4-2021

Essays on empirical asset pricing

Zilin CHEN
Singapore Management University

ESSAYS ON EMPIRICAL ASSET PRICING

ZILIN CHEN

SINGAPORE MANAGEMENT UNIVERSITY

2021

Essays on Empirical Asset Pricing

Zilin Chen

Submitted to Lee Kong Chian School of Business
in partial fulfillment of the requirements for the
Degree of Doctor of Philosophy in Finance

Dissertation Committee:

Jun TU (Supervisor/Chair)
Associate Professor of Finance
Singapore Management University

Dashan HUANG
Assistant Professor of Finance
Singapore Management University

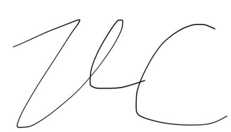
Weikai LI
Assistant Professor of Finance
Singapore Management University

Qianqian DU
Associate Professor of Finance
Southwestern University of Finance and Economics

Singapore Management University
2021

I hereby declare that this PhD dissertation is my original work
and it has been written by me in its entirety.
I have duly acknowledged all the sources of information
which have been used in this dissertation.

This PhD dissertation has also not been submitted for any degree
in any university previously.



Zilin Chen

21 April 2021

Essays on Empirical Asset Pricing

Zilin Chen

Abstract

The dissertation consists of three chapters on empirical asset pricing. The first chapter examines whether the cross-sectional variation in private subsidiaries' information disclosure predicts the cross-sectional dispersion in future equity returns of public parent firms. Information disclosure on private subsidiaries is not mandatory for public firms in the U.S., and thus these subsidiaries could be a good choice for public firms to hide bad news. We construct a private subsidiaries' information disclosure (PSID) measure and find that a value-weighted portfolio that longs stocks in the highest PSID quintile and shorts stocks in the lowest PSID quintile yields a Fama and French (2015) five-factor alpha of 0.60% per month. This return predictability is robust controlling for various firm-specific characteristics and is stronger for stocks that receive less investor attention and stocks that are costlier to arbitrage, consistent with the hypothesis that PSID information is slowly incorporated into stock prices.

The second chapter investigates whether locations of firms' economically-important public subsidiaries contain valuable information about parent firms' stocks returns. Stock returns of firms in the same headquarter state tend to move together (Pirinsky and Wang (2006)). Parsons, Sabbatucci, and Titman (2020) find that the return comovement of firms headquartered in the same state extends to a predictable lead-lag effect because investors are not able to fully process information arising from firms' peers located in the same place. We reexamine whether returns of geographic peers based on the locations of both headquarters and economically relevant subsidiaries are useful for predicting the stock returns of focal firms. We find that focal firms whose geographic peers experience higher (lower) returns in the current month will earn higher (lower) returns in the next month. A strategy exploiting this pattern is distinct from other well-known cross-firm momentum strategies, and it is more pronounced among firms that receive less

investor attention and firms that are more costly to arbitrage, consistent with slow information diffusion in the geographic network into stock prices.

The third chapter focuses on the well-known presidential puzzle, which refers to the striking empirical fact that stock market returns are much higher under Democratic presidencies than Republican ones. Since first noted by Huang (1985) and Hensel and Ziemba (1995) and carefully documented by Santa-Clara and Valkanov (2003), the pattern remains robust. It is only recently that Pastor and Veronesi (2020) provide an ingenious solution to this puzzle. In this paper, we document a different presidential puzzle in the cross-section of individual stocks. We construct a monthly Presidential Economic Approval Rating (PEAR) index from 1981 to 2019, by averaging ratings on president's handling of the economy across various national polls. In the cross-section, stocks with high betas to changes in the PEAR index significantly under-perform those with low betas by 0.9% per month in the future, on a risk adjusted basis. The low-PEAR-beta premium persists up to one year, and is present in various sub-samples (based on industries, presidential cycles, transitions, and tenures) and even in other G7 countries. It is also robust to different risk adjustment models and controls for other related return predictors. Since the PEAR index is negatively correlated with measures of aggregate risk aversion, a simple risk model would predict the low PEAR-beta stocks to earn lower (not higher) expected returns. Contrary to the sentiment-induced overpricing, the premium does not come primarily from the short leg following high sentiment periods. Instead, the premium could be driven by a novel sentiment towards presidential alignment.

Contents

Acknowledgements

First and foremost, I would like to express my special appreciation and thanks to my advisor Professor Jun Tu for the continuous support of my Ph.D. study and research. His immense knowledge and plentiful experience have encouraged me in all the time of my academic research and daily life. He has given me the freedom to pursue various research topics without objection, and also provided valuable and insightful discussions about my research questions. Without his patience and help, I would not have been able to gain so much research experience.

I am also deeply indebted to my research committee, Professor Dashan Huang, who has been guiding me and teaching me a great deal in basically every aspect of academic research. He always encourages me to keep thinking and doing, shares his research experience with me, and inspires me to pursue interesting research topics. It would not have been possible for me to finish my dissertation without his immeasurable support and guidance. I would like to express my sincere thanks to my last research committee member, Professor Weikai Li, who has been providing me valuable and insightful feedbacks. Besides, his diligence and attitude towards research also inspires me to keep working hard. I am also grateful to the faculty at SMU, including Jianfeng Hu, Hao Liang, Roger Loh, and Chishen Wei, who have offered cordial supports and valuable suggestions to my research. Their excellence and strictness deeply motivate me to purse high-quality research.

Finally, my biggest thanks to my family for all the support they have shown me through my five-year Ph.D. study. Words can not describe how thankful I am for the love and encouragement my parents have given me. With their unconditional support and their never-ending faith in me, I won't feel lost in my life.

Chapter 1

Private Subsidiaries' Information Disclosure and the Cross-Sectional Equity Returns of Public Parent Firms*

We investigate the impact of potential information hiding originated from private subsidiaries on the future returns of their public parent firms. We find a significantly positive link between private subsidiaries' information disclosure (PSID) and the cross-section of future equity returns of public parent firms. The economically and statistically significant PSID premium of 0.60% per month is not explained by established factor models and is stronger for stocks that receive less investor attention and that are costlier to arbitrage. Consistent with investor underreaction hypothesis, PSID premium reflects slow diffusion of firm information into stock prices rather than compensation for risk.

1.1 Introduction

Corporate information disclosure is one of the most efficient ways of communicating with the general public and financial regulators. Accurate, complete, and timely financial reporting and

*This is a joint work with Turan G. Bali, Jun Tu, and Ran Zhang information disclosure are essential to convey firm performance to individual and institutional investors. However, regulatory rules on mandatory information release are only applicable to public firms but not to private firms. In an effort to prevent selective disclosure by public firms to market professionals, the Securities and Exchange Commission (SEC) passed a rule, Regulation Fair Disclosure (Reg FD), in October 2000.[*] When a public parent firm wants to withhold information (e.g., Verrecchia (2001) and Kothari, Shu, and Wysocki (2009)), private subsidiaries can be a natural choice.[†] Indeed, in 2018, there are over 2,500 U.S. public firms with private subsidiaries and these public parent firms account for more than 70% of total stock market capitalization of the NYSE, Amex, and Nasdaq. The average number of private subsidiaries is 60 per public parent firm. Given such large scale of usage of private subsidiaries by publicly listed firms, we argue that it is crucial to examine the impact of potential information hiding or disclosure delay originated from private subsidiaries on the performance of their public parent firms.[‡]

In this study, we investigate whether the cross-sectional variation in private subsidiaries' information disclosure (PSID) predicts the cross-sectional dispersion in future equity returns of public parent firms. A high (low) PSID is observed when the parent firm is (not) willing to disclose information about its private subsidiaries. The underlying reason could be that its private subsidiaries have positive (negative) information to disclose so that the magnitude of PSID is positively correlated with private subsidiaries' and their parent firm's fundamentals (or operating

[*] Bailey et al. (2003) find that the quantity of information available to the public is indeed increased with the adoption of Reg FD, and a richer disclosure environment leads to better functioning of financial markets (Bailey, Karolyi, and Salva (2006)).

[†] Earlier studies find that companies have motivations to delay information disclosure, especially bad news. Verrecchia (1983) and Dye (1985) theoretically argue that managers tend to disclose favorable information in a partial-disclosure equilibrium. Patell and Wolfson (1982) examine firms' strategic disclosure behaviors when they announce intraday earnings and dividends. They find that good news appears more frequently during trading but bad news is more likely to be released when the stock markets are close. Verrecchia (2001) finds that there are large costs for firms' information disclosures since competitors may copycat proprietary information. Li (2008) finds that firms with weak performance tend to issue annual reports that are more difficult to understand. Kothari, Shu, and Wysocki (2009) find that firm managers delay in releasing bad news relative to good news due to a range of incentives, such as career concerns and compensations.

[‡] For all public firms and private firms in the market, Asker, Farre-Mensa, and Ljungqvist (2014) estimate that private firms accounted for 86.4 percent of U.S. firms with 500 or more employees, nearly 59 percent of aggregate sales, and nearly 49 percent of aggregate pre-tax profits in 2010. Based on the report of Forbes in 2013, less than 1 percent of the companies in the U.S. are publicly traded on the major exchanges. Meanwhile, private firms are not always lumped into a small business.

performance). Also, given that private subsidiaries are often to be opaque, poorly understood, and attract less investor attention, investors may underreact to the PSID. Therefore, we hypothesize that the PSID positively predicts the cross-section of future stock returns of parent firms. Our empirical results provide strong support for this hypothesis.*

We obtain financial information about private firms from Orbis – the most comprehensive database that contains standardized and comparable data on private firms and corporate ownership, covering over 310 million companies worldwide. Orbis provides seven items of financial information for each private subsidiary under a parent firm, which are operating revenue, total assets, number of employees, income before tax, net income, cash flow, and shareholders' funds.† For each public parent firm, we calculate its PSID. For example, one public parent firm has 100 private subsidiaries, 90 of which disclose operating revenue information to the public, 40 disclose total assets information to the public, and 20 disclose number of employee information to the public. If we consider only the three aforementioned financial variables, the private subsidiaries' average information disclosure is (0.9+0.4+0.2)/3 = 0.5. In our tests, we use a total of seven financial variables to calculate the PSID ratio as our key predictor. This measure is consistent with the level of disaggregation of accounting data proposed by Chen, Miao, and Shevlin (2015), which represents the information disclosure quality of firms.

We show that the PSID ratio has significant cross-sectional predictive power for public parent firms' future stock returns. At the end of June of each year, we sort public parent firms into five quintile portfolios based on the previous year's PSID ratios, and find that public parent firms with a higher (lower) PSID ratio earn higher average (lower) returns in subsequent months. Furthermore, the value-weighted arbitrage portfolio that takes a long position in 20% of the stocks with the highest PSID (quintile 5) and takes a short position in 20% of the stocks with

*In the same spirit, the literature finds that the annual report readability and transparency are positively related to the company's performance (e.g., Subramanian, Insley, and Blackwell (1993), Li (2008), and Dempsey et al. (2012)).

†Due to data limitation, the PSID proposed in this study may miss other information disclosed by private subsidiaries. For instance, a parent firm may be afraid of the copycat issue of information disclosure. Releasing the results of the progress in certain R&D programs may indicate that its private subsidiaries have made significant progress since then the release of the R&D progress will not damage the lead position of the parent firm over its competitors.

the lowest PSID (quintile 1) yields the risk-adjusted returns (alphas) of 0.60%, 0.44%, 0.52%, and 0.54% per month, estimated, respectively, with the six-factor model of Fama and French (2018), the mispricing factor model of Stambaugh and Yuan (2017), the Q-factor model of Hou, Xue, and Zhang (2015), and the behavioral factor model of Daniel, Hirshleifer, and Sun (2020). All alphas are significant at the 1% level, except for the mispricing factor model with the 5% level of significance. Moreover, the PSID ratio shows a robust, positive and statistically significant predictive power on future excess, industry-adjusted, and DGTW-adjusted (Daniel et al. (1997)) returns of their public parent firms in multivariate Fama–MacBeth regressions when we control for a number of firm characteristics and risk factors, including public parent firm's size (SIZE), book-to-market ratio (BM), gross profitability (GP), asset growth (AG), one-month lagged return (STR), medium-term price momentum (MOM), earnings surprise (SUE), Amihud (2002) illiquidity measure (ILLIQ), idiosyncratic volatility (IVOL), turnover ratio (TURNOVER), and the number of private subsidiaries under the parent firm.

If the PSID ratio provides valuable fundamental information, it is supposed to predict high future operating performance of the public parent firm. We find that the PSID ratio does indeed significantly predict the public parent firm's operating performance, such as return-on-asset (ROA), cash flows, and gross margin in the following years. The performance results from accounting fundamentals confirm our hypothesis that high PSID ratio indicates favorable fundamentals of private subsidiaries and their public parent firm, and investors are not able to promptly process and recognize this positive relation.

Hong and Stein (1999) propose a theoretical model in which gradual diffusion of private information among investors explains the observed predictability of stock returns. In their model, at least some investors can process only a subset of publicly available information because either they have limited information-processing capabilities or searching over all possible forecasting models using publicly available information itself is costly (Huberman and Regev (2001), Sims (2003), Hirshleifer and Teoh (2003)), and there are limits to arbitrage (Shleifer and Vishny (1997)). Due to investors' limited attention and costly arbitrage, new informative signals get incorporated into stock prices partially because at least some investors do not adjust their demand by recovering informative signals from firm fundamentals or observed prices. As a result of this failure on the part of some investors, stock returns exhibit predictability.

In this paper, we argue that the return predictability is concentrated in stocks that receive less investor attention and are costlier to arbitrage. Earlier studies show that less sophisticated individual investors have more limited attention (Peng and Xiong (2006)) and hence we argue that the informative signals provided by the PSID for stocks largely held by retail investors are not incorporated into prices quickly. However, sophisticated institutional investors, who are able to detect and process information generated by the PSID, can take advantage of mispricing in these stocks so that the information produced by the PSID will be promptly incorporated into equity prices. Since the information is integrated into the prices much faster in the presence of informed investors, there is little room for predictability among stocks with high institutional ownership and less costly to trade. Thus, the slow diffusion of information and the resulting return predictability should be more pronounced for stocks with low attention-grabbing characteristics and high arbitrage costs.

To provide a better understanding of the economic mechanisms behind the return predictability, we test whether the predictive power of the PSID is driven by investors' limited attention (Hirshleifer and Teoh (2003) and Hirshleifer, Lim, and Teoh (2009)) and/or limits to arbitrage (Shleifer and Vishny (1997)). We find that the abnormal returns on stocks with low attention-grabbing characteristics are indeed larger than the abnormal returns on stocks with high attention-grabbing features, where the proxies of investor attention are the residual media coverage, transient institutional ownership, and absolute SUE.[*] We also find the abnormal returns on stocks with high arbitrage costs are larger than the abnormal returns on stocks with low arbitrage costs, where the proxies for limits to arbitrage include the size-orthogonalized institutional ownership, idiosyncratic volatility, and Amihud (2002) illiquidity measure. Our results indicate that the PSID-based return

[*]See, e.g., Bushee (2001), Hou and Moskowitz (2005), Peng (2005), Peng and Xiong (2006), Hong, Torous, and Valkanov (2007), Cohen and Frazzini (2008), DellaVigna and Pollet (2009), Fang and Peress (2009), Hirshleifer, Lim, and Teoh (2009), Da, Engelberg, and Gao (2011), Hirshleifer, Hsu, and Li (2013), Da, Gurun, and Warachka (2014), Bali et al. (2014), Hirshleifer, Hsu, and Li (2018), and Bali et al. (2018).

predictability is likely due to investors' inattention and limits to arbitrage.

However, the return predictability may still be related to systematic/macroeconomic risk, even if the source of risk has not been clearly identifiable, as argued by Lee and So (2015). To investigate this possibility, we use an alternative methodology to test if the predictive power of the PSID is consistent with a gradual diffusion of information or news relevant for a firm's future cash flows instead of with a change in the discount rate or risk. We use the PSID ratio to forecast public parent firm's standardized unexpected earnings (SUEs). This test is not confounded by the possible existence of non-measurable risks. The results show that the PSID does indeed predict one-quarter-ahead SUEs of public parent firms, but the predictability disappears in the subsequent quarters. This finding provides further support that the PSID-based return predictability is not likely to be attributed to risk.

To further test whether the risk- or mispricing-based factors can explain the predictive power of the PSID, we follow Engelberg, McLean, and Pontiff (2018) and use the PSID ratio to predict the three-day cumulative abnormal returns around the earnings announcement days. If the ratio predicts the cumulative abnormal returns around the earnings announcement days, then investors' misconceptions about a firm's future performance and cash flows become an important driver of the return predictability phenomenon. Alternatively, without predictable cumulative abnormal returns around the earnings announcement days, the more likely explanation would be based on systematic risk factors driving the predictive power of the PSID. Our results show that the return spread on PSID-sorted portfolios is six (four) times higher during a one-day earnings announcement window (a three-day earnings announcement window) than on non-announcement days, indicating that the return predictability is more consistent with the mispricing explanation.

In addition, we conduct a more direct test of a potential risk-based explanation by reporting the average market beta, average total volatility, average idiosyncratic volatility, and ex-ante portfolio exposures to the standard risk factors. If stocks in the highest PSID quintile do not have higher average beta, total or idiosyncratic volatility, or if their exposures to the risk factors are not significantly higher than those in the lowest PSID quintile, it would be harder to argue for the risk-based explanation. We also calculate the beta of ex-post quintile portfolio returns with respect to the monthly and quarterly growth rate of consumption. Our results show that the stocks in the highest PSID quintile have smaller average beta, total and idiosyncratic volatility, and their exposures to most risk factors are smaller than those in the lowest PSID quintile. The stocks in the highest PSID quintile have lower exposures to the consumption growth rate (i.e., lower consumption beta) than those in the lowest PSID quintile. Overall, our results provide no evidence of a risk-based explanation for the predictive power of PSID. Instead, the positive PSID premium reflects slow diffusion of firm information into stock prices consistent with the investor underreaction hypothesis. Thus, we contribute to the literature on the effect of investor inattention on stock price dynamics by introducing a new information disclosure dimension (or disclosure delay originated from private subsidiaries) and by presenting evidence that the theory of investor inattention is important in understanding stock market underreactions to the informative signals provided by private subsidiaries of public parent firms.

