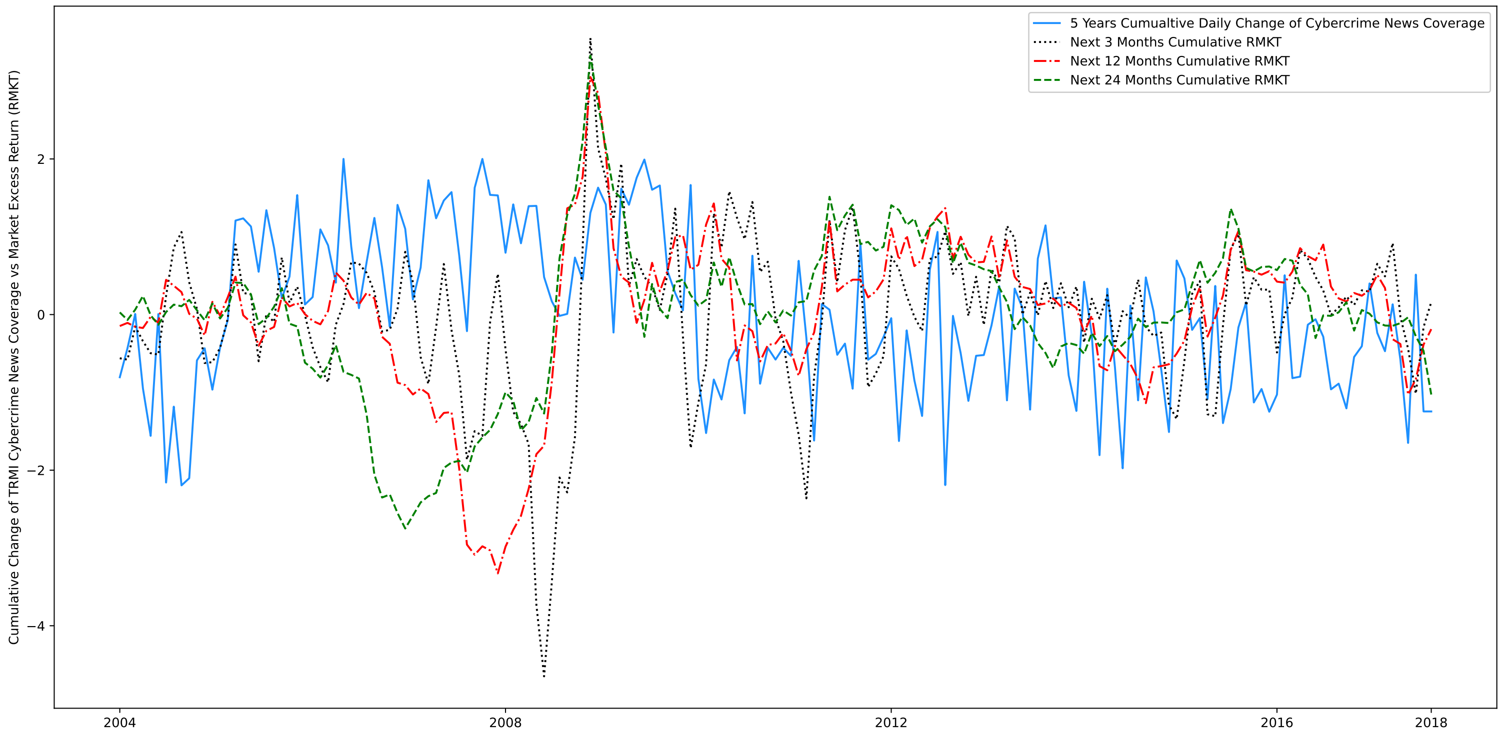

Figure 2-Panel D: Five Years Cumulative Sum of ΔCCB vs. 3, 12, and 24 Months Ahead Market Volatility