The rest of the paper is organized as follows. Section 2 describes the data and variables. Section 3 presents the main empirical results on the cross-sectional return predictability. Section 4 tests whether the PSID is a significant indicator of the future operating performance of public parent firms. Section 5 investigates the sources of predictability. Section 6 distinguishes between risk and mispricing based explanations. Section 7 performs additional analyses. Section 8 concludes the paper.

1.2 Data, Variables, and Summary Statistics

Our main empirical analyses are based on the Orbis database compiled by Bureau van Dijk (BvD). Orbis covers comprehensive ownership information about 30 million ownership/subsidiaries relationships over time.* We collect data from Orbis to identify the ownership links between private subsidiaries and public parent firms listed in the U.S. from 2005 to 2018. At the end of each

*BvD collects the ownership data from a variety of sources including firms' annual reports, the SEC Edgar files, and local data providers.

year for the public firms that own private subsidiaries, we collect company name, ISIN code, ticker symbol, SIC classification, all ownership information including direct and indirect total ownership percentage, as well as all identifying and fundamental information for the subsidiaries. Since we are only interested in the pricing effect of private subsidiaries' information on parent firms, we exclude any public subsidiaries' links based on the listed or unlisted indicator.

We aim to identify the impact of private subsidiaries' information disclosure on the cross-sectional pricing of public parent firms. Claessens, Djankov, and Lang (2000) use a 20% cutoff to determine if a public firm is fully controlled by a unique ultimate owner. We instead follow the ultimate owner classification of Orbis and define the global ultimate owner as of the parent firm which holds more than 50% of the private firm's shares. The public parent firm ultimately controls its private subsidiary if the ownership percentage is larger than 50%. Thus, the information about major ownership of private subsidiaries is essential to the stock returns of the public parent firms. For each parent firm, we then retrieve information about the subsidiaries that are directly or indirectly held by the parent firm. Following Kalemli-Ozcan, Sorensen, Villegas-Sanchez, Volosovych, and Yesiltas (2015), we decode the indicators of percentage owned by parent firm into a specific value since they are not given in a numerical format.[*]

Once we identify the links between public parent firms and private subsidiaries, we use financial information from private subsidiaries' unconsolidated financial accounts to construct the average information disclosure ratio of private subsidiaries. In particular, we extract operating revenue, total assets, number of employees, income before tax (P/L before tax), net income, cash flow, and shareholders funds; seven variables in total for each private subsidiary.[†] As long as a firm updates its financial information, Orbis will promptly record the updated information from a list of reliable information sources. Since Orbis will keep the record for up to five years (Kalemli-

[*]In particular, we replace percentage with a leading “$<$”, “$>$”, or “$\pm$” with the percentage after the symbol; we eliminate possible signs that preceded percentages: “_”, “?”, or “Â”; we replace special codes “WO” (wholly owned) with 100%, “MO” (majority-owned) with 50.01%, “CQP1” (50% plus 1 share) with 50.01%, “NG” (negligible) with 0.01%, “BR” (branch) with 100%; “JO” (jointly owned) with 50%, and “-” (not significant) or “n.a.” (not available) with missing.

[†]Orbis provides only these seven financial variables for all private and public firms in the old version disks before 2012.

Ozcan et al. (2015)), it is possible that one private subsidiary didn't update its financial information promptly and we mistakenly use this stale information. Thus, we convert those stale information into missing value based on the most recent release date of the private subsidiary provided by Orbis. Next, we compute a single ratio for each financial variable defined as the number of private subsidiaries disclosing that financial variable divided by the total number of private subsidiaries under the control of a public firm. As such, we define our main variable of interest, the private information disclosure ratio (PSID), for each public parent firm as the simple average of the ratios of these seven financial variables. For each variable, we scale the raw number by the number of total private subsidiaries to control for the potential size effect since larger firms tend to have more subsidiaries. For the firms listed in the U.S., we extract the CUSIP information from ISIN code and ticker symbol to match with the CRSP monthly stock data using the CRSP names file.

Table 1.1 reports the cross-sectional characteristics for firms that own at least one private subsidiary across the years. Across the sample period, the number of firms that own private subsidiaries are around 2,500 except for the years of 2010 and 2011, implying that the number of firms owning private subsidiaries is stable over time. In terms of the market capitalization, our sample firms comprise 63% to 72% of the market capitalization of the firms listed at the NYSE, AMEX, and NASDAQ. Finally, the averaged PSID ratios are higher before 2013 given that the number of private subsidiaries is also lower in these years. Regarding the distributions of the ratios of seven key financial variables, operating revenue and the number of employees are the two items that most of the private firms choose to disclose to the investors. Specifically, approximately 36% and 33% of private subsidiaries under a public company disclose information about their operating revenue and number of employees to the public. This is not surprising since these two financial statements are the most common items that private firms choose to disclose especially if they need to raise capital from external investors even though they are not forced to disclose by the law. In contrast, it is less likely for public firms to disclose some of the accounting information about their private subsidiaries including cash flow, total assets, income before tax, net income, and shareholders funds.

We obtain monthly stock returns from the Center for Research in Security Prices (CRSP) and accounting information from Compustat. Our sample starts with all firms listed at the NYSE, AMEX, and NASDAQ. We keep common stocks and exclude financial firms and utilities firms. To reduce the effect of micro-cap firms, we exclude firms that are below the 20th percentile of NYSE market capitalization. We follow Shumway (1997) to adjust stock returns for delisting. Specifically, if a delisting return is missing and the delisting event is performance-related, we set the delisting return as -30%. Since some small firms with a few private subsidiaries may naturally have high PSID compared to firms with at least two private subsidiaries, we further restrict our sample to firms with at least five private subsidiaries in our main analysis to ensure that the return predictability is not driven by small and illiquid stocks.* To ensure that the ownership information and other accounting information are fully available to investors, we skip six months until the end of June of next year to form our portfolios. In particular, we match the ownership information and accounting information in year t-1 to the monthly returns from July of year t to June of year t+1.

In the subsequent regression analysis, we also control for other firm characteristics that have been shown to predict future returns. Specifically, SIZE is the firm's market capitalization computed as the logarithm of the market value of the firm's outstanding equity at the end of month t-1. BM is the logarithm of the firm's book value of equity divided by its market capitalization, where the BM ratio is computed following Fama and French (2008). Firms with negative book values are excluded from the analysis. Short-term reversal (STR) is the stock's lagged monthly return. MOM is the stock's cumulative return from the start of month t-12 to the end of month t-2 (skipping the STR month), following Jegadeesh and Titman (1993). Gross Profitability (GP) is the firm's gross profitability, defined as revenue minus cost of goods sold scaled by total assets, following Novy-Marx (2013). Asset Growth (AG) is a percentage of total asset growth between two consecutive fiscal years, following Cooper, Gulen, and Schill (2008). TO is the monthly turnover computed as the number of shares traded divided by the total number of shares outstanding in month t-1. ILLIQ is the monthly illiquidity measure computed as the absolute daily return

*Our results remain quantitatively and qualitatively similar if we include firms with less than five private subsidiaries in our analysis.

divided by daily dollar trading volume, averaged in month t-1, following Amihud (2002). IVOL is the idiosyncratic volatility defined as the standard deviation of daily residuals estimated from the regression of daily excess stock returns on the daily market, size, and value factors of Fama and French (1993) in month t-1, following Ang, Hodrick, Xing, and Zhang (2006). SUE is the standardized unexpected earnings defined as actual earnings in the current quarter minus earnings 4 quarters ago, scaled by stock price in the current quarter following Livnat and Mendenhall (2006).

Our final sample includes 155,591 firm-month observations spanning the period from July 2006 to December 2019. Panel A of Table 1.2 presents descriptive statistics for the main variables. The average number of private subsidiaries (#ofPriSub) owned by public firms is around 66. Concerning our main variable of interest in this paper, the PSID, we note that the mean value of this ratio is 0.2, which means that around 20% of private subsidiaries will release their key financial information in our sample. In Panel B of Table 1.2, we report the time-series averages of the monthly cross-sectional correlations between the PSID and other key characteristics. The Pearson correlations between the PSID and most of the other firm characteristics are quite low with absolute values all below 0.1, suggesting that this ratio is distinct from other well-known return predictors. The corresponding Spearman rank correlations are also very low with most of these firm characteristics. Besides, the PSID is positively correlated with future return and current return. Therefore, our proposed PSID may potentially contain valuable, independent information in predicting the cross-sectional variation in future equity returns.

1.3 Empirical Results

In this section, we test whether the PSID predicts the cross-section of future stock returns using portfolio-level and firm-level regression analyses.

1.3.1 Portfolio-level analysis

To construct the long-short portfolio, at the end of June of each year t from 2006 to 2019, individual stocks of public parent firms are sorted into quintile portfolios based on non-zero PSID at the end of year t-1 from 2005 to 2018 and are held for the next twelve months. We also assign parent firms with zero PSID into a zero group. We skip six months to form the portfolio to make sure that our results are in line with the methodology used by earlier studies. We then compute the value-weighted average excess return of each quintile portfolio and the zero-PSID portfolio over the next twelve months. To examine the cross-sectional relation between the PSID and the future stock returns of public parent firms, we form a long-short portfolio that takes a long position in the highest quintile of PSID and a short position in the lowest quintile of PSID.

In Panel A of Table 1.3, we report the average monthly returns of the zero-PSID portfolio, each quintile portfolio, and the long-short portfolio over the one-month Treasury bill rate. We also report the abnormal returns (alphas) estimated with various factor models, including the capital asset pricing model (CAPM) with the market (MKT) factor, the four-factor model (FFC) of Fama and French (1993) and Carhart (1997) with the MKT, size (SMB), book-to-market (HML), and momentum (MOM) factors, the five-factor model (FFCPS) of Fama and French (1993), Carhart (1997), and Pástor and Stambaugh (2003) with the MKT, SMB, HML, MOM, and the liquidity risk (LIQ) factors, the five-factor model (FF5) of Fama and French (2015) with the MKT, SMB, HML, investment (CMA), and profitability (RMW) factors, the six-factor model (FF6) of Fama and French (2018) the MKT, SMB, HML, CMA, RMW, and MOM factors, the q-factor model (HXZ) of Hou, Xue, and Zhang (2015) with the MKT, size (SMB_Q), investment ($R_{I/A}$), and profitability (R_{ROE}) factors, the mispricing factor model (SY) of Stambaugh and Yuan (2017) with the MKT, SMB, management (MGMT), and performance (PERF) factors, and the behavioral factor model (DHS) of Daniel, Hirshleifer, and Sun (2020) with the MKT, post-earnings-announcement drift (PEAD), and financing (FIN) factors. Controlling for these risk and mispricing factors helps to ensure that the PSID ratio indeed contains incremental predictive power beyond these well-known factor models.

Consistent with our assumption that firms with zero PSID are the most opaque firms, the alphas of this group are negative and larger in absolute magnitude than those in the lowest-PSID quintile, without exception. In general, the excess returns and alphas of five quintile portfolios increase monotonically from quintile 1 to quintile 5. The long-short portfolio that buys 20% of the stocks with the highest PSID (quintile 5) and short-sells 20% of the stocks with the lowest PSID (quintile 1) earns a value-weighted average return of 0.55% per month with a t-statistic of 3.16, translating into an annual return of 6.6%.[*] Controlling for the robust risk and mispricing factors does not change the magnitude and statistical significance of the return spreads on the PSID-sorted portfolios for most of the factor models. The only exception is the alpha of the long-short portfolio under the mispricing factor model, where the alpha decreases from 0.63% (CAPM) to 0.44% (SY model) per month and the corresponding t-statistic decreases from 3.62 to 2.54 for the value-weighted portfolio, suggesting that the return predictability is potentially driven by mispricing rather than compensation for risk.[†] Finally, the significant relation between PSID and future returns is largely coming from the long leg of the arbitrage portfolio as the economic magnitude and statistical significance are larger among the stocks in the long leg than those in the short leg. This implies that high PSID firms are undervalued relative to firms with lower PSID, perhaps due to investors' limited attention.[‡]

Next, we examine the persistence of the rank of PSID and the persistence of the return predictability of PSID. If the rank of PSID is persistent, investors would be able to learn from the past and we would not be able to detect mispricing over a long sample period. Panel B of Table 1.3 presents the probability of staying in the same PSID group or moving to any of the other five PSID groups including the zero-PSID group in the next year. Specifically, we present the average

[*]The t-statistics reported in our portfolio and regression analyses are Newey and West (1987) adjusted with six lags to control for heteroskedasticity and autocorrelation.

[†]We also examine the PSID portfolio performance using the q-factor model augmented with an expected growth factor introduced by Hou et al. (2020), and the results remain intact.

[‡]We report the performance of the equal-weighted portfolios in the online appendix. As shown in Table A1 of the online appendix, the magnitudes of the return and alpha spreads on the equal-weighted portfolio are similar to those on the value-weighted portfolio in Panel A of Table 1.3. Another notable point in Table A1 is that the economic and statistical significance of the short leg is much lower than the long leg.

probability that a stock in quintile i (defined by the rows) in year t will be in quintile j (defined by the columns) in the year $t+1$. All the probabilities in the matrix should be approximately 17% (six portfolios including the zero-PSID portfolio) - 20% (five quintile PSID portfolios) if the evolution for PSID for each stock is random and the relative magnitude of PSID in one period has no implication about the relative PSID values in the next year. However, Panel B of Table 1.3 shows that 71.51% of stocks in the lowest PSID quintile (P1) in year t continue to be in the same quintile in year $t+1$. Similarly, 87.25% of the stocks in the highest PSID quintile (P5) in year t continue to be in the same quintile in year $t+1$. More than half of the stocks (54.17%) in the zero-PSID portfolio in year t continue to be in the same zero-PSID portfolio in year $t+1$. These results overall suggest that PSID is a highly persistent equity characteristic.

The previous analyses show that investors underprice (overprice) securities with the highest (lowest) PSID in the past with the expectation that this behavior will persist in the future. If the expectation of PSID was a characteristic that evolved randomly over the years, we would expect no relation between PSID and future stock returns. The fact that PSID is persistent and it has an anomalous relation with the cross-section of expected equity returns suggests the possibility that investors underestimate the magnitude of the cross-sectional persistence uncovered in Panel B of Table 1.3. We delve further into this possibility in the test of long-term portfolio returns.

We investigate the long-term predictive power of PSID by calculating the six-factor alphas of the PSID quintiles from two to twelve months after portfolio formation. The results are presented in Table 1.4. During the second month after portfolio formation, the quintile that contains the stocks with the highest (lowest) PSID has a value-weighted return of 23 (-31) basis points. The difference is equal to 55 basis points and significant with a t-statistic of 2.72. Similarly, the zero-cost arbitrage portfolio has a return of 46 basis points with a t-statistic of 2.39 during the third month after portfolio formation. The predictive power of PSID on future returns diminishes as one moves further away from the portfolio formation month and becomes insignificant after the eight month. These results show that the positive cross-sectional relation between PSID and future returns is not just a one-month affair and the underreaction to PSID persists several months into the future, which is consistent with the theoretical evidence of continuation by Hong and Stein (1999) as a consequence of the gradual diffusion of private firm information.

Furthermore, we examine the performance of each single ratio constructed based on each of the seven financial items. Again, at the end of June of each year, we sort stocks into quintiles based on the non-zero single ratio at the end of previous year. We then form a value-weighted long-short portfolio that takes a long position in the highest PSID ratio quintile and a short position in the lowest PSID ratio quintile. The factor models used to test the performance are the same as in Table 1.3.

Table A2 presents the results. On average, stocks in the highest quintile outperform stocks in the lowest quintile under various factor models for each of the seven ratios. The magnitudes of these excess returns vary from 0.27% to 0.46%. Moreover, the long-short alphas are generally significant after controlling for the robust factor models, while losing significance for some of the single ratios with respect to the four-factor FFC and the five-factor FFCPS models. Remarkably, the economic magnitudes of the long-short profits are much smaller than that of the long-short portfolio constructed using the PSID. This implies that the comprehensive measure of PSID indeed captures more information about firms' future returns compared to just one single ratio.

We also test if an alternative measure of information disclosure produces similar predictability results. Specifically, we construct a substitute measure of PSID defined as the fraction of number of private subsidiaries that disclose at least one of the seven financial variables relative to the number of private subsidiaries that do not report at all for each public parent firm each year. Again, we sort individual stocks of public parent firms based on this alternative measure of PSID into quintiles and form long-short equity portfolios. Table A3 shows that the main finding is robust to alternative measures of PSID.

We further examine the profits from the long-short PSID portfolio by presenting the value-weighted return spreads on a per annum basis from 2006 to 2019. Figure 1.1 plots the time-series pattern of the annual long-short portfolios. Remarkably, the long-short portfolio returns are negative only in 3 out of the 14 years, and the magnitudes are smaller than 2% in absolute magnitude, while the portfolio returns are above 5% in 8 out of the 14 years, and the value-weighted returns in 2007, 2008, 2014, 2017, and 2018 are above 10%, implying that the PSID-based trading strategy is robust and earns stable positive annual profits. For comparison, the value-weighted long-short strategy earns a monthly return of 0.55% throughout our sample period, which is substantially higher than that of the SMB (-0.002%), the HML (-0.22%), RMW (0.27%), CMA (0.03%), and MOM (0.02%), which are not significant during the same period, except the profitability factor (RMW).

We also compare the cumulative performance of the PSID factor to other individual factors used in Table 1.3. To construct the PSID factor, we follow Fama and French (1993) and sort all stocks into two groups at the June of each year based on their market capitalization with the size breakpoints determined by the median market capitalization of stocks traded on the NYSE. We also independently sort all stocks in our sample into three groups using PSID based on the NYSE breakpoints. The intersection of the two size and three PSID groups constitute six portfolios. The PSID factor is the difference in the average return of the two value-weighted high-PSID portfolios and the average return of the two value-weighted low-PSID portfolios. Figure 1.2 plots the cumulative excess returns of the PSID factor and other factors from July of 2006 to December of 2019. The cumulative returns of the PSID factor are upward trending even through the financial crisis in 2007 and 2008. As of December 2019, only PSID factor and PERF factor can earn more than 100% profits, while most of the established factors earn much lower and even negative profits during the same sample period.

1.3.2 Average portfolio characteristics

We investigate which firm-specific attributes can potentially explain the anomalous significantly positive relation between PSID and expected stock returns. To do so, we sort stocks based on their PSID into quintiles each month and report the time-series averages of the cross-sectional average of various firm-specific characteristics for each quintile. The results are reported in Table 1.5.

We present the average stock characteristics of each PSID quintile portfolio and the long-short portfolio. The characteristics include private subsidiaries' information disclosure (PSID), number of private subsidiaries (NumofPriSub), log book-to-market ratio (BM), log market capitalization (SIZE), gross profitability (GP), illiquidity (ILLIQ), idiosyncratic volatility (IVOL), medium-term stock momentum (MOM), short-term reversal (STR), asset growth (AG), standardized unexpected earnings (SUE), turnover (TO), institutional ownership (IO), media coverage (MediaCov), the three-year moving sum of the absolute value of discretionary accruals (Opacity), and a proxy for readability of 10-K fillings (FOG index).

By construction, the average PSID increases monotonically from portfolio 1 to portfolio 5. The average PSID for stocks in portfolio 1 is 0.06 and the average PSID for stocks in portfolio 5 is 0.40. The difference of average PSID for stocks between portfolio 1 and portfolio 5 (P5-P1) is 0.35 and highly significant (t-statistic = 36.84), indicating significant cross-sectional variation in the PSID ratios of public parent firms. As PSID increases across the quintiles, some characteristics increase too. Such characteristics include market capitalization (SIZE), gross profitability (GP), medium-term stock momentum (MOM), short-term reversal (STR), asset growth (AG), standardized unexpected earnings (SUE), institutional ownership (IO), media coverage (MediaCov), and a proxy for readability of 10-K fillings (FOG index). The increase of average characteristics across the quintiles is economically and statistically significant for almost all of the aforementioned variables. However, the statistical significance is low or absent for asset growth, SUE, and media coverage.

As PSID increases across the quintiles, some characteristics decrease. Such characteristics include the number of private subsidiaries (NumofPriSub), log book-to-market ratio (BM), illiquidity (ILLIQ), idiosyncratic volatility (IVOL), turnover (TO), and the three-year moving sum of the absolute value of discretionary accruals (Opacity). The decrease of average characteristics across the quintiles is economically and statistically significant, except the three-year moving sum of the absolute value of discretionary accruals (Opacity).

Prior literature suggests that the firm-specific attributes considered in Table 1.5 are instrumental in analyzing the cross-section of expected stock returns. Stocks with higher PSID, higher stock momentum returns, higher profitability, and lower idiosyncratic volatility tend to have higher expected returns. Considering the prior findings in the literature and the patterns that the firm-specific attributes exhibit across the PSID quintiles, one may think that momentum, profitability, and/or idiosyncratic volatility drive the significantly positive relation between the PSID and expected stock returns. The fact that stocks with higher PSID have higher stock momentum and lower idiosyncratic volatility suggests that the positive relation between the PSID and future stock returns tends to be in line with the mispricing-based explanation. Furthermore, the stocks with higher PSID have lower book-to-market ratios, larger market capitalization, and higher liquidity, suggesting that the positive relation between the PSID and expected stock returns contradicts with the risk-based explanation. We further analyze these potential driving forces of the return predictability in Section 6.

1.3.3 Fama-MacBeth cross-sectional regressions

In this section, we conduct firm-level Fama-MacBeth regressions to test if the PSID predicts the cross-section of future monthly returns. This test allows us to examine the predictive power of the key variable of interest (PSID) more precisely while controlling for other known return predictors. Each month, we run a cross-sectional regression of stock returns in that month on the past PSID as well as a number of control variables, including lagged size, book-to-market, gross profitability, asset growth, and earnings surprise. We control for the short-term return reversal, medium-term price momentum, idiosyncratic volatility, illiquidity, and turnover ratio. We also control for the number of private subsidiaries under the parent firm as the predictive power of PSID may be correlated with the number of private firms that public firms own. Following Fama and French (1992), we skip six months between the accounting-related control variables and stock returns to ensure that the accounting information is publicly available to investors. To minimize the effect of outliers, we winsorize all independent variables each month at the 1% level. In Table A4, we report the summary statistics of the PSID ratios across Fama-French 48 industries. Even though the PSID are distributed evenly across all industries, we cannot rule out the possibility that the return predictability is attributed to the industry momentum since some good news contained in the same industry that private subsidiaries belong to may affect private subsidiaries' financial reporting as well as the parent firms' future returns. Therefore, we also control for the industry fixed effects following the 48-industry classification scheme of Fama and French (1997). The stock-level cross-sectional regressions are run each month and the time-series standard errors are corrected for heteroskedasticity and autocorrelation following Newey and West (1987).

Table 1.6 reports the results for firms with at least five private subsidiaries. In column 1, we include the PSID as well as other well-known return predictors in the cross-sectional regressions. Consistent with the portfolio results, we find a positive and significant relation between the PSID and one-month-ahead returns controlling for a large number of predictors. The average slope coefficient on the PSID ratio is 0.75 with a t-statistic of 2.65. The spread in the average standardized PSID between quintiles 5 and 1 is approximately 0.55, and multiplying this spread by the average slope of 0.75 yields an economically significant return difference of 0.41% per month, controlling for all else. In most cases, the slope coefficients on the control variables are consistent with prior literature: Short term reversal (STR) and asset growth (AG) are negatively correlated with the future return, and gross profitability and earnings surprise (SUE) are positively related to the next month's return. However, the sign of momentum (MOM) is negative and insignificant, which is due to the momentum crash in 2009 (Daniel and Moskowitz (2016)) and it becomes positive when we exclude the year 2009 from our sample. In addition, the coefficient on the number of private subsidiaries is positive but insignificant, indicating that the PSID predictability is not driven by this number. In column 2, we further control the industry fixed effect using Fama-French 48-industry classifications. However, the PSID retains significant predictive power, although the magnitude of the average slope coefficient decreases slightly to 0.57.

In column 3, we include $INDRET_{t+1}$, which is computed as the value-weighted Fama-French 48-industry portfolio returns, as a control variable in our main regression to further control for the industry effect. Specifically, we adjust the dependent variable, by subtracting the firm's value-weighted Fama-French 48-industry return $INDRET_{t+1}$ from the firm's current month return. Doing so allows us to tease out the return predictive power from the PSID rather than the one-month industry momentum effect (Moskowitz and Grinblatt (1999)). The coefficient of the PSID remains similar controlling for the industry return directly. In column 4, we further control for the common characteristics that are shown to affect stock returns systematically. Specifically, we follow Daniel et al. (1997) to compute the characteristics-adjusted returns, which is the difference between the firm's return and the corresponding DGTW benchmark portfolio returns. We replace the firm's raw return with this characteristics-adjusted return as the dependent variable and run the same monthly cross-sectional regressions. Again, the magnitude of the slope coefficient on PSID becomes slightly weaker, but it remains highly significant.

Overall, these results indicate that the PSID provides incrementally value-relevant information. The predictive power of the PSID is distinct and robust to the inclusion of other well-known return predictors and asset pricing specifications.

1.3.4 Bivariate portfolio-level analysis

An alternative explanation of the PSID effect is that the return predictability is related to the information environment of the firms proxied by firm's earnings management (Collins and Hribar (2000); Teoh, Welch, and Wong (1998); Teoh, Welch, and Wong (1998)). By no means we should expect a more transparent environment predicts higher returns, but still we examine the interaction with opacity and test if the PSID effect is independent of the information transparency. Following Hutton, Marcus, and Tehranian (2009), the opacity is measured as the three-year moving sum of the absolute value of discretionary accruals, which is a proxy for the opacity of financial statement information. In addition, even if public firms have to disclose their private subsidiaries' information in their annual reports, to some extent these firms may increase the complexity of the vocabularies or syntax in the reports such that investors cannot easily and accurately interpret the information contained in the reports. Hence, the return predictive power of PSID may be correlated with the readability of financial reports. We measure the readability of the financial report with a fog index, which is a well-known measure of readability in the literature. We focus on the readability of the 10-K annual report as public firms disclose all relevant information in this report.

We perform independent bivariate sort analysis to examine these two alternative explanations. At the end of June of year t from 2006 to 2019, we independently sort firms into quintiles based on the non-zero PSID and into two groups based on these two characteristics using the information at the end of year t-1. This intersection produces ten portfolios for each characteristic. We then form a long-short PSID portfolio in each subgroup. The portfolios are held for the next twelve months. We compute the value-weighted average monthly excess returns and alphas estimated from alternative factor models.

Panel A of Table 1.7 presents results from the bivariate portfolios of PSID and the measure of opacity. Across the two opacity groups, the differences in the number of firms and the PSID of each quintile portfolio are quite small, while the size of each PSID-sorted portfolio is generally smaller in those high opacity firms, consistent with the literature that smaller firms are less transparent compared to large firms. The monthly average excess returns and alphas of the long-short portfolio are much larger and significant in the high opacity groups. Specifically, the long-short excess return is 0.69% and the alphas range from 0.53% to 0.69% in the high opacity groups. In contrast, the excess return is 0.46% per month and the alphas are in the range of 0.42% and 0.60% per month in the low opacity groups. However, these returns and alphas are all statistically significant at the 5% significance level or better, suggesting that our main finding is not driven by the opacity effect.

Panel B of Table 1.7 presents results from double sorting on PSID and the fog index. In general, the high fog index group consists of firms with similar size compared to the low fog index group across the five PSID-sorted portfolios. The abnormal returns on the long-short PSID portfolio in the high fog index group range from 0.32% to 0.51% and statistically significant at the 10% level, except for the mispricing factor model (SY). For the low fog index group, the magnitude of the abnormal returns on the long-short PSID portfolio is much larger and they are all statistically significant at the 1% level. This suggests that the return predictability of the PSID measure is also not driven by the effect of the readability of financial reports.

In short, the independent double sorts provide strong evidence that the PSID does contain robust, valuable information about future equity returns.

1.4 Subsequent operating performance

In this section, we examine whether the PSID ratio actually contains valuable information about fundamental performance of the company. If the disclosing patterns of the private subsidiaries are indeed positively related to the firm's real operating activities, we should expect public firms that are disclosing more accounting information on their private subsidiaries continue to perform well in the future. Thus, we conduct yearly Fama-MacBeth regressions of the measures of operating performance on the PSID as well as the control variables used in Table 1.6. Specifically, we run the following cross-sectional regressions for each year:

$$OP_{i,t+1} = \alpha + \beta_1 * PSID_{i,t} + \beta_2 * OP_{i,t} + \beta_3 * \Delta OP_{i,t} + controls_{i,t} + industry_{i,t} + e_{i,t+1} \quad (1.1)$$

where $OP_{i,t+1}$ is the firm i's operating performance in year t+1, $\Delta OP_{i,t}$ is the change in operating performance between year t and year t-1, and $industry_i$ is the dummy variable that equals one for the industry that firm i belongs to and zero otherwise based on the Fama-French 48-industry classifications. We include the past operating performance in the model to account for persistence in operating performance. We also include the change in operating performance to control for the mean-reversion of operating performance (Fama and French (2000)). We further control for size, book-to-market, momentum, earnings surprise, idiosyncratic volatility, illiquidity, and asset growth in the regressions. To reduce the influence of outliers, we winsorize all variables at the 1% and 99% levels and standardize all independent variables to have a zero mean and standard deviation of one.

To measure the operating performance, we use three proxies that are prevalent in the literature, namely return-on-asset (ROA), cash flow (CF), and gross margin (GM). We measure the ROA as income before extraordinary items plus interest expenses divided by lagged total assets. Cash flow is measured as income before extraordinary items minus total accruals divided by average total assets. Gross margin is computed as sales minus cost of goods sold divided by current sales. To reduce the noise in gross margin, we follow Kothari, Laguerre, and Leone (2002) to truncate gross margin at 1 or -1. These three measures all reflect the real operating performance of a company (Hirshleifer, Hsu, and Li (2013), Hirshleifer, Hsu, and Li (2018)).

Table 1.8 presents the average slope coefficients and intercepts and the corresponding Newey-West t-statistics from the yearly Fama-MacBeth cross-sectional regressions. The results show a significantly positive relationship between the PSID and the proxies of operating performance in the next year. Specifically, in the first two columns, we regress ROA in year t+1 on the PSID as well as the ROA and the change in ROA in year t. The coefficient of the PSID is 0.20 and significant at the 5% level after accounting for the control variables and the industry effects, meaning that a one standard deviation increase in the PSID leads to 0.20% increase in the ROA in the next year. Similarly, the coefficients between the PSID and other measures of operating performance are 0.37 for cash flow and 0.30 for gross margin, and are all statistically significant. Furthermore, the significantly positive coefficients on the proxies of operating performance in the current year and the significantly negative coefficients on the change of proxies of operating performance are consistent with the literature on the persistence and mean-reversion in the operating performance. Overall, the results indicate that the PSID indeed contains valuable information about the firm's future operating performance.

1.5 Sources of return predictability

Having established that the predictive power of the PSID may be driven by slow dissemination of disclosure-related information due to investors' underreaction, we seek to understand the cross-sectional sensitivity of our main result to proxies of investors' limited attention and limits to arbitrage. To this end, we perform the multivariate regression analysis on the proxies and the PSID. Specifically, we split the sample into two groups based on the median value of each proxy and run regressions separately for each group to examine whether the PSID effect varies in these two groups.

1.5.1 Investors' limited attention

Our work is related to the growing research on investors' limited attention to information disclosure. Patell and Wolfson (1982) find that investors have lower attention to news released during the market closing time than to news released during the market trading time. Barber and Odean (2007) argue that individual investors can only process limited investment choices due to limited time and resources they have. DellaVigna and Pollet (2009) find investors' inattention to earnings announcements on Fridays, compared to the earnings announcements on other weekdays. Hirshleifer, Lim, and Teoh (2009) find that investors' underreaction to earning surprises and post-earnings-announcement drift are stronger for firms that announce earnings on days that many other firms announce earnings due to investors' limited attention. Cohen and Frazzini (2008) find that suppliers' have delayed responses to the information disclosure of their customers. Cohen and Lou (2012) find that single-segment firm returns predict returns of multi-segment firms operating in the same industry, consistent with the limited attention argument. Lee et al. (2019) use patent technology class to define technology-linked firms and find return predictability across these firms.

Following the aforementioned studies, we argue that investors pay limited attention to public firms' disclosure to private subsidiaries' financial information. If investors were fully aware of this information, the stock price of a public firm would quickly adjust to the information reflected in disclosing patterns of private subsidiaries. Following the literature, we use three proxies of investor attention: media coverage (Fang and Peress (2009)), transient institutional ownership (Bushee (2001)), and absolute SUE (Bali et al. (2018)). As argued by Bushee (2001) and Hirshleifer, Hsu, and Li (2018), transient institutional investors trade stocks based on short-term strategies and are thus less likely to pay as much attention to firms' fundamentals as long-term-orientated dedicated institutional investors.* Bali et al. (2018) shows that firms with greater absolute earnings surprises are more likely to attract investor attention, increasing investor awareness of firms' specific characteristics. Therefore, firms with lower media coverage, higher transient institutional

*As discussed by Bushee (2001), transient institutions with short investment horizons are characterized as having high portfolio turnover and highly diversified portfolio holdings. They not only exhibit strong preferences for near-term earnings, but these preferences also translate into significant misvaluations.

ownership, or lower absolute SUE receive less attention from investors and should exhibit more sluggish stock price reactions to the information contained in private subsidiaries' information disclosure and greater predictability of stock returns.

The media coverage is defined as the number of news articles covering the stock in one month, using news from Thomson Reuters News Analytics. If the number of media news is missing, we set it to zero. To purge out the size effect, we regress logarithm of the number of media news on logarithm of firm's market capitalization and use the residual as our proxy of investor attention. The transient institutional investors are classified following Bushee (2001). For absolute SUE, we use the last non-missing SUE value that is released prior to the June of each year during the past 12 months. Panel A of Table 1.9 reports the regression results accounting for the same set of control variables used in Table 1.6 and the industry effects. For brevity, we just report the coefficients of our main variable PSID. Consistent with our hypothesis, the results show that the return predictability of the PSID is stronger among stocks with lower investor attention. The magnitudes of coefficients of PSID in the low attention group are much larger and statistically significant, while the magnitudes of coefficients are smaller and insignificant in the high attention group. For example, the average slope of PSID in low residual media coverage subsample is 0.88 with a t-statistic of 2.40, while the one in high residual media coverage subsample is only 0.3 with a t-statistic of 1.01. Overall, the results support our hypothesis that the return predictability is driven by investors' limited attention to the information contained in the PSID.

1.5.2 Limits to arbitrage

Results in the previous section suggest that investors' inattention is a source of the return predictability, but we do not fully understand what sustains this return predictability. In this section, we further explore the role of limits to arbitrage. If the predictive power of the PSID is driven by mispricing to some extent, then we should expect the return predictability to be more pronounced for stocks with high arbitrage costs. In our next test, we use three proxies of limits-to-arbitrage that are prevalent in the literature.

Following Nagel (2005), we use the residual institutional ownership (i.e., size-orthogonalized institutional ownership) at the end of June of each year as a proxy for limits-to-arbitrage. Following Pontiff (2006), we use idiosyncratic volatility as an alternative proxy for costly arbitrage. We rely on Ang et al. (2006) and measure the monthly IVOL as the standard deviation of the daily residuals from the regression of daily excess stock returns on the three factors of Fama and French (1993) over the past one month. Finally, following Amihud (2002), we construct the illiquidity measure in the current month as our third proxy. Panel B of Table 1.9 reports the regression results controlling for the usual suspects in Table 1.6 and the industry effects. Consistent with the limits-to-arbitrage hypothesis, the coefficients are higher and significant in high idiosyncratic volatility group and high illiquidity group, and lower in the high institutional ownership group, while the coefficients are smaller and only marginally significant in high institutional ownership and low illiquidity groups. Thus, slow diffusion of private information into stock prices due to limits to arbitrage may provide a partial explanation to the return predictability of the PSID-based trading strategy.