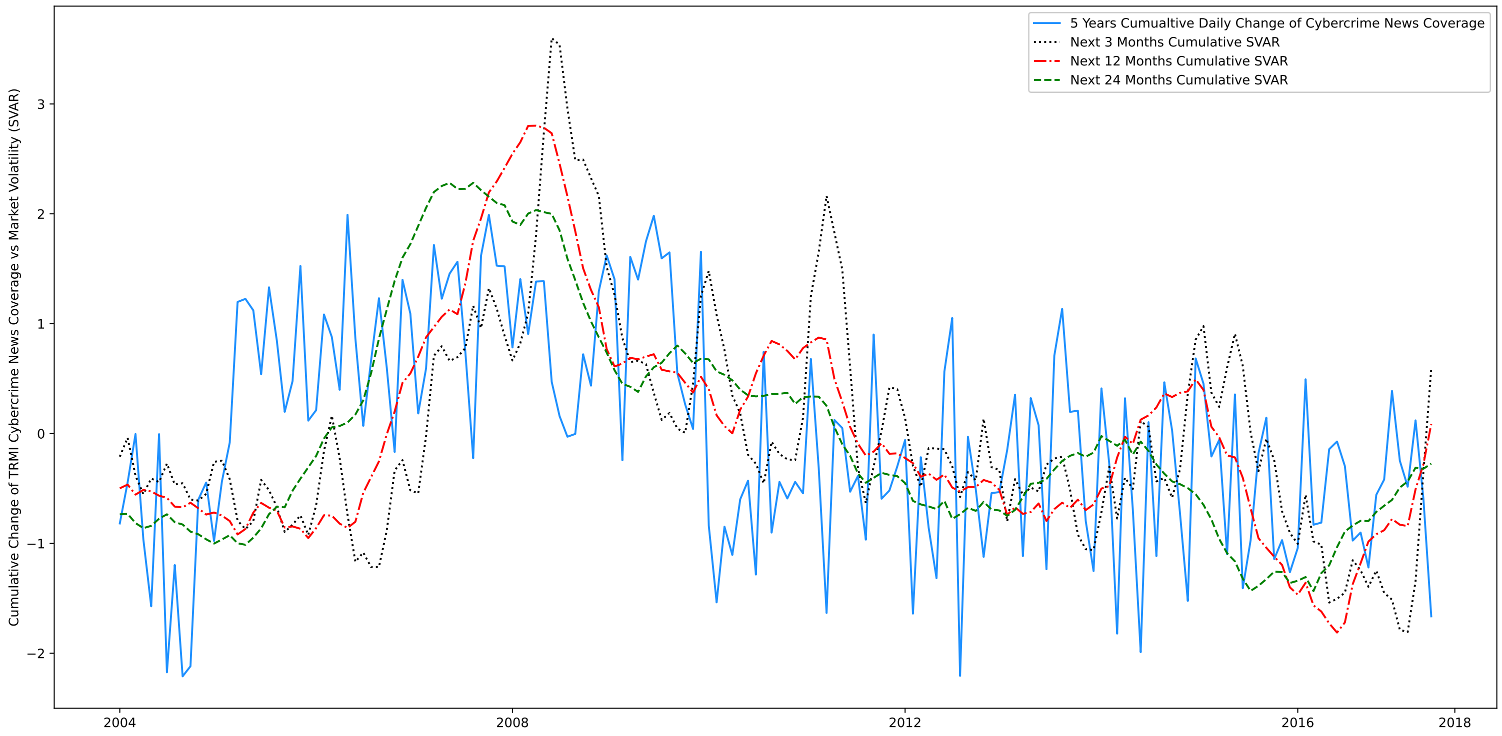

Figure 3-Panel A: Pearson Correlation between CCA and TCCA in each Rolling Sample Period

Figure 3-Panel B: Time-Varying Weights for the Base Assets formed by Cybercrime Beta