1.5.3 Anomaly-based mispricing and PSID

The results so far suggest that stocks with high PSID tend to be undervalued relative to stocks with low PSID, but we have not yet provided any formal empirical evidence that high-PSID stocks are indeed undervalued. Thus, we investigate this conjecture by assessing the mispricing score of the stocks directly.

Specifically, we employ the composite mispricing measure originally constructed by Stambaugh, Yu, and Yuan (2015a) to identify if high-PSID stocks are indeed undervalued. The composite mispricing measure is the average of percentiles of 11 prominent anomalies, including net stock issues, composite equity issues, accruals, net operating assets, asset growth, investment-to-assets, distress, O-score, momentum, gross profitability, and return on assets. At the end of June of each year, we conduct independent double sorts based on a stock's composite mispricing measure and its PSID. We then compute the average composite mispricing measure for stocks in each of the 25 portfolios.

Table 1.10 shows that high-PSID stocks indeed tend to have a lower average mispricing score compared to low-PSID stocks. Furthermore, the difference in the composite mispricing measures between the low- and high-PSID stocks within each composite mispricing score quintile is statistically significant. This evidence supports our mispricing argument that stocks with high-PSID value are truly undervalued.

1.6 Risk versus mispricing explanation

The results so far suggest that the standard asset pricing models of risk do not explain the cross-sectional variation in returns associated with the private information disclosure. However, there is still the possibility of a risk-based mechanism that leads to the return predictability. For example, the PSID can predict the future change in risk, which would lead to a change in the firm's expected return. In this case, the high abnormal return of high-PSID stocks is justified as a means of investors' compensation for high risk, instead of an underreaction to the PSID-related information. In this section, we conduct tests to explore whether alternative measures of risk could plausibly explain our results.

1.6.1 Earnings prediction

If investors could not fully capture the implication of the private information disclosure on the firm's profitability, they would be surprised by the earnings realizations in the future. Thus, we examine whether the PSID can predict the future earnings controlling for the past earnings. We use standardized unexpected earnings (SUE), defined as actual earnings in the current quarter minus earnings 4 quarters ago, scaled by stock price in the current quarter, following Livnat and Mendenhall (2006), to proxy for earnings surprise. We conduct Fama-MacBeth regressions of the SUE from quarter q+3 in year t+1 to quarter q+2 in year t+2 on the PSID and other accounting variables at the end of year t as well as other priced-based controls in last month prior to each quarter. We also control for the industry effects following the 48-industry classification of Fama and French (1997). We winsorize all variables at the 1% and 99% levels and standardize all independent variables to have a zero mean and standard deviation of one to reduce the effect of outliers. Finally, we also examine the future SUEs over longer time periods, while keeping all independent variables the same. If the PSID contains information about future earnings, we should expect the slope coefficient to be positive and significant.

Consistent with our expectation, the first column of Table 1.11 shows that the coefficient on the PSID is 0.10 with a t-statistic of 2.37 accounting for past SUE, control variables, and the industry effects. Moreover, consistent with Bernard and Thomas (1989), the lagged SUE at quarter q is strongly positively correlated with the future SUE. In columns 2 to 4, we repeat the Fama-MacBeth regressions with the same independent variables but replace the dependent variable (SUEs) in subsequent quarters. The coefficients on the PSID decrease monotonically from column 2 to column 4, and they all become statistically insignificant, indicating that the earnings predictability of the PSID decays quickly after one quarter. This is consistent with the underreaction hypothesis that the PSID reflects slow diffusion of cash flow news into stock prices rather than a change in the future discount rate or compensation for risk.

1.6.2 Return patterns around earnings announcements

To further differentiate the underreaction-to-information mechanism from a risk-based explanation, we examine stock price reactions around earnings announcements. If the return predictability were explained by underlying risk, we would expect the returns to be evenly affected in the subsequent periods. In contrast, if the effect is consistent with mispricing, then the returns must be disproportionately affected around earnings announcements, meaning that the return prediction around earnings announcement should be stronger than that around non-earnings announcement period if investors are surprised by the good or bad news during that period.

We test these two distinct hypotheses by examining stock price reactions around subsequent earnings announcements. This approach is widely used in the literature (see, e.g., Bernard and Thomas (1989); La Porta et al. (1997); Engelberg, McLean, and Pontiff (2018); Lee et al. (2019)).

Following Engelberg, McLean, and Pontiff (2018), we conduct a panel regression analysis of daily stock returns (DLYRET) on the last available PSID, an earnings announcement window dummy (EDAY), and the interaction term between the two variables. We also include a set of control variables, consisting of the lagged values for each of the past ten days for stock returns, stock returns squared, and trading volume. We also control for day fixed effects and cluster the standard errors by day.

The earnings announcement date is defined as in Engelberg, McLean, and Pontiff (2018). Specifically, we examine the firm's trading volume scaled by market trading volume for the day before, the day of, and the day after the reported earnings announcement date, which is obtained from Compustat quarterly database. We then define the day with the highest scaled trading volume as the earnings announcement day. We select one-day or three-day earnings announcement window centered on the earnings announcement date in our analysis.

Columns 1 and 2 in Table 1.12 report the regression results for one-day window, and columns 3 and 4 present the results for three-day window. In all cases, the coefficients are all positive and significant. Consistent with the mispricing explanation, returns to the high PSID stocks are much larger during earnings news releasing dates. In column 2, the coefficient on PSID is 0.059, while the coefficient of PSID*EDAY interaction term is 0.341, meaning that the return spread for the hedged PSID strategy is 5.78 times higher during an earnings announcement window than on non-announcement days. Analogously, based on column 4, the PSID premium is 4.23 times higher during a three-day earnings announcement window than on non-announcement days. This is comparable to the findings of Engelberg, McLean, and Pontiff (2018) that the anomaly returns on average are six times higher on the earning announcement day and three times higher in the three-day earning announcement window. Thus, the evidence supports our mispricing argument that investors do not fully incorporate the PSID-driven return predictability information into their earnings forecasts and are therefore surprised when earnings are realized.

1.6.3 Testing potential risk-based explanations

The results have so far shown that the standard factor models or traditional measures of risk do not explain the cross-sectional variation in stock returns associated with the PSID effect. In this section, we provide comprehensive evidence from testing alternative risk-based explanations. Specifically, we rely on the established rational asset pricing models and investigate whether these models' implied measures of risk can be the driving force of the PSID-return relation.

We first test whether the CAPM explains the PSID premium. Specifically, we report total volatility, idiosyncratic volatility, and market beta for each PSID-sorted quintile portfolio as well as the zero-PSID portfolio. The CAPM implied measures of market beta, total volatility, and idiosyncratic volatility are estimated for each month using the past 60-month individual stock returns. Table 1.13 shows that the CAPM does not explain the PSID premium as the high-PSID stocks have lower total volatility, lower idiosyncratic volatility, and lower market beta than the low-PSID stocks.

Next, we investigate if the PSID effect can be explained by the intertemporal CAPM (ICAPM) of Merton (1973) and/or the consumption CAPM (CCAPM) of Breeden (1979). Following Ang et al. (2006) and Campbell et al. (2018), we use the change in VIX – S&P500 index option implied volatility – as the second factor of the two-factor ICAPM model.[*] Specifically, we estimate the VIX beta for each stock and each month by running the time-series regressions of excess stock returns on the excess market returns and the change in VIX in the past 60 months. To test for the CCAPM explanation, we compute the consumption beta for each stock and each month by regressing the excess stock returns on the consumption growth rate in the past 60 months.[†] We convert the quarterly consumption data to monthly frequency using linear and cubic spline interpolation

[*] Campbell et al. (2018) extend Merton's original model by proposing a two-factor ICAPM with stochastic volatility in which an unexpected increase in future market volatility represents deterioration in the investment opportunity set.

[†] The central implication of the CCAPM is that the expected return on an asset is related to "consumption risk," that is, how much uncertainty in consumption would come from holding the asset. Assets that lead to a large amount of uncertainty offer large expected returns, as investors want to be compensated for bearing consumption risk. Thus, the expected excess return on a risky asset is proportional to the covariance of its return and consumption in the period of the return.

methods and the consumption beta estimates turn out to be similar from both methods.* Results in Table 1.13 show that neither the ICAPM nor the CCAPM explains the PSID effect. Specifically, the high-PSID stocks tend to have a higher VIX beta than the low-PSID stocks, implying lower future return for the high-PSID stocks in the ICAPM framework. Also, as presented in Table 1.13, the VIX beta difference between the low-PSID and high-PSID groups is statistically insignificant. In addition, the high-PSID stocks have a lower consumption beta than the low-PSID stocks, rejecting the CCAPM explanation for the PSID premium.

Finally, we investigate the magnitude of the factor exposures to see if the PSID-driven return spread is positively loaded on these factors. Specifically, we estimate stock exposure to each factor (ex-ante factor beta) for each month by regressing the excess stock returns on each of these well-established factors in the past 60 months. Generally, the stocks in the highest PSID quintile have lower factor exposures than those in the lowest PSID quintile. The exceptions are the MOM beta, PERF beta, ROE beta, and PEAD beta, and only the differences on two behavioral factors, PERF beta and PEAD beta, between the low-PSID and high-PSID stocks are significant. Overall, these results indicate that the predictive power of the PSID is not explained by alternative measures of risk.

1.6.4 PSID vs. low-risk anomalies

Table 1.13 shows that the average total volatility (TVOL), idiosyncratic volatility (IVOL), and market beta (BETA) of the high-PSID stocks is somewhat lower than the average TVOL, IVOL, and BETA of the low-PSID stocks, rejecting the standard risk-based explanation. However, these results suggest that the PSID premium may potentially be explained by the betting-against-beta, idiosyncratic volatility, or lottery demand effects. Contrary to the fundamental principle that higher risk is compensated with higher expected return, Ang et al. (2006) and Frazzini and Pedersen (2014) show that high-volatility (high-beta) stocks underperform low-volatility (low-beta) stocks. Bali, Cakici, and Whitelaw (2011) and Bali et al. (2017) show that the market beta, idiosyncratic

*The quarterly consumption data (CAY) are obtained from Martin Lettau's online data library: https://sites.google.com/view/martinlettau/data?authuser=0.

volatility, and demand for lottery-like stocks are highly correlated and retail investors' preference for lottery stocks is a driving factor in these well-established low-risk anomalies. To test whether the cross-sectional relation between PSID and the future equity returns of public parent firms is explained by the low-risk anomalies, we control for the betting-against-beta (BAB), idiosyncratic volatility (IVOL), and the lottery demand (MAX) factors of Ang et al. (2006), Frazzini and Pedersen (2014), and Bali, Cakici, and Whitelaw (2011); Bali et al. (2017).[*] Specifically, we estimate the abnormal returns of the PSID-sorted portfolios reported in Table 3, Panel A, by extending the well-established factor models with the BAB, IVOL, and MAX factors. Table A5 of the online appendix shows that the alpha spreads on the long-short portfolios of PSID remain economically and statistically significant after controlling for the BAB, IVOL, and MAX factors, indicating that the low-risk anomalies do not explain the PSID premium.

1.7 Additional analyses

1.7.1 PSID and analyst forecast errors

Given that the private subsidiaries' information disclosure (PSID) contains value-relevant information about firm's future performance, we next examine if professionals, such as financial analysts, can fully understand the value relevance of PSID. If these financial analysts indeed underreact to such information, it is also likely for investors who rely on financial analysts to suffer from the same bias.

Specifically, we conduct Fama-MacBeth regressions of analyst forecast errors (AFE) in year t+1 on PSID and control variables in year t. The analyst forecast error is measured as the difference

[*]The BAB and MAX factors are borrowed respectively from Frazzini and Pedersen (2014) and Bali et al. (2017): http://www.lhpedersen.com/data and https://sites.google.com/a/georgetown.edu/turan-bali. For the idiosyncratic volatility (IVOL) factor, we follow Fama and French (1993) and sort all stocks into two groups at the end of each month based on their market capitalization with the breakpoints determined by the median market capitalization of stocks traded on the NYSE. We also independently sort all stocks in our sample into three groups using IVOL based on the NYSE breakpoints. The intersection of the two size and three IVOL groups constitute six portfolios. The IVOL factor is the difference in the average return of the two value-weighted high-IVOL portfolios and the average return of the two value-weighted low-IVOL portfolios.

between actual earnings per share (EPS) and the latest analyst consensus forecast before the fiscal year end of the year being forecasted, scaled by lagged total assets. Control variables include lagged analyst forecast errors (AFE), size, book-to-market, gross profitability, asset growth, earnings surprise, short-term reversal (STR), momentum (MOM), idiosyncratic volatility (IVOL), illiquidity (ILLIQ), turnover ratio (TURNOVER), and the number of private subsidiaries. We also control for the industry effects. All independent variables are based on the last non-missing observation for each year t and are standardized to have a zero mean and standard deviation of one. We conduct the cross-sectional analysis for each year. To reduce the effect of outliers, we winsorize all variables at the 1% and 99% levels each year.

Model 1A and Model 1B in Table 1.14 report the time-series averages of the slope coefficients and the corresponding t-statistics without and with the industry effects, respectively. The results show that analysts indeed underreact to the PSID related information as the PSID does positively predict the future analyst forecast errors. More specifically, in Model 1B, one standard deviation increase in PSID induces a 0.03% increase in the next year's analyst forecast errors, with a t-statistic of 2.01. This is also economically significant as the sample mean (median) of the forecast errors is 0.09% (0.00%).

In short, these results suggest that even professionals such as financial analysts cannot fully incorporate the valuable information contained in the PSID. As many investors make investment decisions based on the analyst earnings forecasts, we would expect that these investors also underreact to such information, which is again consistent with the limited attention explanation for the predictive power of the PSID.

1.7.2 PSID and institutional trading

Next we examine if sophisticated investors, such as institutional investors, exploit such value-relevant information from PSID in their investment decisions, i.e., if institutional investors incorporate the information contained in PSID, they should trade in the direction indicated by PSID. Similar to the regressions with analyst forecast errors, we conduct annual Fama-MacBeth regressions of net purchases by institutional investors on lagged PSID, controlling for the same set of firm characteristics. The net purchases by institutional investors are computed as yearly change in the fraction of a firm's shares outstanding held by institutional investors. Model 2A and Model 2B in Table 1.14 present the results. Consistent with our hypothesis, the insignificant coefficients on the PSID ratio suggest that more informed institutional investors do not respond to the information contained in the private subsidiaries' information disclosure, providing further support for retail investors' limited attention and underreaction to such information.

1.8 Conclusions

This paper examines the asset pricing implications of private information disclosure about public parent firms. We find that the information contained in the disclosing behaviors of public firms' private subsidiaries is slowly incorporated into stock prices due to investors' limited attention and arbitrage costs. A proxy for such private information disclosure does predict the cross-sectional variation in future equity returns of public parent firms significantly, and the established factor models do not explain the predictive power of this measure. Further analyses show that the proxies for investors' inattention and limits-to-arbitrage are associated with stronger return predictability, suggesting that the PSID-return relation is consistent with the mispricing explanation.

We conduct comprehensive analyses to differentiate the risk vs. mispricing explanations. First, we examine the market reactions around earnings announcements and find that the predictive power of the PSID is stronger around earnings announcements than that around non-earnings announcement periods. Second, we find that market professionals such as financial analysts cannot fully process the information contained in PSID. Third, the stocks in the highest PSID quintile portfolio have lower average beta, total and idiosyncratic volatility, and their exposures to the established risk factors are lower than those in the lowest PSID quintile portfolio. These results suggest that the return predictability is driven by mispricing rather than compensation for risk.

Overall, we contribute to the literature on the impact of investor inattention on equity returns by proposing a new dimension of information disclosure (or disclosure delay originated from private subsidiaries) and by presenting novel evidence that the theory of investor inattention is important in understanding stock market underreactions to the informative signals provided by private subsidiaries of public parent firms. Our findings may also have implications on the valuation of corporate disclosure on its private subsidiaries' information. If the markets are slow in responding to firms' value-relevant disclosing patterns, there will be a potential misallocation of resources across these firms. Thus, a better understanding of the mechanism of the investors' underreaction to the PSID related information may shed lights on facilitating information incorporation and achieving higher market efficiency.

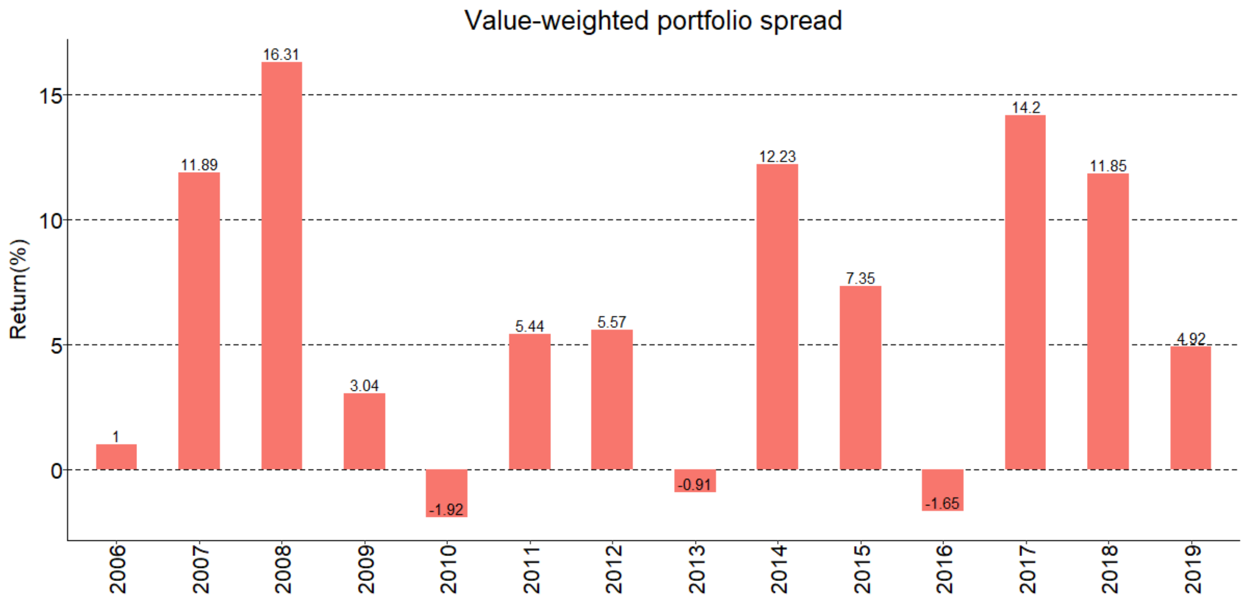


Fig. 1.1 Annual value-weighted returns of the long-short portfolio

This figure shows the annual value-weighted returns of the long-short portfolio of stocks sorted by non-zero PSID. At the end of June of each year t from 2006 to 2019, individual stocks of public parent firms are sorted into quintile portfolios based on non-zero PSID at the end of year t-1 from 2005 to 2018, and are held for the next twelve months (July of year t to June of year t+1). The long-short portfolio buys stocks in the top PSID quintile and sells stocks in the bottom PSID quintile. There are six months in 2006 and twelve months in other years.