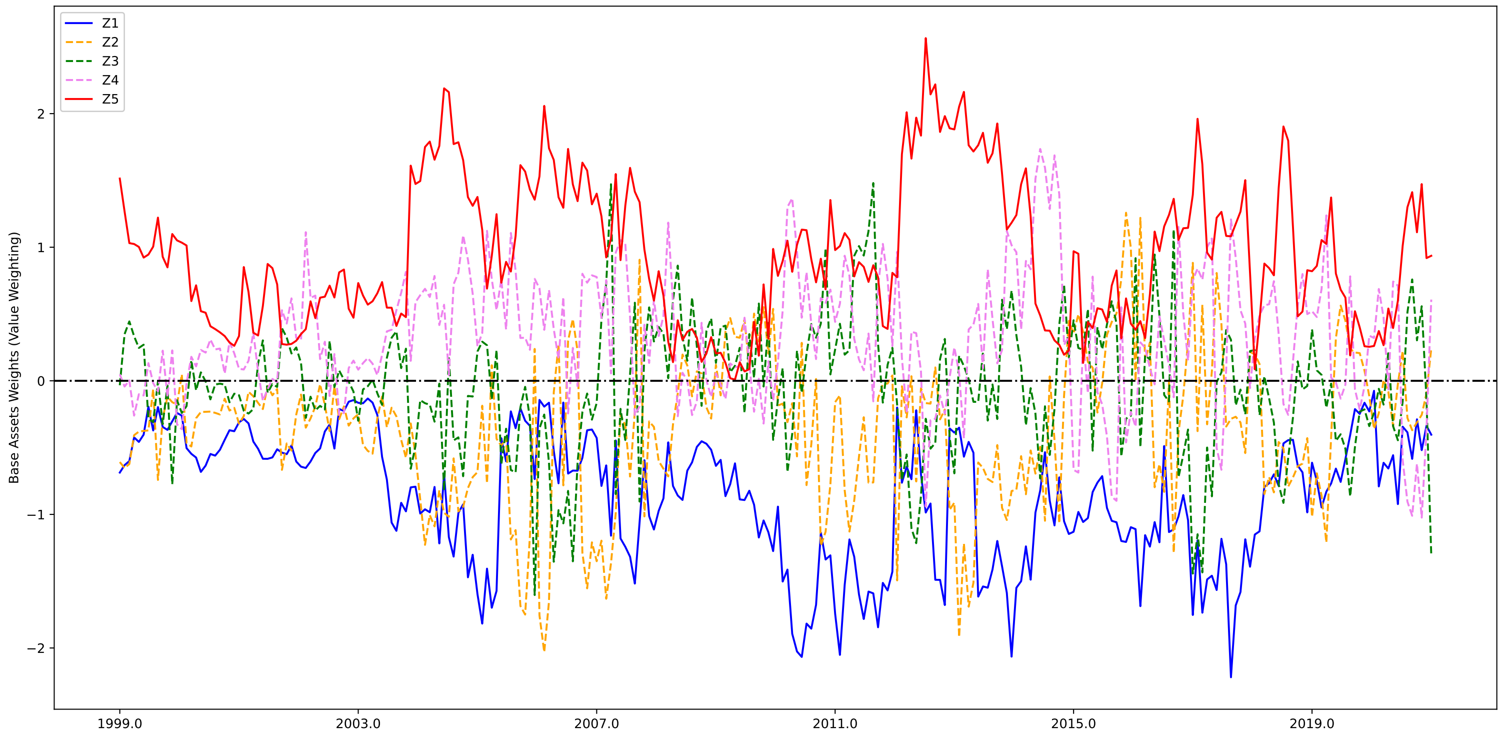

Table1: State Variable Validation for ΔCCB

This table reports the state variable validation test results for cybercrime news coverage. ΔCCB is calculated as the cumulative sum of daily change of cybercrime news coverage in a past five-year rolling window with monthly updates. Pearson correlations between ΔCCB and proxies of future economic activity (IPG and $CFNAI$) and investment opportunity set (r_{MKT} and $SVAR$) at horizons 1, 3, 12, 24, 36, and 48 months ahead are presented in Panel A. Results for multiple long-horizon regressions are reported in Panel B. The other well-known state variables are added into the model as control. TS denotes the term spread measured by the difference between the ten-year bond yield and the risk-free rate. DS denotes the default spread measured by the difference between BAA and AAA corporate bond yield. DP denotes for the dividend-to-price ratio that is calculated by taking the log difference between dividend and price from the market portfolio. $infl$ is the inflation that is the Consumer Price Index. The original sample is from 1998:01 to 2021:12, and q observations are lost in each of the respective q-horizon regressions. R^2 is the adjusted R-squared. For each regression, the estimated coefficients are reported in line 1, and Newey-West standard t-statistics are reported in parentheses computed with $q-1$ lags. The unit of coefficients for the tests of IPG and R_{MKT} is 100 basis points for easy display. $*$, $**$ and $***$ denote statistical significance at the 10%, 5% and 1% levels, respectively.

Panel A Pearson Correlation between ΔCCB and Future Economic Activity and Investment Opportunity Set

	IPG	$CFNAI$	r_{MKT}	$SVAR$
$q=1$	-0.12	-0.20***	0.01	0.22***
$q=3$	-0.14	-0.25***	-0.07	0.25***
$q=12$	-0.20***	-0.37***	-0.24***	0.39***
$q=24$	-0.23***	-0.41***	-0.27***	0.51***
$q=36$	-0.23***	-0.49***	-0.34***	0.60***
$q=48$	-0.14*	-0.49***	-0.21***	0.58***

Panel B Multiple predictive regressions for validation of ΔCCB as ICAPM State Variable

Forecasting Future Economic Activity

	IPG						$CFNAI$					
	ΔCCB	TS	DS	DP	$infl$	R^2	ΔCCB	TS	DS	DP	$infl$	R^2
$q=1$	-0.04	0.11	-0.4	0.17	0.09	0.23	-0.01	-0.09	-0.52	-0.24	0.08	0.57
	(-1.26)	(1.28)	(1.23)	(1.23)	(1.51)		(-0.22)	(-1.31)	(-3.91)	(-1.73)	(1.13)	
$q=3$	-0.1	0.11	-0.74	0.06	0.22	0.29	-0.0622	-0.1597	-0.3511	-0.30	0.08	0.44
	(-0.92)	(0.40)	(-2.92)	(0.15)	(1.17)		(-0.91)	(-1.2)	(-2.09)	(-1.38)	(0.77)	
$q=12$	-0.93	0.6	-1.17	0.99	-0.01	0.03	-0.26	-0.07	-0.30	-0.01	-0.04	0.2
	(-1.35)	(0.58)	(-1.28)	(0.60)	(-0.02)		(-1.61)	(-0.51)	(-1.85)	(-0.03)	(-0.31)	
$q=24$	-2.2	1.13	-0.63	3	0.32	0.10	-0.38	-0.06	-0.26	0.27	0.00	0.18
	(-1.88)	(0.78)	(-0.35)	(1.15)	(0.82)		(-1.98)	(-0.42)	(-1.33)	(0.90)	(0.04)	
$q=36$	-2.62	1.08	1.3	2.17	0.62	0.20	-0.46	-0.18	-0.08	0.23	0.06	0.30
	(-2.43)	(0.68)	(0.63)	(0.84)	(2.30)		(-2.57)	(-1.15)	(-0.37)	(0.83)	(1.60)	
$q=48$	-2.13	2.19	-0.68	4.87	0.16	0.25	-0.43	-0.20	-0.32	0.49	0.01	0.38
	(-2.22)	(1.97)	(-0.47)	(2.36)	(0.44)		(-2.45)	(-2.04)	(-1.9)	(2.19)	(0.22)	

Forecasting Future Investment Opportunity

	r_{MKT}						$SVAR$					
	ΔCCB	TS	DS	DP	$infl$	R^2	ΔCCB	TS	DS	DP	$infl$	R^2
$q=1$	0.13	-0.25	-1.39	1.73	0.3	0.03	-0.01	0.39	0.31	0.37	-0.07	0.43
	(0.42)	(-0.42)	(-2.03)	(1.83)	(0.91)		(-0.11)	(3.81)	(2.79)	(2.48)	(-1.23)	
$q=3$	-0.54	-0.29	-2.71	4.48	0.83	0.07	0.02	1.04	1.12	0.60	-0.10	0.44
	(-0.92)	(-0.19)	(-2.10)	(1.94)	(0.86)		(0.09)	(2.91)	(3.21)	(1.17)	(-0.55)	
$q=12$	-4.54	0.16	-2.23	10.15	-1.15	0.26	1.65	3.37	3.21	0.49	0.22	0.35
	(-2.06)	(0.06)	(-0.97)	(2.97)	(-0.77)		(1.43)	(2.16)	(2.19)	(0.23)	(0.29)	
$q=24$	-7.35	-6.47	6.00	9.3	1.74	0.49	4.66	8.48	1.51	2.30	-0.56	0.42
	(-2.59)	(-1.79)	(1.53)	(1.41)	(1.53)		(1.92)	(3.44)	(0.65)	(0.53)	(-0.74)	
$q=36$	-9.92	-7.19	2.19	14.65	2.38	0.49	7.57	11.33	2.62	1.00	-1.00	0.54
	(-2.92)	(-1.69)	(0.45)	(2.07)	(2.48)		(2.64)	(3.64)	(0.94)	(0.20)	(-1.72)	
$q=48$	-7.99	-7.27	-10.79	35.09	0.87	0.56	8.52	11.30	7.60	-7.17	-0.41	0.57
	(-3.49)	(-1.38)	(-2.85)	(7.55)	(0.78)		(3.00)	(4.94)	(1.89)	(-1.60)	(-0.54)	

Table 2: AR(p) coefficients of TRMI cyberCrime news measure

This table reports the results from equation (5), an AR(6) model. The left panel presents the autocorrelation coefficients subject to 6 lags. The right panel presents results by adding ΔVIX and ΔEPU as additional control variables. Dicker-Fuller test statistics and innovation AR(1) coefficients are reported for the AR(6) model. The original sample is from 1998:01 to 2021:12, and observations are lost in the right panel subject to the data availability of VIX and EPU. For each regression, the estimated coefficients are reported in line 1, and Newey-West t-statistics are reported in parentheses and computed with 6 lags. N is the number of observations in each regression and $\bar{R}$ is the Adjusted R-squared.

TRMI	CCB_t	CCB_t
AR(1)	0.68	0.63
	(14.05)	(11.69)
AR(2)	-0.09	0.06
	(-2.52)	(1.14)
AR(3)	0.06	0.09
	(1.65)	(1.78)
AR(4)	0.01	0.07
	(0.29)	(0.98)
AR(5)	-0.03	0.11
	(-1.05)	(2.61)
AR(6)	0.23	0.11
	(6.65)	(4.24)
ΔVIX		0.93
		(0.66)
ΔEPU		-0.01
		(-0.26)
DF	-10.35	
Innovation AR(1)	-0.03	
N	8300.00	7206.00
$\bar{R}$	0.58	0.67

Table 3: Univariate Portfolios of Stocks Sorted by β_{CCA} or β_{TCCA}

This table reports univariate portfolio sorting based on the β_{CCA} and β_{TCCA} in the left and right panels, respectively. First, for each month from December 1998, we form quintile portfolios every month by using NYSE breakpoints, β_{CCA} is estimated from equation (7), and β_{TCCA} is estimated from equation (10), using the last 12 months daily data. Second, we calculate the value-weighted returns for the next month. The first column in each panel reports individual stocks' average cybercrime beta and average cybercrime tracking beta in each relative beta quintile. The remaining columns in each panel present the average excess returns (RET-RF) and risk-adjusted returns (α_3,α_5,α_6, and α_8) for the quintile value-weighted portfolios and the high minus low portfolio in the last row. α_3 is estimated from Fama and French (1993)) three-factor model; α_5 is estimated from Fama and French (2015) five-factor model; α_6 is estimated from Fama and French (2015) five-factor model augmented with the momentum factor; α_8 is estimated from Fama and French (2015) five-factor model augmented with the momentum, short-term and long-term reversal factor. Newey-West adjusted t-statistics are reported in parentheses. The sample period is from 01/01/1998 to 12/31/2021.

		CCA Beta						TCCA Beta				
	β^{CCA}	Excess Return	α_3	α_5	α_6	α_8	β^{TCCA}	Excess Return	α_3	α_5	α_6	α_8
Low	-0.23	0.94	0.23	0.29	0.29	0.27	-0.66	1.08	0.41	0.40	0.39	0.39
		(3.06)	(2.19)	(2.50)	(2.56)	(2.28)		(3.70)	(2.80)	(2.85)	(2.85)	(2.75)
2	-0.06	0.77	0.19	0.08	0.08	0.07	-0.14	0.83	0.25	0.14	0.14	0.11
		(2.90)	(2.69)	(1.27)	(1.23)	(1.14)		(3.19)	(2.64)	(1.72)	(1.63)	(1.38)
3	0.001	0.64	0.04	-0.04	-0.04	-0.04	0.09	0.65	0.06	-0.06	-0.05	-0.07
		(2.51)	(0.58)	(-0.62)	(-0.55)	(-0.55)		(2.35)	(0.81)	(-0.76)	(-0.68)	(-0.96)
4	0.06	0.52	-0.09	-0.14	-0.13	-0.14	0.30	0.69	0.04	-0.04	-0.03	-0.03
		(1.85)	(-1.41)	(-1.92)	(-1.75)	(-1.85)		(2.37)	(0.42)	(-0.46)	(-0.30)	(-0.27)
High	0.23	0.34	-0.41	-0.36	-0.34	-0.35	0.81	0.14	-0.69	-0.53	-0.50	-0.48
		(0.87)	(-3.01)	(-2.52)	(-2.49)	(-2.37)		(0.32)	(-3.49)	(-2.92)	(-2.94)	(-2.64)
High-Low		-0.60	-0.64	-0.64	-0.63	-0.62		-0.95	-1.09	-0.92	-0.88	-0.86
		(-2.66)	(-3.01)	(-2.83)	(-2.90)	(-2.66)		(-2.93)	(-3.44)	(-3.12)	(-3.17)	(-2.90)