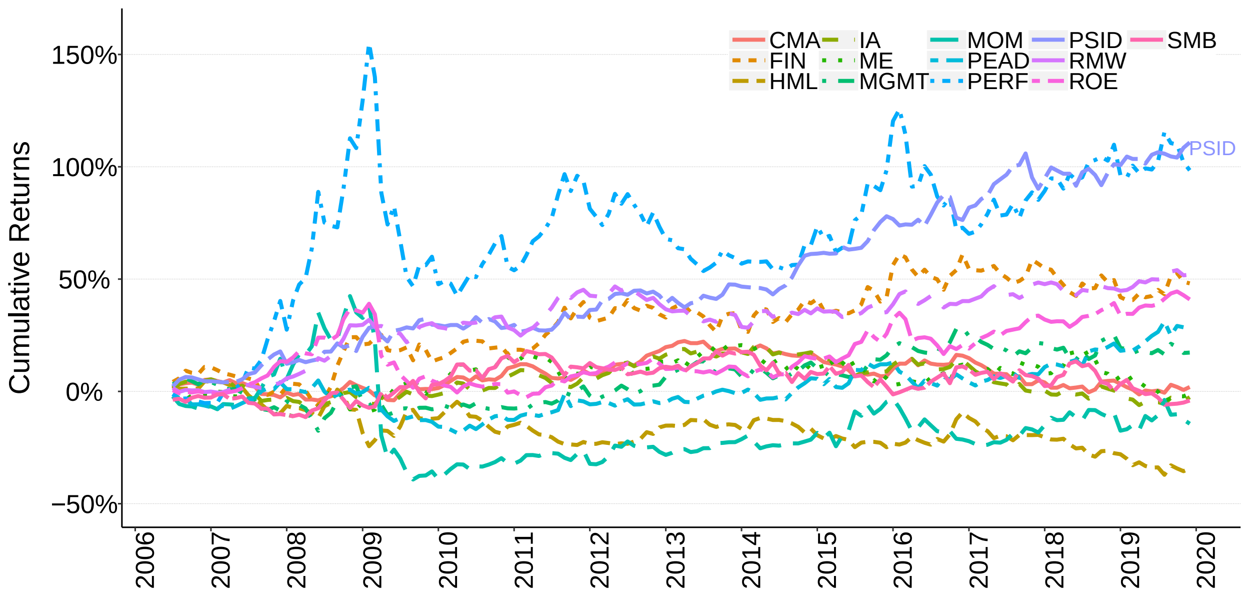


Fig. 1.2 Cumulative returns of the PSID factor and other factors

This figure plots the cumulative value-weighted returns of the PSID factor and other individual factors from July 2006 to December 2019. To construct the PSID factor, we follow Fama and French (1993) and sort all stocks into two groups at the June of each year based on their market capitalization with the size breakpoints determined by the median market capitalization of stocks traded on the NYSE. We also independently sort all stocks in our sample into three groups using PSID based on the NYSE breakpoints. The intersection of the two size and three PSID groups constitute six portfolios. The PSID factor is the difference in the average return of the two value-weighted high-PSID portfolios and the average return of the two value-weighted low-PSID portfolios.

Table 1.1 Summary Statistics for Public Parent Firms

This table presents the descriptive statistics of public parent firms with at least one private subsidiary across years. Table reports, respectively, the number of public parent firms, the ratio of total market capitalization of public parent firms to the total market capitalization of all firms trading at the NYSE/NASDAQ/AMEX (Mkt Share), the average PSID ratio, the average number of private subsidiaries, and the average information disclosure ratio for each of the seven financial variables (operating revenue, total assets, number of employees, income before tax, net income, cash flow, and shareholders' funds). The sample period is from 2005 to 2018.

Year	# of firms	Mkt Share(%)	PSID	# of PriSub	Revenue	Assets	Number of employees	Income before tax	Net income	Cash flow	Shareholder funds
2005	2506	63.11	0.29	21	0.43	0.14	0.42	0.12	0.12	0.11	0.14
2006	2860	69.28	0.18	39	0.34	0.11	0.35	0.09	0.09	0.08	0.11
2007	2617	66.09	0.18	29	0.36	0.14	0.33	0.11	0.11	0.08	0.12
2008	2519	67.49	0.23	30	0.35	0.14	0.35	0.11	0.11	0.07	0.14
2009	2457	66.94	0.21	31	0.26	0.15	0.34	0.12	0.12	0.08	0.15
2010	2315	65.72	0.22	32	0.28	0.17	0.33	0.13	0.13	0.08	0.16
2011	2361	66.25	0.24	34	0.35	0.16	0.36	0.13	0.14	0.08	0.16
2012	2480	67.78	0.20	37	0.32	0.15	0.31	0.12	0.13	0.07	0.15
2013	2481	64.40	0.19	39	0.28	0.14	0.29	0.11	0.12	0.07	0.14
2014	2745	67.87	0.14	51	0.24	0.11	0.27	0.08	0.08	0.05	0.10
2015	2593	71.39	0.10	53	0.13	0.12	0.13	0.10	0.10	0.06	0.12
2016	2522	69.87	0.11	53	0.13	0.13	0.14	0.10	0.11	0.07	0.13
2017	2559	71.29	0.11	57	0.14	0.13	0.15	0.11	0.11	0.07	0.13
2018	2579	71.33	0.13	60	0.16	0.15	0.17	0.11	0.11	0.07	0.14

Table 1.2 Summary Statistics for the Cross-Sectional Variables

This table reports the summary statistics for the cross-sectional variables. The sample consists of all common stocks (share codes equal to 10 or 11) that are listed on NYSE, Nasdaq, and Amex. Financial firms (with one-digit SIC = 6), utility firms (with two-digit SIC = 49), and stocks that are below the 20th percentile of NYSE market capitalization are excluded from the analysis. The sample is further restricted to firms with at least five private subsidiaries. PSID is computed as the equal-weighted average of the information disclosure ratios of seven financial variables (operating revenue, total assets, number of employees, income before tax, net income, cash flow, and shareholders' funds) disclosed by the private subsidiaries of public firms scaled by the total number of private subsidiaries. #ofPriSub is the total number of private subsidiaries under each public firm. RET_{t+1} is the one-month-ahead return. SIZE is the firm's market capitalization computed as the logarithm of the market value of the firm's outstanding equity at the end of month t-1. BM is the logarithm of the firm's book value of equity divided by its market capitalization, where the BM ratio is computed following Fama and French (2008). Firms with negative book values are excluded from the analysis. Short-term reversal (STR) is the stock's one-month lagged return. MOM is the stock's cumulative return from the start of month t-12 to the end of month t-2 following Jegadeesh and Titman (1993). Gross Profitability (GP) is the firm's gross profitability following Novy-Marx (2013), calculated as revenue minus cost of goods sold scaled by total assets. Asset Growth (AG) is a percentage of total asset growth between two consecutive fiscal years following Cooper, Gulen, and Schill (2008). TO is the monthly turnover computed as the number of trading shares divided by the total number of shares outstanding in month t-1. ILLIQ is the monthly illiquidity measure following Amihud (2002), computed using daily data in month t-1. IVOL is the idiosyncratic volatility following Ang et al. (2006) over month t-1. SUE is the standardized unexpected earnings defined as actual earnings in the current quarter minus earnings 4 quarters ago, scaled by stock price in the current quarter following Livnat and Mendenhall (2006). IO is the total shares held by institutions from 13F filings in each quarter scaled by total shares outstanding. MediaCov is the number of media news covering the firm. Opacity is defined as the three-year moving sum of the absolute value of discretionary accruals following Hutton, Marcus, and Tehranian (2009). Fog index is a proxy for readability of 10-K fillings. All variables except PSID and NumofPrisub in the sample are winsorized at the 1% level for both tails to mitigate the effect of outliers. The mean, standard deviation (SD), minimum, median, and maximum of each variable are shown in Panel A, and their pairwise correlations with the PSID are presented in Panel B. The sample period is from July 2006 to December 2019.

Panel A: Descriptive statistics

	Mean	Sd	Min	Med	Max
PSID	0.20	0.14	0.00	0.19	0.91
#ofPriSub	66.45	122.68	5.00	28.00	4115.00
SIZE	8.25	1.35	4.13	8.03	12.48
BM	-0.94	0.78	-4.09	-0.88	1.28
GP	0.34	0.21	-0.34	0.30	1.13
ILLIQ	0.14	0.42	0.00	0.04	22.87
IVOL	0.02	0.01	0.00	0.01	0.13
MOM	0.24	0.48	-0.94	0.15	4.42
STR	0.01	0.10	-0.63	0.01	1.46
AG	1.11	0.26	0.60	1.06	3.20
SUE	0.12	2.07	-51.67	0.11	29.25
TO	0.47	0.85	0.02	0.21	10.99
IO	77.63	21.71	0.00	82.92	100.00
MediaCov	32.36	48.41	0.00	17.00	458.83
Opacity	0.16	0.15	0.01	0.11	1.27
FOG Index	20.11	1.06	17.41	20.03	25.37
RET_{t+1}	0.01	0.10	-0.63	0.01	1.46

Panel B: Pearson (Spearman) correlations below (above) the diagonal

	PSID	# of PriSub	SIZE	BM	GP	ILLIQ	IVOL	MOM	STR	AG	SUE	TO	IO	MediaCov	Opacity	FOG Index	RET_{t+1}
PSID		-0.04	0.12	-0.12	0.14	-0.09	-0.15	0.06	0.04	0.02	0.05	-0.13	0.03	0.01	-0.03	0.04	0.03
#ofPriSub	-0.10		0.38	0.01	-0.03	-0.34	-0.16	0.01	-0.00	-0.11	-0.01	-0.09	-0.08	0.27	-0.12	0.08	-0.00
SIZE	0.12	0.30		-0.25	-0.07	-0.93	-0.34	0.17	0.06	0.04	0.11	-0.04	-0.19	0.55	-0.14	0.09	-0.01
BM	-0.12	0.02	-0.23		-0.47	0.22	0.02	0.02	0.02	-0.19	-0.16	-0.01	-0.03	-0.05	-0.11	0.01	0.02
GP	0.08	-0.05	-0.03	-0.43		0.03	0.04	0.00	0.00	0.04	0.02	0.02	0.11	-0.05	0.11	-0.12	0.00
ILLIQ	-0.04	-0.06	-0.37	0.09	0.00		0.31	-0.17	-0.04	-0.02	-0.08	-0.17	0.11	-0.59	0.10	-0.09	0.01
IVOL	-0.12	-0.08	-0.32	0.03	0.02	0.35		-0.24	-0.04	0.07	-0.08	0.58	0.13	0.14	0.20	-0.06	-0.03
MOM	0.04	-0.02	0.12	0.04	-0.00	-0.17	-0.23		-0.02	-0.07	0.29	-0.14	0.00	-0.04	0.02	0.03	-0.01
STR	0.03	-0.01	0.05	0.02	0.01	-0.01	-0.05	-0.03		-0.02	0.03	-0.06	-0.00	0.01	0.01	0.01	-0.01
AG	-0.01	-0.06	0.00	-0.09	-0.08	-0.00	0.09	-0.06	-0.02		0.04	0.05	0.07	-0.04	0.09	-0.01	-0.02
SUE	0.05	-0.01	0.10	-0.09	0.02	-0.15	-0.13	0.25	-0.00	0.01		-0.07	-0.00	-0.04	-0.03	-0.00	-0.02
TO	-0.12	-0.06	-0.11	0.01	0.01	-0.10	0.53	-0.05	-0.05	0.075	-0.06		0.27	0.27	0.18	-0.05	-0.04
IO	0.06	-0.04	-0.13	-0.04	0.08	-0.07	0.04	-0.02	-0.01	0.02	0.01	0.13		-0.13	0.13	0.02	-0.01
MediaCov	0.04	0.24	0.61	-0.06	-0.02	-0.14	0.05	-0.04	-0.00	-0.02	-0.02	0.12	-0.14		-0.02	0.06	-0.01
Opacity	-0.06	-0.08	-0.05	-0.07	0.01	0.00	0.12	0.03	0.00	0.10	-0.02	0.14	0.02	0.02		0.03	0.01
FOG Index	0.03	0.02	0.09	0.03	-0.13	-0.04	-0.06	0.02	0.01	-0.00	0.00	-0.04	-0.01	0.04	0.00		0.01
RET_{t+1}	0.03	-0.01	-0.03	0.02	0.01	0.09	-0.01	-0.02	0.04	-0.02	-0.07	-0.04	-0.01	-0.01	0.00	0.01	

Table 1.3 Univariate Portfolio Analysis

Panel A reports the average monthly excess returns and alphas on the value-weighted portfolios of stocks sorted by the PSID. At the end of June of each year t from 2006 to 2019, individual stocks of public parent firms are sorted into quintiles based on non-zero PSID at the end of year t-1 from 2005 to 2018, and are held for the next twelve months (July of year t to June of year t+1). "Zero" is the portfolio of firms with zero-PSID (no information disclosure). P1 is the portfolio of stocks with the lowest PSID and P5 is the portfolio of stocks with the highest PSID. L/S is a zero-cost portfolio that buys stocks in quintile 5 (highest PSID) and sells stocks in quintile 1 (lowest PSID). All returns and alphas are expressed in percentage. Excess return is the raw return of the portfolio over the risk-free rate. Alpha is the intercept from a time-series regression of monthly excess returns on the factors of alternative models: the CAPM, a four-factor model including Fama and French (1993) three-factor and Carhart (1997) momentum factor (FFC), a five-factor model including FFC and the liquidity factor of Pástor and Stambaugh (2003) (FFCPS), Fama and French (2015) five-factor model (FF5), Fama and French (2018) six-factor model (FF6), Stambaugh and Yuan (2017) mispricing-factor model (SY), Hou, Xue, and Zhang (2015) q-factor model (HXZ), and Daniel, Hirshleifer, and Sun (2020) behavior factor model (DHS). Newey and West (1987) adjusted t-statistics are given in parentheses. The sample period is from July 2006 to December 2019. Panel B reports the transition probabilities for PSID at a lag of one year between 2005 and 2018. For each PSID quintile in year t, the percentage of stocks that fall into each of the year t+1 PSID quintile is calculated, and the time-series averages of these one-year-ahead transition probabilities are presented.

Panel A: Returns and alphas on PSID-sorted quintile portfolios

Rank	Excess Return	CAPM	FFC	FFCPS	FF5	FF6	SY	HXZ	DHS
Zero	0.28	-0.67	-0.52	-0.49	-0.50	-0.49	-0.42	-0.37	-0.46
	(0.51)	(-1.88)	(-1.51)	(-1.43)	(-1.38)	(-1.38)	(-1.19)	(-1.07)	(-1.24)
P1	0.51	-0.27	-0.28	-0.25	-0.38	-0.38	-0.21	-0.19	-0.20
	(1.42)	(-2.23)	(-2.34)	(-2.20)	(-3.11)	(-3.13)	(-1.73)	(-1.73)	(-1.60)
P2	0.87	0.19	0.18	0.19	0.03	0.03	0.06	0.18	0.15
	(2.81)	(2.08)	(1.96)	(2.05)	(0.41)	(0.39)	(0.69)	(1.93)	(1.56)
P3	0.80	0.02	-0.05	-0.04	-0.05	-0.05	0.01	0.10	0.00
	(2.24)	(0.14)	(-0.47)	(-0.39)	(-0.53)	(-0.52)	(0.08)	(0.96)	(0.02)
P4	0.88	0.14	0.08	0.10	0.08	0.04	0.04	0.14	0.12
	(2.65)	(1.68)	(1.04)	(0.94)	(0.49)	(0.48)	(0.43)	(1.56)	(1.32)
P5	1.06	0.36	0.27	0.26	0.22	0.22	0.23	0.32	0.34
	(3.33)	(3.73)	(3.21)	(3.12)	(2.59)	(2.59)	(2.66)	(3.41)	(3.44)
L/S	0.55	0.63	0.55	0.51	0.60	0.60	0.44	0.52	0.54
	(3.16)	(3.62)	(3.26)	(3.14)	(3.42)	(3.46)	(2.54)	(3.03)	(3.03)

Panel B: One-year-ahead transition matrix of PSID-sorted portfolios

	PSID rank in year t+1					
PSID rank in year t	Zero	P1	P2	P3	P4	P5
Zero	54.17%	33.00%	4.17%	8.00%	0.00%	0.00%
P1	0.05%	71.51%	17.72%	5.78%	3.66%	1.28%
P2	0.05%	13.74%	63.74%	15.25%	5.21%	2.01%
P3	0.03%	2.06%	15.57%	67.80%	11.55%	2.99%
P4	0.03%	0.98%	1.73%	15.50%	72.21%	9.55%
P5	0.00%	0.44%	0.53%	1.05%	10.73%	87.25%

Table 1.4 Long-Term Portfolio Performance

This table presents the long-term predictive power of the PSID. "Zero" is the portfolio of firms with zero-PSID (no information disclosure). P1 is the portfolio of stocks with the lowest PSID and P5 is the portfolio of stocks with the highest PSID. L/S is a zero-cost portfolio that buys stocks in quintile 5 (highest PSID) and sells stocks in quintile 1 (lowest PSID). Table reports Fama and French (2018) six-factor alphas for zero-PSID portfolio and for each of the non-zero PSID-sorted quintile portfolios from two to twelve months ahead after portfolio formation. The last column shows the differences of monthly Fama and French (2018) six-factor alphas between quintiles 5 and 1. Newey and West (1987) adjusted t-statistics are presented in parentheses.