Table 4: *CCA* Triggering Trading Demand and β_{CCA}

This table reports the cybercrime news-driven to-buy order test results. First, for each β_{CCA} estimation sample, we estimate the sensitivity of buy orders to cybercrime news attention (γ_{CCA}) from equation (6). Three measures of investor buy orders. First, the buy fraction (BF) is calculated as the total buy trading volume over the total volume. Second, the buy and sell spread (BS) is the difference between the buy and sell trading volumes. Lastly, the number of buy orders (BV) is the logarithm of the total buy trading volume. Panel A reports the average γ_{CCA} in each quintile portfolio sorted by β_{CCA}. Fama-Macbeth cross-sectional regressions results are presented in Panel B. The control variables include firm size (SIZE) measured by market capitalization in millions of dollars, book-to-market ratio (BM), operating profitability (OP), investment (I/A), market beta (β_{MKT}), momentum (MOM), last month return (LRET), illiquidity (ILLIQ), idiosyncratic volatility (IVOL). For easy display, the unit in Panel B is percentage points. Newey-West t-statistics (Panel A) and standard errors (Panel B) are in given parentheses. $*$, $**$, and $***$ denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Panel A: Average γ_{CCA} Respect to Different β_{CCA} Stocks						
	P1	P2	P3	P4	P5	HL
BF	-0.22	-0.01	0.00	0.11	0.27	0.50
	(-9.19)	(-4.55)	(0.15)	(6.22)	(11.78)	(1.47)
BS	-6.68	-4.99	0.17	3.35	7.23	13.98
	(-8.65)	(-3.81)	(0.15)	(3.86)	(6.18)	(8.78)
BV	0.90	0.78	0.81	0.88	1.27	0.37
	(8.98)	(7.81)	(7.74)	(8.90)	(10.52)	(6.08)
Panel B: Average Cross-Sectional Relationship between β_{CCA} and γ_{CCA}						
	(1)	(2)	(3)	(4)	(5)	(6)
	BF	BF	BS	BS	BV	BV
β_{CCA}	0.1***	0.1***	20.06***	22.02***	0.82***	0.95***
	(0.05)	(0.05)	(2.39)	(2.93)	(0.14)	(0.13)
Size		0.01		-3.82		0.07*
		(0.01)		(3.16)		(0.04)
BM		-0.01		-1.02		-0.04
		(0.01)		(0.9)		(0.13)
I/A		-0.01**		-0.69		-0.02
		(0.00)		(0.43)		(0.02)
OP		-0.00		-0.03		0.02
		(0.00)		(0.15)		(0.02)
β_{MKT}		0.01*		-1.17		0.16*
		(0.01)		(1.56)		(0.08)
MOM		-0.01		-0.62		0.49***
		(0.01)		(1.20)		(0.18)
LRET		0.00		0.28		0.19*
		(0.02)		(2.74)		(0.01)
IVOL		0.00		-0.05		0.00
		(0.00)		(0.04)		(0.00)
ILLIQ		0.01		-1.79		-0.1***
		(0.01)		(1.38)		(0.03)
Intercept	0.02	0.05	-0.09	12.35	0.96	-0.8
	(0.02)	(0.05)	(0.93)	(9.97)	(0.1)	(0.1)
R-Squared	3	4	4	3	1	13
Obs	276	276	276	276	276	276

Table 5: β_{TCAA} Portfolio Returns Controlled by Industries Divided by Demand for Cybersecurity

Stocks are divided into 12 industries based on the four-digit SIC code. We group industries into those that demand less cybersecurity and those with high cybersecurity demand based on the World Economic Forum (2016) report and Cybersecurity.org that, analyze the importance of each industry's demand for cybersecurity. The group of firms with lower demand for cybersecurity includes industries from consumer durables, manufacturing, shops, utilities, and others. The group of cybersecurity demanding firms includes industries from consumer nondurables, energy, high-tech, telecommunication, healthcare, and finance. For each month, stocks are sorted into each cybersecurity dependence group. Then, within each group, we sort stocks into quintile portfolios based on β_{TCCA}, where quintile 1 (5) contains stocks with the lowest (highest) β_{TCCA} from the previous month. The results for the high minus low portfolio are reported in the last row. α_5 is estimated from Fama and French (2015) five-factor model; α_6 is estimated from Fama and French (2015) five-factor model augmented with the momentum factor; α_8 is estimated from Fama and French (2015) five-factor model augmented with the momentum, short-term and long-term reversal factor. Newey-West adjusted t-statistics are reported in parentheses. The sample period is from 01/01/1998 to 12/31/2021.

Portfolio	Low Cybersecurity Demand			High Cybersecurity Demand		
	α_5	α_6	α_8	α_5	α_6	α_8
Low	0.43	0.42	0.42	0.50	0.49	0.48
	(2.58)	(2.55)	(2.53)	(3.28)	(3.27)	(3.02)
2	-0.02	0.00	-0.01	0.24	0.24	0.21
	(-0.19)	(0.004)	(-0.12)	(2.23)	(2.10)	(1.83)
3	-0.03	-0.01	-0.03	-0.04	-0.04	-0.07
	(-0.29)	(-0.12)	(-0.33)	(-0.32)	(-0.36)	(-0.55)
4	-0.27	-0.25	-0.27	-0.03	-0.02	-0.01
	(-2.13)	(-1.98)	(-2.09)	(-0.25)	(-0.13)	(-0.09)
High	-0.56	-0.53	-0.52	-0.44	-0.41	-0.40
	(-2.75)	(-2.73)	(-2.64)	(-1.91)	(-1.87)	(-1.73)
High-Low	-0.97	-0.95	-0.94	-0.64	-0.89	-0.88
	(-3.20)	(-3.20)	(-3.16)	(-2.71)	(-2.75)	(-2.50)

Table 6: β_{TCCA} and Average Stock Characteristics

This table reports the time-series average of slop coefficients from Fama-Macbeth cross-sectional regressions of β_{TCCA} on stock-level characteristics. The slope coefficients for each monthly cross-sectional regression are estimated from equation (11). The stock-level characteristic variables include the market beta (β_{MKT}), the market volatility beta (β_{VIX}), the economic policy uncertainty beta (β_{EPU}), the firm size measured by the logarithm of market capitalization, book-to-market ratio (BM), operating profitability (OP), investment (I/A), last month return (LRET), illiquidity (ILLIQ), idiosyncratic volatility (IVOL), momentum (MOM) and analyst forecast dispersion (DISP). The cross-sectional regressions are conducted monthly from January 1999 to December 2021. Newey-West standard errors are reported in parentheses. $*$,$**$, and $***$ denote statistical significance at the 10%, 5%, and 1% levels, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
β_{MKT}	0.18^{**}												0.22^{**}
	(0.08)												(0.09)
β_{VIX}		0.01											-0.002
		(0.01)											(0.009)
β_{EPU}			0.003										-0.002
			(0.09)										(0.08)
Size				-0.012^{**}									0.006
				(0.006)									(0.01)
BM					0.07^{**}								0.03^{*}
					(0.03)								(0.02)
OP						0.004							0.006
						(0.003)							(0.004)
I/A							0.01						0.003
							(0.01)						(0.008)
LRET								-0.03					-0.01
								(0.05)					(0.04)
ILLIQ									0.005^{**}				0.02^{**}
									(0.003)				(0.008)
IVOL										0.001			-0.001
										(0.001)			(0.001)
MOM											-0.11^{***}		-0.07^{***}
											(0.04)		(0.02)
DISP												0.31	0.02
												(0.23)	(0.04)
Intercept	-0.10^{*}	0.09^{**}	0.09^{**}	0.19^{**}	0.06	0.10^{**}	0.10^{*}	0.09^{**}	0.15^{**}	0.06	0.10^{*}	0.10^{**}	-0.004
	(0.05)	(0.05)	(0.04)	(0.09)	(0.04)	(0.05)	(0.05)	(0.05)	(0.07)	(0.05)	(0.05)	(0.05)	(0.08)