Post-sorting months	Zero	P1(Low)	P2	P3	P4	P5(High)	L/S
m+2	-0.49	-0.31	0.05	-0.08	0.08	0.23	0.55
	(-0.43)	(-2.35)	(0.57)	(-0.82)	(0.87)	(2.39)	(2.72)
m+3	-0.38	-0.25	0.07	-0.01	0.08	0.21	0.46
	(-0.35)	(-2.06)	(0.75)	(-0.12)	(0.95)	(2.22)	(2.39)
m+4	-0.28	-0.15	0.09	0.02	0.12	0.20	0.35
	(-0.26)	(-1.46)	(0.98)	(0.22)	(1.30)	(2.01)	(2.02)
m+5	0.18	-0.09	0.09	0.01	0.15	0.19	0.29
	(0.18)	(-0.96)	(1.00)	(0.11)	(1.67)	(1.87)	(1.73)
m+6	-0.23	-0.12	0.08	0.04	0.12	0.19	0.31
	(-0.23)	(-1.20)	(0.92)	(0.50)	(1.32)	(1.84)	(1.87)
m+7	-0.94	-0.1	0.04	0.04	0.17	0.17	0.27
	(-1.05)	(-0.94)	(0.39)	(0.43)	(1.89)	(1.64)	(1.68)
m+8	-0.69	-0.12	0.03	-0.01	0.22	0.16	0.28
	(-0.70)	(-1.14)	(0.36)	(-0.13)	(2.08)	(1.60)	(1.73)
m+9	-1.05	-0.11	0.03	-0.04	0.22	0.14	0.25
	(-1.07)	(-0.95)	(0.28)	(-0.33)	(2.29)	(1.39)	(1.51)
m+10	-1.44	-0.09	0.01	-0.03	0.21	0.13	0.22
	(-1.59)	(-0.83)	(0.11)	(-0.21)	(2.00)	(1.33)	(1.46)
m+11	-1.62	-0.10	-0.05	-0.02	0.28	0.10	0.20
	(-1.90)	(-1.02)	(-0.58)	(-0.17)	(2.76)	(0.99)	(1.41)
m+12	-1.24	-0.14	-0.05	-0.02	0.27	0.11	0.25
	(-1.52)	(-1.33)	(-0.58)	(-0.17)	(2.90)	(1.12)	(1.48)

Table 1.5 Average Portfolio Characteristics

This table presents the average stock characteristics of the portfolios formed based on the PSID. "Zero" is the portfolio of firms with zero-PSID (no information disclosure). P1 is the portfolio of stocks with the lowest PSID and P5 is the portfolio of stocks with the highest PSID. L/S is a zero-cost portfolio that buys stocks in quintile 5 (highest PSID) and sells stocks in quintile 1 (lowest PSID). Table reports the time-series averages of the monthly averages for PSID and various firm-specific characteristics for each portfolio. The last two columns show the differences for the firm-specific characteristics between P1 and P5 and the associated Newey and West (1987) adjusted t-statistics. The PSID and other firm-specific characteristics are defined in Table 1.2. The sample period is from July 2006 to December 2019.

Variables	Zero	P1	P2	P3	P4	P5	P5-P1	t-stat
PSID	0.00	0.06	0.13	0.20	0.28	0.40	0.35	(36.84)
NumofPriSub	16.89	91.89	55.53	63.10	72.69	53.08	-38.81	(-11.16)
SIZE	7.48	8.08	8.17	8.21	8.38	8.47	0.39	(7.89)
BM	-1.17	-0.83	-0.88	-0.88	-0.98	-1.12	-0.30	(-9.64)
GP	0.38	0.32	0.33	0.32	0.35	0.38	0.06	(17.55)
ILLIQ	0.20	0.15	0.15	0.14	0.13	0.11	-0.04	(-2.36)
IVOL	0.02	0.02	0.02	0.02	0.01	0.01	-0.00	(-7.18)
MOM	0.24	0.21	0.23	0.23	0.24	0.25	0.04	(3.75)
STR	0.00	0.01	0.01	0.01	0.01	0.01	0.00	(4.98)
AG	1.17	1.12	1.11	1.11	1.12	1.12	0.00	(0.22)
SUE	0.17	0.11	0.11	0.06	0.17	0.18	0.07	(1.65)
TO	0.56	0.51	0.47	0.44	0.42	0.41	-0.10	(-4.66)
IO	74.76	75.70	76.27	77.87	78.52	80.08	4.38	(8.78)
MediaCov	17.73	30.44	31.08	32.59	34.33	34.23	3.79	(1.77)
Opacity	0.19	0.17	0.16	0.15	0.15	0.15	-0.02	(-1.20)
FOG Index	20.01	20.05	20.05	20.10	20.09	20.19	0.14	(5.41)

Table 1.6 Fama-MacBeth Cross-Sectional Regressions

This table reports the Fama and MacBeth (1973) cross-sectional regression results. The sample period is from July 2006 to December 2019. The PSID and other accounting variables in year t are matched to monthly stock returns from July of year t+1 to June of year t+2. The monthly price-based variables are calculated using the last non-missing observations prior to each month. The dependent variable is the firm's future raw return in the first two columns, the firm's future excess return over its value-weighted industry peers' return (Column 3), or the firm's DGTW adjusted return (Column 4). We include industry dummies and classify each firm's industry peers based on the Fama-French 48-industry classifications. All returns are expressed in percentage. The PSID and other firm-specific characteristics are defined in Table 1.2. All explanatory variables are generated using the last non-missing available observation for each month t-1. Cross-sectional regressions are run every calendar month, and the time-series standard errors are corrected for heteroskedasticity and autocorrelation. Newey and West (1987) adjusted t-statistics are reported in parentheses. Coefficients marked with *, **, and *** are significant at the 10%, the 5%, and the 1% level, respectively.

Independent Variables	*RET*	*RET*	*RET -INDRET*	DGTW-adj. *RET*
PSID	0.75***	0.57**	0.79***	0.57***
	(2.65)	(2.22)	(2.71)	(2.83)
STR	-1.77**	-1.98***	-1.54*	-1.92**
	(-2.16)	(-2.80)	(-1.79)	(-2.56)
MOM	-0.54	-0.68	-0.51	-0.64
	(-0.94)	(-1.21)	(-0.82)	(-1.07)
AG	-0.06	-0.05	-0.11	-0.09
	(-0.43)	(-0.42)	(-0.75)	(-0.82)
BM	-0.14*	-0.09	-0.11	-0.07
	(-1.83)	(-1.36)	(-1.45)	(-0.98)
GP	0.21	0.24	0.31	0.37
	(0.62)	(0.87)	(0.87)	(1.29)
SIZE	-0.05	-0.05	-0.05	-0.05
	(-1.07)	(-1.08)	(-0.96)	(-1.05)
SUE	0.09***	0.10***	0.06**	0.07***
	(2.63)	(3.31)	(2.04)	(2.76)
TO	-0.51	-0.07	-0.55	-0.11
	(-1.24)	(-0.23)	(-1.24)	(-0.32)
ILLIQ	2.36	8.73	11.40	25.31
	(0.07)	(0.25)	(0.36)	(0.77)
IVOL	-13.04*	-13.71**	-9.67	-11.91*
	(-1.78)	(-2.19)	(-1.29)	(-1.79)
NumofPriSub	0.00	0.00	0.00	0.00
	(0.63)	(0.77)	(0.56)	(0.71)
Intercept	1.19*	0.99*	1.03	0.82
	(1.83)	(1.80)	(1.44)	(1.37)
Industry FEs	No	Yes	No	Yes
N	155591	155591	155591	154833
Adj.R^2	0.081	0.153	0.078	0.152

Table 1.7 Bivariate Portfolio Analysis

At the end of June of year t from 2006 to 2018, individual stocks are independently sorted into quintile portfolios based on non-zero PSID of year t-1 from 2005 to 2017 and a control variable (opacity and fog index). Opacity is constructed as the three-year moving sum of the absolute value of discretionary accruals following Hutton, Marcus, and Tehranian (2009), and it is a proxy for the opacity of the financial reports. Fog index is a proxy for the readability of 10-K fillings. We construct a long-short PSID portfolio within each control quintile of opacity (or fog index) and hold for next twelve months (July of year t to June of year t+1). Value-weighted portfolio excess returns and alphas, expressed in percentage, are reported. Excess return is the raw return of the portfolio over the risk-free rate. Alpha is the intercept from a time-series regression of monthly excess returns on the factors of alternative models: Fama and French (2015) five-factor model, Fama and French (2018) six-factor model, Stambaugh and Yuan (2017) mispricing-factor model, Hou, Xue, and Zhang (2015) q-factor model, and Daniel, Hirshleifer, and Sun (2020) behavioral factor model. Number of firms, average PSID, and average market capitalization are also reported for each portfolio. Newey and West (1987) adjusted t-statistics are given in parentheses. The sample period is from July 2006 to December 2019.

Panel A: Double sort on PSID and opacity

Low opacity

PSID Rank	# of firms	PSID	Size	Excess Return	FF5	FF6	SY	HXZ	DHS
P1	89.00	0.06	10362.32	0.48	-0.40	-0.40	-0.26	-0.21	-0.24
				(1.37)	(-2.92)	(-2.90)	(-2.08)	(-1.90)	(-1.97)
P2	91.00	0.13	13180.09	0.84	0.10	0.10	0.09	0.18	0.17
				(2.82)	(1.13)	(1.15)	(0.94)	(2.11)	(1.79)
P3	89.00	0.20	12144.76	0.78	-0.03	-0.03	-0.00	0.10	-0.00
				(2.22)	(-0.28)	(-0.26)	(-0.03)	(0.91)	(-0.02)
P4	93.00	0.28	17415.93	0.84	-0.01	-0.01	0.02	0.11	0.10
				(2.28)	(-0.12)	(-0.12)	(0.17)	(1.10)	(0.88)
P5	89.00	0.40	16388.17	0.94	0.20	0.20	0.17	0.26	0.23
				(2.88)	(1.75)	(1.75)	(1.51)	(2.25)	(2.15)
L/S				**0.46**	**0.60**	**0.60**	**0.47**	**0.47**	**0.49**
				(2.96)	(2.95)	(2.94)	(2.30)	(2.72)	(2.86)

High opacity

PSID Rank	# of firms	PSID	Size	Excess Return	FF5	FF6	SY	HXZ	DHS
P1	91.00	0.06	6846.57	0.56	-0.34	-0.34	-0.14	-0.13	-0.12
				(1.28)	(-1.91)	(-1.91)	(-0.99)	(-0.91)	(-0.83)
P2	89.00	0.13	9880.30	0.71	-0.20	-0.20	-0.11	0.01	-0.09
				(1.74)	(-1.27)	(-1.27)	(-0.67)	(0.03)	(-0.54)
P3	90.00	0.20	13863.65	0.94	0.07	0.07	0.09	0.20	0.14
				(2.05)	(0.41)	(0.41)	(0.65)	(1.35)	(0.91)
P4	86.00	0.28	14549.62	0.72	-0.14	-0.14	-0.12	-0.03	0.11
				(2.04)	(-1.02)	(-1.05)	(-0.91)	(-0.17)	(0.71)
P5	90.00	0.41	20338.75	1.25	0.35	0.35	0.39	0.47	0.52
				(3.64)	(2.31)	(2.26)	(2.46)	(2.77)	(3.11)
L/S				**0.69**	**0.69**	**0.68**	**0.54**	**0.60**	**0.64**
				(2.71)	(2.61)	(2.57)	(2.37)	(2.66)	(2.85)

Panel B: Double sort on PSID and fog index									
Low fog index									
PSID Rank	# of firms	PSID	Size	Excess Return	FF5	FF6	SY	HXZ	DHS
P1	103.00	0.06	7783.57	0.36	-0.52	-0.51	-0.36	-0.36	-0.33
				(0.88)	(-3.14)	(-3.12)	(-2.51)	(-2.45)	(-2.33)
P2	105.00	0.13	11410.73	0.90	0.05	0.05	0.06	0.17	0.16
				(2.83)	(0.41)	(0.40)	(0.44)	(1.28)	(1.10)
P3	99.00	0.20	10638.81	0.79	-0.05	-0.05	0.02	0.12	0.08
				(2.02)	(-0.33)	(-0.31)	(0.16)	(0.98)	(0.58)
P4	99.00	0.28	14775.79	0.77	-0.09	-0.09	-0.09	0.01	0.03
				(2.23)	(-0.71)	(-0.74)	(-0.77)	(0.12)	(0.22)
P5	90.00	0.40	18048.61	1.02	0.22	0.22	0.21	0.29	0.30
				(3.13)	(2.14)	(2.10)	(2.03)	(2.43)	(2.58)
L/S				**0.66**	**0.74**	**0.73**	**0.58**	**0.65**	**0.64**
				(3.26)	(3.62)	(3.56)	(3.08)	(3.29)	(3.45)
High fog index									
PSID Rank	# of firms	PSID	Size	Excess Return	FF5	FF6	SY	HXZ	DHS
P1	95.00	0.06	8991.02	0.66	-0.27	-0.27	-0.08	-0.05	-0.08
				(1.67)	(-1.80)	(-1.81)	(-0.61)	(-0.35)	(-0.54)
P2	92.00	0.13	10430.33	0.80	0.01	0.01	0.05	0.18	0.11
				(2.16)	(0.09)	(0.09)	(0.38)	(1.45)	(0.95)
P3	99.00	0.20	14334.54	0.89	0.03	0.03	0.07	0.16	0.02
				(2.12)	(0.23)	(0.24)	(0.51)	(1.12)	(0.13)
P4	98.00	0.28	15735.90	0.97	0.17	0.17	0.15	0.25	0.21
				(2.62)	(1.77)	(1.83)	(1.50)	(2.67)	(2.18)
P5	107.00	0.40	17996.14	1.09	0.22	0.22	0.24	0.34	0.38
				(3.29)	(1.45)	(1.43)	(1.51)	(2.06)	(2.42)
L/S				**0.42**	**0.50**	**0.50**	0.32	**0.38**	**0.46**
				(1.97)	(1.83)	(1.82)	(1.34)	(1.68)	(2.05)

Table 1.8 Subsequent Operating Performance

This table reports the results from the Fama and MacBeth (1973) cross-sectional regressions of individual firm's operating performance measured in year t+1 on the PSID and other control variables in year t. ROA is income before extraordinary items plus interest expenses divided by lagged total assets. Cash flows (CF) is income before extraordinary items minus total accruals (i.e., changes in current assets plus changes in short-term debt and minus changes in cash, changes in current liabilities, and depreciation expenses) divided by average total assets. GM (gross margin) is measured by sales minus cost of goods sold divided by current sales. If GM exceeds 1, it is set to 1. If GM is lower than – 1, it is set to – 1. ΔROA_t ($\Delta Cash_t$ and ΔGM_t) is the change in ROA (Cash and GM) from year t-1 to year t. We classify each firm's industry peers based on Fama-French 48-industry classifications. We winsorize all variables at the 1% and 99% levels and standardize all independent variables to have a zero mean and standard deviation of one. The control variables include firm size (SIZE), book-to-market (BM), momentum (MOM), asset growth (AG), earnings surprise (SUE), idiosyncratic volatility (IVOL), and Amihud's illiquidity (ILLIQ). Cross-sectional regressions are run every calendar year, and the time-series standard errors are corrected for heteroskedasticity and autocorrelation. Newey and West (1987) adjusted t-statistics are given in parentheses. Coefficients marked with *, **, and *** are significant at the 10%, the 5%, and the 1% level, respectively.

Independent Variables	ROA_{t+1}	ROA_{t+1}	CF_{t+1}	CF_{t+1}	GM_{t+1}	GM_{t+1}
PSID	0.38***	0.20**	0.69***	0.37**	0.40***	0.30***
	(3.56)	(2.35)	(3.64)	(2.42)	(4.36)	(4.03)
ROA_t	4.82***	4.37***				
	(17.10)	(14.10)				
ΔROA_t	-0.98***	-0.91***				
	(-6.94)	(-5.97)				
$Cash_t$			5.01***	4.35***		
			(7.88)	(6.86)		
$\Delta Cash_t$			-2.71***	-2.15***		
			(-9.72)	(-7.94)		
GM_t					19.46***	19.34***
					(95.25)	(91.22)
ΔGM_t					-0.42	-0.49*
					(-1.68)	(-2.12)
SIZE		0.21***		0.66***		0.18**
		(4.01)		(3.52)		(2.81)
BM		-1.30***		-1.16***		-0.63***
		(-7.40)		(-3.73)		(-4.12)
MOM		0.98***		0.63***		0.60***
		(10.81)		(4.14)		(7.32)
AG		-0.82***		-0.63***		-0.17
		(-11.26)		(-5.86)		(-1.64)
SUE		1.34***		1.10***		0.44***
		(17.17)		(8.94)		(4.03)
IVOL		-0.41***		-1.10***		-0.04
		(-4.86)		(-8.85)		(-0.25)
ILLIQ		0.15***		0.07		-0.03
		(3.58)		(0.68)		(-0.52)
Intercept	6.86***	6.54***	4.34***	4.12***	41.90***	41.81***
	(15.67)	(17.54)	(6.17)	(7.58)	(45.38)	(46.21)
Industry FEs	Yes	Yes	Yes	Yes	Yes	Yes
N	12136	12136	12674	12674	13021	13021
Adj. R^2	0.447	0.571	0.179	0.250	0.896	0.900

Table 1.9 Limited Attention and Limits to Arbitrage

This table presents the Fama and MacBeth (1973) regression results from subsample analysis. The sample period is from July 2006 to December 2019. The PSID and other accounting variables in year t are matched to monthly stock returns from July of year t+1 to June of year t+2. Panel A splits the stock sample into two subsamples based on whether the proxies of investor attention are below or above the median value. Panel B splits the stock sample into two subsamples based on whether the proxies of limits to arbitrage are below or above the median value. Proxies of investor attention are the residual media coverage, transient institutional ownership, and absolute SUE. Proxies of limits to arbitrage are the residual institutional ownership, idiosyncratic volatility, and Amihud's illiquidity measure. Media coverage is the number of media news covering the firm in a month, using data from Thomson Reuters News Archive. Transient institutional investors are classified following Bushee (2001). Absolute SUE is defined as the absolute value of SUE based on the last non-missing SUE during the 12 months preceding June, following Bali et al. (2018). Residual institutional ownership is the residual of institutional ownership orthogonalized with respect to the firm market capitalization, following Nagel (2005). Idiosyncratic volatility is constructed following Ang et al. (2006). The illiquidity measure is calculated following Amihud (2002). Cross-sectional regressions are run every calendar month, and the time-series standard errors are corrected for heteroskedasticity and autocorrelation. Newey and West (1987) adjusted t-statistics are given in parentheses. Coefficients marked with *, **, and *** are significant at the 10%, the 5%, and the 1% level, respectively.