Table 7: α^8 in Bivariate Portfolio Sorts

This table presents bivariate portfolio sorting results. First, we sort stocks based on each control variable into quintiles. Second, stocks within each control variable are sorted into quintiles based on β_{TCCA}. This table reports the next month's value-weighted portfolio return alphas (α^8) estimated by the Fama and French (2015) five-factor model augmented with momentum and short and long-term reversal factors for each β_{TCCA} quintile, averaged across the five control groups. The control variables include firm size (SIZE) measured by market capitalization in millions of dollars, book-to-market ratio (BM), operating profitability (OP), investment (I/A), market beta (β_{MKT}), market volatility beta (β_{VIX}), economic policy uncertainty beta (β_{EPU}), momentum (MOM), last month return (LRET), illiquidity (ILLIQ), idiosyncratic volatility (IVOL), and analyst forecast dispersion (DISP). The differences in α^8between quintile 5 (High) and quintile 1 (Low) are presented in the last row. Newey-West adjusted t-statistics are given in parentheses.

β^{TCCA}	SIZE	BM	OP	I/A	β_{MKT}	β_{VIX}	β_{EPU}	MOM	LRET	ILLIQ	IVOL	DISP
Low	0.31	0.36	0.49	0.43	0.31	0.44	0.34	0.37	0.38	0.33	0.36	0.42
	(1.90)	(2.25)	(2.53)	(2.45)	(2.09)	(2.75)	(2.05)	(2.75)	(2.38)	(2.03)	(2.11)	(2.63)
P_2	0.05	0.12	0.05	0.10	0.10	0.05	0.08	0.16	0.04	0.03	0.04	0.07
	(0.68)	(1.51)	(0.51)	(1.22)	(0.98)	(0.58)	(1.04)	(1.85)	(0.54)	(0.43)	(0.45)	(0.88)
P_3	-0.04	0.05	-0.04	-0.04	0.06	-0.06	-0.05	-0.08	-0.02	-0.02	-0.03	-0.15
	(-0.51)	(0.57)	(-0.52)	(-0.46)	(0.71)	(-0.84)	(-0.65)	(-0.98)	(-0.19)	(-0.26)	(-0.29)	(-1.83)
P_4	-0.12	-0.10	-0.11	-0.11	-0.22	-0.15	-0.10	-0.14	-0.15	-0.14	-0.17	-0.17
	(-1.66)	(-1.15)	(-1.16)	(-1.20)	(-2.59)	(-1.67)	(-1.03)	(-1.43)	(-1.68)	(-1.82)	(-1.70)	(-1.65)
High	-0.36	-0.45	-0.45	-0.37	-0.44	-0.44	-0.44	-0.34	-0.41	-0.40	-0.37	-0.40
	(-2.39)	(-2.62)	(-2.74)	(-2.20)	(-2.68)	(-2.60)	(-2.74)	(-2.22)	(-2.52)	(-2.49)	(-2.07)	(-2.39)
High-Low	-0.67	-0.81	-0.93	-0.80	-0.75	-0.87	-0.78	-0.70	-0.79	-0.72	-0.73	-0.82
	(-2.33)	(-2.65)	(-2.84)	(-2.51)	(-2.89)	(-2.93)	(-2.60)	(-2.77)	(-2.70)	(-2.42)	(-2.35)	(-2.70)

Table 8: Stock-Level Fama-MacBeth Cross-Sectional Regressions on β_{TCCA}

This table reports the time-series averages of the slope coefficients from regressing stock one month ahead of excess returns (in percentage) on the cybercrime tracking beta (β_{TCCA}) and a set of control variables with return predictability using Fama-Macbeth cross-sectional regressions. The left panel shows results without industry control, and the right panel reports results with further industry control. The control variables include firm size (SIZE) measured by market capitalization in millions of dollars, book-to-market ratio (BM), operating profitability (OP), investment (I/A), market beta (β_{MKT}), market volatility beta (β_{VIX}), economic policy uncertainty beta (β_{EPU}), momentum (MOM), last month return (LRET), illiquidity (ILLIQ), idiosyncratic volatility (IVOL), and analyst forecast dispersion (DISP). Since illiquidity and size may be strongly correlated, the second-column and third-column report results for controlling these two variables separately. Newey-West adjusted standard errors are given in parentheses. $***$: $p < 0.01$, $**$: $p < 0.05$, $*$: $p < 0.1$

	Without Inudustry Control			With Industry Control		
	$Re_{i,t+1}$	$Re_{i,t+1}$	$Re_{i,t+1}$	$Re_{i,t+1}$	$Re_{i,t+1}$	$Re_{i,t+1}$
β_{TCCA}	-0.46^{**}	-0.29^{**}	-0.30^{**}	-0.44^{**}	-0.27^{**}	-0.28^{**}
	(-0.18)	(-0.13)	(-0.13)	(-0.17)	(-0.13)	(-0.13)
β_{MKT}		0.18	0.15		0.17	0.15
		(-0.23)	(-0.23)		(-0.23)	(-0.23)
β_{VIX}		-0.13^{**}	-0.13^{**}		-0.13^{**}	-0.13^{**}
		(-0.06)	(-0.06)		(-0.06)	-0.06
β_{EPU}		-0.13	-0.13		-0.12	-0.12
		(-0.20)	(-0.20)		(-0.20)	(-0.19)
Size		-0.09^{*}			-0.09^{*}	
		(-0.05)			(-0.05)	
BM		-0.04	-0.04		-0.04	-0.03
		(-0.14)	(-0.14)		(-0.14)	(-0.14)
OP		0.14^{**}	0.14^{**}		0.14^{**}	0.14^{**}
		(-0.07)	(-0.07)		(-0.07)	(-0.07)
I/A		-0.17^{***}	-0.17^{***}		-0.17^{***}	-0.16^{**}
		(-0.06)	(-0.06)		(-0.06)	(-0.06)
LRET		-1.27^{***}	-1.33^{***}		-1.28^{***}	-1.35^{***}
		(-0.45)	(-0.45)		(-0.44)	(-0.44)
IVOL		-0.01	-0.01		-0.01	-0.01
		(-0.01)	(-0.01)		(-0.01)	(-0.01)
MOM		0.01	-0.01		0.01	-0.01
		(-0.23)	(-0.22)		(-0.22)	(-0.22)
DISP		-1.33^{**}	-1.36^{**}		-1.31^{**}	-1.33^{**}
		(-0.65)	(-0.65)		(-0.65)	(-0.65)
ILLIQ			0.04			0.05
			(-0.03)			(-0.03)
Intercept	0.76^{**}	1.36^{***}	1.12^{**}	0.92^{**}	1.48^{***}	1.25^{**}
	(-0.34)	(-0.47)	(-0.50)	(-0.40)	(-0.46)	(-0.49)
R-squared	0.02	0.09	0.09	0.02	0.10	0.10
Obs	276	276	276	276	276	276

Table 9: β_{TCCA} Estimated by 349 Portfolios with FF5 model

Thie table reports univariate portfolio sorting based on the β_{TCCA} that is estimated with 349 equity portfolios. First, for each of the 49 industry portfolios and 100 portfolios (10$\times$ 10 bivariate) formed on size and book-to-market, size and investment, and size and profitability, we estimate the cybercrime tracking beta by using the ex-ante tracking factor ($TCCA$) with daily data using equation (11). Second, we form quintile portfolios from January 1999 to December 2021. The first column reports the equity portfolio's average cybercrime new tracking beta in each relative beta quintile. The remaining columns in each panel present the average portfolio excess returns (RET-RF) and risk-adjusted returns (α_5,α_6, and α_8) for the quintile value-weighted portfolios and the high minus low portfolio in the last row. α_5 is estimated from Fama and French (2015) five-factor model; α_6 is estimated from Fama and French (2015) five-factor model augmented with the momentum factor; α_8 is estimated from Fama and French (2015) five-factor model augmented with the momentum, short-term and long-term reversal factor. α_q is estimated from Hou et al. (2015) q-factor model. Newey-West adjusted t-statistics are reported in parentheses.

	Beta TCAA	Excess Return	α_5	α_6	α_8	α_q
Low	-0.24	0.73	0.09	0.08	0.07	0.03
		(2.34)	(0.97)	(0.88)	(0.74)	(0.35)
P_2	-0.07	0.65	-0.03	-0.02	-0.03	-0.05
		(2.32)	(-0.51)	(-0.38)	(-0.55)	(-0.83)
P_3	0.008	0.67	-0.04	-0.02	-0.03	-0.02
		(2.38)	(-0.62)	(-0.41)	(-0.49)	(-0.24)
P_4	0.09	0.70	-0.02	-0.01	-0.02	0.002
		(2.46)	(-0.28)	(-0.19)	(-0.30)	(0.03)
High	0.27	0.43	-0.17	-0.17	-0.16	-0.23
		(1.41)	(-2.22)	(-2.21)	(-2.16)	(-2.65)
High-Low		-0.30	-0.26	-0.25	-0.23	-0.26
		(-2.68)	(-2.30)	(-2.22)	(-2.05)	(-2.23)

Table 10: Ex-post Cybercrime Mimicking Value-Weighted Pricing Factor

This table reports the results of the ex-post cybercrime mimicking factor pricing test. First, we estimate the weights b' by using equation (7) on a monthly rolling basis. Second, we multiply b' by the vector of one-month ahead base asset returns that are the five value-weighted portfolio returns sorted by β_{CCA} from equation (5) to obtain the cybercrime ex-post pricing factor return from January 1999 to December 2021. The first column is the average return of the ex-post cybercrime tracking factor. The remaining columns present results based on different pricing models. α^{CAPM} is eatimated from the CAPM model. α_3 is estimated from Fama and French (1993)) three-factor model; α_5 is estimated from Fama and French (2015) five-factor model; α_6 is estimated from Fama and French (2015) five-factor model augmented with the momentum factor; α_8 is estimated from Fama and French (2015) five-factor model augmented with the momentum, short-term and long-term reversal factor. Newey-West adjusted t-statistics are reported in parentheses.