Panel A: Stock subsamples splitted by the proxies of investor attention

	Residual Media Coverage		Transient Institutions		Absolute SUE	
	Low	High	Low	High	Low	High
PSID	0.88**	0.30	0.53*	1.03***	0.81**	0.45
	(2.40)	(1.01)	(1.70)	(2.77)	(2.41)	(1.39)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Industry FEs	Yes	Yes	Yes	Yes	Yes	Yes
N	47604	47651	72189	72107	75949	76030
Adj. R^2	0.077	0.086	0.094	0.069	0.086	0.086

Panel B: Stock subsamples splitted by the proxies of limits to arbitrage

	Residual IO		IVOL		ILLIQ	
	Low	High	Low	High	Low	High
PSID	0.88***	0.61*	0.35	0.99***	0.57*	0.71**
	(2.61)	(1.91)	(1.01)	(2.88)	(1.67)	(2.12)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Industry FEs	Yes	Yes	Yes	Yes	Yes	Yes
N	77839	77785	77810	77781	75949	76030
Adj. R^2	0.102	0.070	0.084	0.075	0.086	0.086

Table 1.10 Anomaly-based Mispricing Measure and PSID

This table presents the average composite mispricing measure from the bivariate portfolios (independent double sorts) of stocks sorted by the composite mispricing score and PSID. At the end of June of year t from 2006 to 2018, individual stocks are independently sorted into quintiles based on the composite mispricing measure and quintiles by non-zero PSID of year t-1 from 2005 to 2017. The composite mispricing score is the average of the ranking percentiles produced by 11 anomaly variables following Stambaugh, Yu, and Yuan (2015a). The last two columns show the differences of monthly composite mispricing measures and the corresponding t-statistics between the PSID quintiles within each mispricing quintile. Newey and West (1987) adjusted t-statistics are given in parentheses.

	P1 (Low PSID)	P2	P3	P4	P5 (High PSID)	H-L	t-stat
Low Misp	30.50	30.55	30.53	30.49	30.16	-0.34	(-3.43)
2	39.33	39.37	39.30	39.40	39.27	-0.05	(-1.75)
3	45.88	45.76	45.70	45.79	45.78	-0.10	(-3.05)
4	52.75	52.55	52.79	52.63	52.38	-0.36	(-3.44)
High Misp	65.20	64.42	63.70	64.02	64.61	-0.60	(-4.36)

Table 1.11 Earnings Prediction

This table reports the results from the Fama and MacBeth (1973) regressions of individual firm's one-quarter- to four-quarter-ahead SUE on the past PSID and other control variables. All independent variables are calculated using the last non-missing observations prior to each quarter. We classify each firm's industry peers based on the Fama-French 48-industry classifications. The PSID and other firm-specific characteristics are defined in Table 1.2. We winsorize all variables at the 1% and 99% levels and standardize all independent variables to have a zero mean and standard deviation of one. Cross-sectional regressions are run every calendar quarter, and the time-series standard errors are corrected for heteroskedasticity and autocorrelation. Newey and West (1987) adjusted t-statistics are given in parentheses. Coefficients marked with *, **, and *** are significant at 10%, 5%, and 1%, respectively.

Independent Variables	SUE_{q+1}	SUE_{q+2}	SUE_{q+3}	SUE_{q+4}
PSID	0.10**	0.07	0.06	0.00
	(2.37)	(1.32)	(1.02)	(0.13)
SUE_q	0.78***	0.57***	0.67***	0.86***
	(3.91)	(5.47)	(5.64)	(3.85)
Dividend	-0.11***	-0.05**	-0.02	0.06
	(-3.88)	(-2.30)	(-1.27)	(1.52)
NGE	-0.02	-0.33	-0.53	-0.62*
	(-0.17)	(-1.27)	(-1.44)	(-1.82)
MOM	0.62***	0.39***	0.18**	0.22
	(3.57)	(2.75)	(2.65)	(1.43)
STR	0.28**	0.18**	0.21*	0.11***
	(2.12)	(2.36)	(1.89)	(3.10)
BM	-0.28*	-0.22	-0.22	-0.16
	(-1.91)	(-1.05)	(-1.05)	(-0.76)
SIZE	-0.15*	-0.04	-0.01	-0.00
	(-1.97)	(-1.32)	(-0.52)	(-0.12)
AG	-0.16**	-0.12**	-0.11*	-0.04
	(-2.41)	(-2.02)	(-1.94)	(-1.27)
ILLIQ	-0.57	-0.13	-0.06*	-0.08
	(-1.48)	(-1.47)	(-1.73)	(-0.85)
IVOL	-0.40	-0.10*	-0.16	-0.06
	(-1.67)	(-1.68)	(-1.24)	(-0.37)
TO	-0.31**	-0.15*	-0.06	-0.10
	(-2.60)	(-1.92)	(-1.34)	(-1.10)
Intercept	-0.54	-0.32	-0.22	-0.22
	(-1.23)	(-1.10)	(-1.09)	(-0.75)
Industry FEs	Yes	Yes	Yes	Yes
N	52013	50474	48941	47497
Adj. R^2	0.099	0.145	0.144	0.147

Table 1.12 Earnings Announcement Returns Prediction

This table presents the results from the panel regressions of earnings announcement window daily returns (DLYRET) on the PSID, earnings day dummy variables (EDAY), day-fixed effects, and other lagged control variables (coefficients unreported). An earnings announcement window is defined as the one-day or three-day window centered on an earnings release, i.e., days t-1, t, and t+1. EDAY is a dummy variable that equals one if the daily observation is during an announcement window, and zero otherwise. Following Engelberg et al. (2018), we obtain earnings announcement dates from the Compustat quarterly database and examine the firm's trading volume scaled by market trading volume for the day before, the day of, and the day after the reported earnings announcement date. An earnings announcement day is defined as the day with the highest scaled trading volume. Control variables include lagged values for each of the past ten days for stock returns, squared stock returns, and trading volume. Standard errors are clustered by day. The t-statistics are reported in parentheses and the coefficients marked with * , ** , and *** are significant at 10%, 5%, and 1%, respectively. The sample period is from July 2006 to December 2019.

	Panel A: One-day window		Panel B: Three-day window	
Dep.variable (%)	DLYRET	DLYRET	DLYRET	DLYRET
PSID	0.046**	0.059***	0.047**	0.060***
	(2.35)	(3.05)	(2.42)	(3.11)
PSID * EDAY	0.343***	0.341***	0.254***	0.254***
	(3.39)	(3.13)	(2.63)	(2.63)
EDAY	0.060	0.058	0.010	0.010
	(0.93)	(0.89)	(0.13)	(0.33)
Lagged controls	No	Yes	No	Yes
Day fixed effects	Yes	Yes	Yes	Yes
N	3473445	3451066	3473445	3451066
Adj. R^2	0.286	0.286	0.286	0.287

Table 1.13 Testing Risk-based Explanations

This table presents the results from testing potential risk-based explanations. Table reports the average portfolio risk attributes for zero-PSID portfolio, each quantile portfolio sorted by PSID, and the differences for the firm-specific risk attributes between quintiles 5 and 1 and the associated Newey and West (1987) adjusted t-statistics. "Zero" is the portfolio of firms with zero-PSID (no information disclosure). P1 is the portfolio of stocks with the lowest PSID and P5 is the portfolio of stocks with the highest PSID. The CAPM implied measures of total volatility (TVOL), idiosyncratic volatility (IVOL), and market (MKT) Beta are estimated for each month using the past 60-month individual stock returns. The individual stock exposures (Betas) to the change in VIX, consumption growth rate, and each risk factor are estimated for each month using the past 60-month observations. The sample period is from July 2006 to December 2019.

Risk	Zero	P1	P2	P3	P4	P5	P5-P1	t-stat
CAPM								
TVOL	0.12	0.11	0.10	0.10	0.10	0.10	-0.01	(-3.60)
IVOL	0.02	0.02	0.02	0.02	0.01	0.01	-0.00	(-7.18)
MKT Beta	1.12	1.07	1.01	1.07	1.08	1.07	-0.00	(-0.16)
ICAPM								
VIX beta	-0.08	-0.06	-0.03	-0.04	-0.05	-0.04	0.02	(0.44)
CCAPM								
Consumption Growth Beta	0.64	0.69	0.64	0.73	0.67	0.53	-0.14	(-1.33)
Factor exposures								
SMB Beta	1.00	0.68	0.64	0.65	0.56	0.54	-0.14	(-3.99)
HML Beta	-0.15	0.04	-0.03	-0.03	-0.08	-0.18	-0.22	(-5.76)
RMW Beta	0.04	0.23	0.11	0.10	-0.01	-0.11	-0.33	(-6.54)
CMA Beta	-0.38	0.07	0.10	-0.01	-0.02	0.02	-0.04	(-0.59)
MOM Beta	-0.06	-0.14	-0.13	-0.12	-0.11	-0.09	0.05	(1.56)
LIQ Beta	-0.01	0.10	0.07	0.07	0.05	0.04	-0.06	(-2.90)
MGMT Beta	-0.73	-0.75	-0.70	-0.79	-0.73	-0.78	-0.03	(-0.53)
PERF Beta	-0.70	-0.74	-0.69	-0.71	-0.70	-0.68	0.06	(2.30)
IA Beta	-0.43	-0.15	-0.17	-0.31	-0.30	-0.32	-0.17	(-2.16)
ROE Beta	-1.32	-1.39	-1.30	-1.37	-1.31	-1.31	0.09	(1.28)
FIN Beta	-1.01	-0.91	-0.87	-0.94	-0.96	-1.00	-0.09	(-1.86)
PEAD Beta	-0.88	-0.95	-0.79	-0.86	-0.83	-0.82	0.14	(2.94)

Table 1.14 PSID, Analyst Forecast Errors, and Institutional Trading

This table reports the results from the Fama and MacBeth (1973) regressions of analyst forecast errors and institutional trading in year t+1 on the PSID and control variables in year t. In Model 1A and 1B, the dependent variable is analyst forecast errors (AFE). In Model 2A and 2B, the dependent variable is institutional net buys (INB). Analyst forecast errors (AFE) is defined as the difference between actual earnings per share (EPS) and the latest analyst consensus forecast before the fiscal year end of the year being forecasted, scaled by lagged total assets. Institutional net buys (INB) is defined as the yearly change in institutional investors holding on a stock, with holding is a fraction of a firm's ownership. We classify each firm's industry peers based on the Fama-French 48-industry classifications. The PSID and other firm-specific characteristics are defined in Table 1.2. We winsorize all variables at the 1% and 99% levels and standardize all independent variables to have a zero mean and standard deviation of one. Cross-sectional regressions are run every calendar year, and the time-series standard errors are corrected for heteroskedasticity and autocorrelation. Newey and West (1987) adjusted t-statistics are given in parentheses. Coefficients marked with *, **, and *** are significant at the 10%, the 5%, and the 1% level, respectively.

	Analyst Forecast Errors		Institutional Net Buys	
Independent Variables	Model 1A	Model 1B	Model 2A	Model 2B
PSID	0.0002*	0.0003**	0.1322	0.0739
	(1.97)	(2.01)	(0.91)	(0.61)
AFE_t	0.0140***	0.0140***		
	(4.01)	(4.15)		
INB_t			-0.7260***	-0.7102***
			(-3.38)	(-3.34)
SIZE	0.0010***	0.0009***	-0.0963**	-0.1418*
	(2.62)	(3.82)	(-2.06)	(-1.73)
BM	0.0000	-0.0001	-0.2670**	-0.2659*
	(0.10)	(-0.62)	(-2.14)	(-1.84)
MOM	-0.0011**	-0.0011**	0.3779***	0.4192***
	(-2.32)	(-2.22)	(4.81)	(5.49)
AG	-0.0006***	-0.0006***	0.0614	0.0770
	(-3.16)	(-3.31)	(0.69)	(0.86)
SUE	-0.0003	-0.0003	-0.1890**	-0.1852***
	(-1.59)	(-1.54)	(-2.31)	(-4.22)
GP	-0.0010***	-0.0011***	-0.1812*	-0.0764
	(-3.15)	(-3.17)	(-1.70)	(-0.81)
TO	-0.0005	0.0001	-0.1166	-0.0486
	(-1.57)	(0.33)	(-0.87)	(-0.37)
IVOL	0.0003***	0.0004***	0.2867***	0.2661***
	(2.96)	(2.81)	(5.39)	(5.48)
ILLIQ	0.0004**	0.0004**	0.1598	0.2241
	(2.03)	(2.01)	(1.12)	(1.50)
STR	-0.0002***	-0.0003**	0.3908***	0.3520***
	(-1.51)	(-2.50)	(6.10)	(7.09)
NumofPriSub	0.0001**	0.0001*	-0.0311	0.0166
	(2.05)	(1.88)	(-0.42)	(0.18)
Intercept	0.0039	-0.0012***	-0.2914	0.3284
	(1.78)	(-3.31)	(-0.50)	(0.42)
Industry FEs	No	Yes	No	Yes
N	12897	12897	12319	12319
Adj. R^2	0.065	0.067	0.040	0.054

Appendix

"Private Subsidiaries' Information Disclosure and the Cross-Sectional Equity Returns of Public Parent Firms"

This Online Appendix includes tables referred to but not included in the main body of the paper, which provide robustness checks and additional findings.

Table A1 Equal-Weighted Portfolios of Stocks Sorted by PSID

This table reports the average monthly excess returns and alphas on the equal-weighted portfolios of stocks sorted by the PSID. At the end of June of each year t from 2006 to 2019, individual stocks are sorted into quintiles based on non-zero PSID at the end of year t-1 from 2005 to 2018, and are held for the next twelve months (July of year t to June of year t+1). "Zero" is the portfolio of firms with zero-PSID (no information disclosure). P1 is the portfolio of stocks with the lowest PSID and P5 is the portfolio of stocks with the highest PSID. L/S is a zero-cost portfolio that buys stocks in quintile 5 (highest PSID) and sells stocks in quintile 1 (lowest PSID). All returns and alphas are expressed in percentage. Excess return is the raw return of the portfolio over the risk-free rate. Alpha is the intercept from a time-series regression of monthly excess returns on the factors of alternative models. Newey and West (1987) adjusted t-statistics are given in parentheses. The sample period is from July 2006 to December 2019.

Rank	Excess Return	CAPM	FFC	FFCPS	FF5	FF6	SY	HXZ	DHS
Zero	0.41	-0.58	-0.36	-0.35	-0.40	-0.39	-0.30	-0.23	-0.26
	(0.76)	(-1.78)	(-1.29)	(-1.23)	(-1.33)	(-1.35)	(-0.99)	(-0.66)	(-0.79)
P1	0.59	-0.36	-0.20	-0.18	-0.29	-0.28	-0.11	-0.06	-0.19
	(1.33)	(-2.14)	(-1.84)	(-1.63)	(-2.14)	(-2.54)	(-0.85)	(-1.23)	(-1.11)
P2	0.91	0.00	0.12	0.15	0.04	0.05	0.14	0.24	0.13
	(2.18)	(0.01)	(1.53)	(1.88)	(0.45)	(0.56)	(1.38)	(2.93)	(1.01)
P3	0.96	-0.02	0.10	0.12	0.07	0.08	0.15	0.25	0.15
	(2.13)	(-0.15)	(1.14)	(1.40)	(0.68)	(0.87)	(1.41)	(2.78)	(1.03)
P4	1.07	0.12	0.20	0.22	0.18	0.19	0.23	0.35	0.26
	(2.46)	(0.90)	(2.46)	(2.62)	(1.90)	(2.21)	(2.32)	(3.81)	(2.02)
P5	1.10	0.17	0.23	0.25	0.23	0.24	0.30	0.39	0.33
	(2.60)	(1.36)	(2.89)	(3.24)	(2.40)	(2.91)	(3.24)	(4.36)	(2.77)
L/S	0.51	0.53	0.43	0.42	0.52	0.52	0.42	0.45	0.53
	(3.76)	(3.84)	(3.35)	(3.32)	(3.98)	(4.03)	(3.13)	(3.43)	(3.85)

Table A2 Portfolio Returns for Each Single Ratio of Private Information Disclosure

This table reports the average monthly excess returns and alphas on the value-weighted long-short portfolios of stocks sorted by each single ratio of private information disclosure. At the end of June of each year t from 2006 to 2019, individual stocks are sorted into quintiles based on the ratio at the end of year t-1 from 2005 to 2018, and are held for the next twelve months (July of year t to June of year t+1). All returns and alphas are expressed in percentage. Excess return is the raw return of the portfolio over the risk-free rate. Alpha is the intercept from a time-series regression of monthly excess returns on the factors of alternative models. Newey and West (1987) adjusted t-statistics are given in parentheses. The sample period is from July 2006 to December 2019.