	FCCA Monthly Pricing Factor Test							
	Factor	α^{CAPM}	α^3	α^5	α^6	α^8		α^q
Models	-0.44	-0.44	-0.46	-0.47	-0.46	-0.45		-0.42
	(-3.51)	(-3.47)	(-3.60)	(-3.47)	(-3.48)	(-3.30)		(-3.16)
MKT		0.01	-0.005	-0.005	-0.01	-0.03	R_{MKT}	-0.03
		(0.23)	(-0.12)	(-0.12)	(-0.23)	(-0.60)		(-0.54)
SMB			0.09	0.10	0.10	0.06	R_{ME}	0.04
			(1.81)	(1.61)	(1.64)	(0.99)		(0.72)
HML			0.09	0.07	0.07	0.02	R_{IA}	0.07
			(1.85)	(1.25)	(1.06)	(0.26)		(1.12)
RMW				0.02	0.03	0.06	R_{ROE}	-0.10
				(0.35)	(0.41)	(1.01)		(-1.80)
CMA				-0.03	-0.03	-0.09		
				(-0.34)	(-0.30)	(-0.92)		
UMD					-0.014	-0.021		
					(-0.51)	(-0.87)		
ST						0.02		
						(0.50)		
LT						0.15		
						(2.49)		
$\bar{R}^2$		-0.003	0.02	0.02	0.01	0.03		0.02

Table 11: Fama-Macbeth Risk Premium Tests on 349 Portfolios

This reports the Fama–Macbeth (1973) factor premiums on 349 portfolios, including 49 industry portfolios and 100 portfolios (10×10 bivariate) formed on size and book-to-market, size and investment, and size and profitability. We augment the baseline models that are Fama and French (2015) five-factor model and Hou et al. (2015) q-factor model with the ex-post cybercrime pricing factor ($FCCA$). We report ex-post factor loadings for $FCCA$ and other factors with perchance units. Newey-West t-statistics are given in parentheses. The sample period is from January 1999 to December 2021.

	FCCA Risk Premimum				
r_p	*FF5*			*q-Factor*	
$FCCA$		-2.82	$FCCA$		-3.95
		(-4.20)			(-4.38)
MKT	-0.99	-0.85	R_{MKT}	-4.06	-3.08
	(-2.12)	(-1.89)		(-3.99)	(-3.69)
SMB	1.00	1.00	R_{ME}	0.45	0.66
	(3.52)	(3.52)		(1.62)	(2.26)
HML	-1.24	-1.00	R_{IA}	1.31	1.54
	(-2.99)	(-2.66)		(3.49)	(3.81)
RMW	0.67	0.54	R_{ROE}	-1.36	-0.70
	(2.25)	(1.81)		(-2.26)	(-1.35)
CMA	2.01	1.82			
	4.2	4.08			
Intercept	1.44	1.25		4.29	3.37
	(6.04)	(5.80)		(5.32)	(5.40)
$\bar{R}$	0.27	0.28		0.26	0.28

Table 12: Univariate Portfolios of Stocks Sorted by Google-Search Trend Based Cybercrime Beta

This table reports univariate portfolio sorting based on the β_{SVI} and β_{TSVI} in the left and right panels, respectively. First, for each month from December 1998, we form quintile portfolios every month by using NYSE breakpoints, β_{SVI} is estimated from equation (14), and β_{TSVI} is estimated from equation (10) by replacing $TCCA$ with $TSVI$, using the last 12 months daily data. Noted that the $TSVI$ is constructed as the same procedure with $TCCA$ by using Google Search Trend data. Second, we calculate the value-weighted returns for the next month. The first column in each panel reports individual stocks' average cybercrime beta and average cybercrime tracking beta in each relative beta quintile. The remaining columns in each panel present the average excess returns (RET-RF) and risk-adjusted returns (α_3,α_5,α_6, and α_8) for the quintile value-weighted portfolios and the high minus low portfolio in the last row. α_3 is estimated from Fama and French (1993)) three-factor model; α_5 is estimated from Fama and French (2015) five-factor model; α_6 is estimated from Fama and French (2015) five-factor model augmented with the momentum factor; α_8 is estimated from Fama and French (2015) five-factor model augmented with the momentum, short-term and long-term reversal factor. Newey-West adjusted t-statistics are reported in parentheses. The sample period is from 01/01/2007 to 12/31/2021.

		SVI Beta						TSVI Beta				
	β^{CCA}	Excess Return	α_3	α_5	α_6	α_8	β^{TCCA}	Excess Return	α_3	α_5	α_6	α_8
Low	-0.58	1.14	0.10	0.15	0.15	0.15	-2.55	1.34	0.29	0.37	0.38	0.36
		(2.43)	(0.92)	(1.40)	(1.46)	(1.40)		(2.33)	(1.41)	(1.99)	(2.09)	(2.11)
2	-0.16	1.06	0.17	0.11	0.11	0.11	-0.93	1.18	0.26	0.23	0.23	0.23
		(2.80)	(2.38)	(1.55)	(1.59)	(1.52)		(2.92)	(2.56)	(2.34)	(2.50)	(2.49)
3	-0.01	0.97	0.10	0.05	0.05	0.04	-0.17	0.92	0.02	-0.00	-0.01	-0.00
		(2.79)	(1.68)	(0.89)	(0.87)	(0.90)		(2.60)	(0.19)	(-0.02)	(-0.01)	(-0.02)
4	0.15	0.83	-0.10	-0.10	-0.10	-0.11	0.55	0.89	-0.06	-0.42	-0.03	-0.03
		(2.35)	(-1.26)	(-1.31)	(-1.33)	(-1.37)		(2.51)	(-0.51)	(-0.26)	(-0.27)	(-0.24)
High	0.59	0.69	-0.49	-0.37	-0.38	-0.38	1.97	0.57	-0.5	-0.42	-0.42	-0.41
		(1.42)	(-2.88)	(-2.25)	(-2.33)	(-2.20)		(1.38)	(-3.01)	(-2.39)	(-2.55)	(-2.38)
High-Low		-0.45	-0.59	-0.53	-0.53	-0.52		-0.76	-0.79	-0.79	-0.78	-0.77
		(-2.04)	(-2.77)	(-2.41)	(-2.55)	(-2.31)		(-2.12)	(-2.45)	(-2.51)	(-2.73)	(-2.63)