Single ratio	Excess Return	CAPM	FFC	FFCPS	FF5	FF6	SY	HXZ	DHS
Revenue	0.46	0.48	0.43	0.44	0.49	0.49	0.38	0.43	0.44
	(2.97)	(3.13)	(2.55)	(2.45)	(2.89)	(2.81)	(2.33)	(2.67)	(2.81)
Total assets	0.42	0.39	0.32	0.35	0.44	0.44	0.36	0.38	0.46
	(3.11)	(2.66)	(2.07)	(2.47)	(2.49)	(2.47)	(2.22)	(2.55)	(2.77)
NumofEmployee	0.40	0.42	0.29	0.27	0.36	0.36	0.32	0.35	0.45
	(2.46)	(2.60)	(1.90)	(1.73)	(2.34)	(2.32)	(2.09)	(2.40)	(2.41)
Income before tax	0.27	0.23	0.15	0.17	0.28	0.28	0.24	0.23	0.22
	(2.09)	(1.73)	(1.18)	(1.27)	(1.90)	(1.95)	(1.73)	(1.84)	(1.51)
Net income	0.42	0.34	0.28	0.29	0.43	0.43	0.35	0.40	0.41
	(1.75)	(1.50)	(0.99)	(1.08)	(1.74)	(1.78)	(1.40)	(1.58)	(1.73)
Cash flow	0.31	0.30	0.22	0.22	0.37	0.37	0.33	0.31	0.39
	(1.97)	(1.83)	(1.27)	(1.30)	(1.96)	(1.97)	(1.88)	(1.88)	(1.97)
Shareholders funds	0.31	0.27	0.19	0.22	0.34	0.34	0.27	0.28	0.31
	(2.43)	(1.96)	(1.64)	(1.66)	(2.07)	(2.07)	(1.79)	(2.02)	(1.99)

Table A3 Alternative Measure of PSID

This table reports the average monthly excess returns and alphas on the value-weighted long-short portfolios of stocks sorted by an alternative measure of PSID, defined as the fraction of number of private subsidiaries that disclose at least one of the seven financial variables relative to the number of private subsidiaries that do not report at all for each public parent firm each year. At the end of June of each year t from 2006 to 2019, individual stocks are sorted into quintiles based on the ratio at the end of year t-1 from 2005 to 2018, and are held for the next twelve months (July of year t to June of year t+1). All returns and alphas are expressed in percentage. Excess return is the raw return of the portfolio over the risk-free rate. Alpha is the intercept from a time-series regression of monthly excess returns on the factors of alternative models. Newey and West (1987) adjusted t-statistics are given in parentheses. The sample period is from July 2006 to December 2019.

Rank	Excess Return	CAPM	FFC	FFCPS	FF5	FF6	SY	HXZ	DHS
Zero	0.28	-0.67	-0.52	-0.49	-0.50	-0.49	-0.42	-0.37	-0.46
	(0.51)	(-1.88)	(-1.51)	(-1.43)	(-1.38)	(-1.38)	(-1.19)	(-1.07)	(-1.24)
P1	0.74	-0.01	-0.01	0.01	-0.11	-0.11	-0.01	0.03	0.03
	(2.16)	(-0.07)	(-0.10)	(0.12)	(-1.00)	(-1.00)	(-0.13)	(0.25)	(0.28)
P2	0.88	0.17	0.13	0.10	0.07	0.07	0.08	0.19	0.08
	(2.70)	(1.62)	(1.24)	(0.89)	(0.64)	(0.63)	(0.71)	(1.69)	(0.70)
P3	0.81	0.07	-0.04	-0.05	-0.05	-0.05	-0.03	0.04	0.07
	(2.39)	(0.68)	(-0.38)	(-0.51)	(-0.47)	(-0.50)	(-0.29)	(0.41)	(0.62)
P4	0.82	0.10	0.03	0.07	-0.02	-0.02	-0.02	0.09	0.09
	(2.50)	(1.05)	(0.35)	(0.71)	(-0.19)	(-0.21)	(-0.27)	(0.92)	(0.88)
P5	1.13	0.40	0.34	0.36	0.27	0.27	0.28	0.41	0.42
	(3.40)	(4.20)	(3.76)	(3.74)	(2.93)	(2.93)	(3.12)	(4.22)	(4.14)
L/S	0.38	0.41	0.35	0.34	0.38	0.38	0.30	0.39	0.39
	(3.11)	(3.35)	(2.74)	(2.62)	(2.70)	(2.68)	(2.29)	(3.04)	(2.76)

Table A4 PSID within Industries

This table reports the pooled mean, standard deviation (Sd), minimum (Min), median (Med), and Maximum (Max) of the private information disclosure ratio (PSID) for firms with at least one private subsidiary in industries according to the Fama-French 48-industry classifications. The information disclosure ratio is defined as the number of private subsidiaries disclosing the particular financial variable divided by the total number of private subsidiaries under the control of a public parent firm. A firm's PSID is the equal-weighted average of the information disclosure ratios from seven financial variables (operating revenue, total assets, number of employees, income before tax, net income, cash flow, and shareholders' funds). The sample period is from 2005 to 2018.

FF48	Industry	Mean	Sd	Min	Med	Max
1	Agriculture	0.09	0.10	0.00	0.05	0.43
2	Food Products	0.15	0.13	0.00	0.13	1.00
3	Candy & Soda	0.16	0.12	0.00	0.14	0.64
4	Beer & Liquor	0.12	0.10	0.00	0.10	0.40
5	Tobacco Products	0.19	0.16	0.00	0.14	0.60
6	Recreation	0.21	0.14	0.00	0.20	0.64
7	Entertainment	0.12	0.12	0.00	0.09	0.86
8	Printing and Publishing	0.15	0.11	0.00	0.13	0.76
9	Consumer Goods	0.21	0.16	0.00	0.18	1.00
10	Apparel	0.19	0.13	0.00	0.17	0.64
11	Healthcare	0.13	0.12	0.00	0.10	0.71
12	Medical Equipment	0.26	0.17	0.00	0.26	1.00
13	Pharmaceutical Products	0.22	0.19	0.00	0.17	1.00
14	Chemicals	0.23	0.15	0.00	0.22	1.00
15	Rubber and Plastic Products	0.24	0.17	0.00	0.22	0.86
16	Textiles	0.19	0.15	0.00	0.17	0.67
17	Construction Materials	0.21	0.15	0.00	0.20	0.86
18	Construction	0.13	0.14	0.00	0.07	0.71
19	Steel Works Etc	0.18	0.13	0.00	0.17	0.69
20	Fabricated Products	0.22	0.13	0.00	0.21	0.57
21	Machinery	0.26	0.14	0.00	0.25	1.00
22	Electrical Equipment	0.22	0.17	0.00	0.19	1.00
23	Automobiles and Trucks	0.22	0.14	0.00	0.21	1.00
24	Aircraft	0.22	0.12	0.00	0.21	0.64
25	Shipbuilding, Railroad Equipment	0.20	0.14	0.00	0.19	0.52
26	Defense	0.17	0.13	0.00	0.14	0.57
27	Precious Metals	0.11	0.10	0.00	0.09	0.43
28	Non-Metallic and Industrial Metal Mining	0.13	0.14	0.00	0.08	0.71
29	Coal	0.10	0.08	0.00	0.08	0.50
30	Petroleum and Natural Gas	0.12	0.13	0.00	0.08	0.86
31	Utilities	0.16	0.14	0.00	0.12	0.86
32	Communication	0.16	0.15	0.00	0.13	1.00
33	Personal Services	0.17	0.15	0.00	0.14	1.00
34	Business Services	0.22	0.16	0.00	0.20	1.00
35	Computers	0.25	0.16	0.00	0.24	1.00
36	Electronic Equipment	0.24	0.17	0.00	0.21	1.00
37	Measuring and Control Equipment	0.28	0.17	0.00	0.27	1.00
38	Business Supplies	0.21	0.15	0.00	0.18	0.88
39	Shipping Containers	0.21	0.12	0.00	0.20	0.57
40	Transportation	0.15	0.14	0.00	0.12	0.86
41	Wholesale	0.19	0.15	0.00	0.16	1.00
42	Retail	0.13	0.13	0.00	0.09	0.76
43	Restaurants, Hotels, Motels	0.10	0.11	0.00	0.07	0.64
44	Banking	0.08	0.10	0.00	0.05	0.62
45	Insurance	0.10	0.11	0.00	0.07	0.86
46	Real Estate	0.08	0.13	0.00	0.03	0.86
47	Trading	0.07	0.11	0.00	0.01	0.86
48	Almost Nothing	0.15	0.16	0.00	0.11	1.00

Table A5 Alternative Factor Models

This table reports the monthly alphas on the value-weighted portfolios of stocks sorted by the PSID. At the end of June of each year t from 2006 to 2019, individual stocks are sorted into quintiles based on non-zero PSID at the end of year t-1 from 2005 to 2018, and are held for the next twelve months (July of year t to June of year t+1). L/S is a zero-cost portfolio that buys stocks in quintile 5 (highest PSID) and sells stocks in quintile 1 (lowest PSID). The alphas, expressed in percentage, are obtained from the time-series regressions of monthly excess returns on the factors of alternative models augmented with the MAX factor (Panel A), the BAB factor (Panel B), and the IVOL factor (Panel C). Newey and West (1987) adjusted t-statistics are given in parentheses. The sample period is from July 2006 to December 2019.

Panel A: Factor models augmented with the BAB factor								
Rank	CAPM	FFC	FFCPS	FF5	FF6	SY	HXZ	DHS
L/S	0.63	0.58	0.57	0.64	0.65	0.46	0.54	0.55
	(3.56)	(3.48)	(3.31)	(3.58)	(3.72)	(2.64)	(3.13)	(3.05)
Panel B: Factor models augmented with the IVOL factor								
Rank	CAPM	FFC	FFCPS	FF5	FF6	SY	HXZ	DHS
L/S	0.67	0.53	0.50	0.67	0.63	0.41	0.51	0.52
	(3.84)	(2.61)	(2.62)	(2.89)	(2.88)	(2.27)	(2.72)	(3.01)
Panel C: Factor models augmented with the MAX factor								
Rank	CAPM	FFC	FFCPS	FF5	FF6	SY	HXZ	DHS
L/S	0.43	0.41	0.43	0.45	0.47	0.38	0.43	0.40
	(2.45)	(2.36)	(2.41)	(2.58)	(2.69)	(2.14)	(2.49)	(2.26)

Chapter 2

Far Away From Home: Investors' Underreaction to Geographically Dispersed Information*

Using a novel data set to identify geographic peer firms based on the locations of both firms' headquarters and material subsidiaries, we show that the returns of geographical peers have strong predictive power for focal firm returns. A value-weighted long-short strategy that buys stocks with highest geo-peer returns and shorts stocks with lowest geo-peer returns generates a Fama and French (2015) five-factor alpha of 0.6% per month. This strategy is distinct from other cross-firm momentum stratigies and cannot be explained by existing asset pricing models. It is more pronounced among firms that receive less investor attention and that are more costly to arbitrage, consistent with slow information diffusion in the geographic network into stock prices.

2.1 Introduction

Geographic locations are essential for the pricing of assets. Stock returns of firms in the same headquarter state tend to move together (Pirinsky and Wang (2006)). Parsons, Sabbatucci, and

*This is a joint work with Liya Chu and Jun Tu

Titman (2020) show that the return comovement of firms headquarterd in the same state extends to a predictable lead-lag effect. The underlying assumption of these two papers is that all economic activities of one specific firm occur in a single place. However, in the real world, many public firms do not just operate in a single location. A public firm may headquarter in a state but operate subsidiaries in other states, and so value-relevant information about the firm is also likely to be geographically dispersed (Bernile, Kumar, and Sulaeman (2015)). Meanwhile, the information flow inside geographically dispersed firms is not as efficient as the geographically concentrated firms (Aarland et al. (2007); Giroud (2013)). Investing in these firms may present challenges to investors with limitations in investors' ability to process value-relevant information, and this may lead to market inefficiencies in the form of gradual incorporation of information into stock prices (Hong and Stein (1999); Hong and Stein (2007)). In particular, we conjecture that valuable information arising from geographically distributed economic activities is slowly incorporated into stocks prices.

In this paper we directly test this hypothesis by examining whether the returns of geographic peers based on the locations of both headquarters and economically relevant subsidiaries are useful for predicting the stocks returns of focal firms. Consistent with our hypothesis, we find that the stock prices of focal firms are adjusted to the shocks of such geographic peers up to four months. Focal firms whose geographic peers experience higher (lower) returns in the current month will earn higher (lower) returns in the next month. A value-weighted trading strategy that buys stocks with highest geo-peer returns and short stocks with lowest geo-peer returns earns a monthly Fama and French (2015) five-factor alpha of around 0.6%. This return predictability is robust to an extensive set of controlling variables and other cross-sectional momentum effects, and is hard to square with risk explanations. We find that the return predicative power is stronger among firms that receive less attention and that are more costly to arbitrage. Therefore, our results are consistent with slow adjustment of value-relevant information from firm's geographic peers to the stock prices of focal firms.

In the empirical analysis, we obtain the locations of firms' headquarters and economically relevant subsidiaries from Exhibit 21, a section within or attached to the 10-K filling that provides the names and locations of firms' headquarters and significant subsidiaries, namely, material subsidiaries.* This data enables us to connect firms using the time-varying locations of both headquarters and material subsidiaries that firms disclose in the Exhibit 21, while the headquarter location variable in Compustat captures only the most recent information. We then construct a proxy of geographic information using returns of geographic peer firms that are linked through the locations of material subsidiaries or headquarters. Specifically, we constructed a frequency-weighted return of state-level portfolio returns of all geo-linked firms (GPRET). The state-level portfolio return is computed as equal-weighted average of stock returns of the geographic peers. The frequency, which captures the economic importance of one state for the focal firm, is the number of times of that state mentioned in the annual reports. By construction, this GPRET captures general good or bad news for the focal firm via aggregating all the relevant information of geographically linked firms. If investors have limited attention or are constrained in processing such kind of information, then the GPRET should predict focal firms' stock returns.

We test the predictive power of the geographic information by forming a trading strategy. Specifically, at the end of each month, we sort firms into deciles based on GPRET. We then form a long-short portfolio that goes long on the decile portfolio with the highest GPRET and shorts on the decile portfolio with the lowest GPRET. After controlling for the Fama and French (2015) five-factor factors, we obtain a 0.45% (t=3.08) monthly abnormal return from an equal-weighted long-short portfolio, or 0.60% (t=2.53) abnormal return from a value-weighted portfolio. We refer to this return predictability as ”geographic momentum”. We further confirm the return predictability of the long-short GPRET portfolio is robust to a number of prominent factor models.

Furthermore, the GPRET effect is more pronounced than a similar geographic momentum effect documented by Parsons, Sabbatucci, and Titman (2020), which find that returns of peer firms headquartered in the same state are significantly associated with focal firm's future returns. If we follow Parsons, Sabbatucci, and Titman (2020) and define geographic peers based on the

*A subsidiary which accounts for more than a certain percentage (usually 5% or 10%) of the consolidated assets or revenues of the parent firm and its subsidiaries is considered as material subsidiary.

locations of headquarters only, an equal-weighted portfolio based on the returns of these peers can earn similar significant returns as in Parsons, Sabbatucci, and Titman (2020), while a value-weighted portfolio cannot cannot generate any significant results. In sharp contrast, if we use the locations of material subsidiaries only to construct the GPRET, the return predictability of this GPRET is still significant when portfolios are value-weighted but sightly weaker compared to the one based on the locations of both headquarters and material subsidiaries. Hence, the geographic momentum effect documented by Parsons, Sabbatucci, and Titman (2020) may be a special case of our more comprehensive GPRET effect.

We also conduct Fama and MacBeth (1973) regressions to confirm that the return predictive power of GPRET is not driven by other firm's characteristics. After controlling for other well-known anomalies, including market beta, size, book-to-market ratio, medium term momentum, short term reversal, gross profitability, asset growth, turnover ratio, idiosyncratic volatility, and illiquidity ratio, the return predictability of GPRET remains significant. In particular, one standard deviation increase of GPRET leads to 0.09% increase in the focal firm's return, which is also economically significant as well. To further distinguish the GPRET effect from previous findings, we follow Moskowitz and Grinblatt (1999) and Daniel and Titman (1997) to control for effects related to industry and firm characteristics. Specifically, we adjust stock returns using the industry returns and DGTW matched portfolio's returns in our regression. We obtain similar and significant coefficients on the GPRET in these settings, suggesting that the GPRET return predictive power is not driven by these two effects.

To ensure that our geographic momentum effect is not a rediscovery of other well-known cross-sectional momentum effects, such as industry momentum effect (Moskowitz and Grinblatt (1999)), customer and supplier industry momentum effects (Menzly and Ozbas (2010)), complicated firm momentum (Cohen and Lou (2012)), technology momentum effect (Lee et al. (2019)), connected-firm momentum effect (Ali and Hirshleifer (2020)), common board momentum effect (Burt, Hrdlicka, and Harford (2020)), and common instutional momentum effect (Gao, Moulton, and Ng (2017)), we also conduct Fama and MacBeth (1973) regressions controlling for these momentum effects directly. In general, controlling for these momentum effects, especially connected-firm momentum does not have any significant effect on our geographic momentum effect, while the geographic momentum effect based on locations of headquarters is explained by the connected-firm momentum effect according to Ali and Hirshleifer (2020). The only exception is the customer momentum effect, and one possible reason is that the sample size becomes too small to have meaningful results if we include the customer return in our regression.

To better understand the underlying mechanism of the GPRET effect, we examine the sensitivity of the GPRET effect to firm-specific characteristics associated with investors' inattention and costs of arbitrage. We find that the return predictability is more pronounced among focal firms that receive less attention (i.e., firms that have lower media coverage and analyst coverage), and that are more costly to arbitrage (i.e., firms that have higher idiosyncratic volatility and higher illiquidity ratio). These results are consistent with sluggish price adjustment of information embedded in the geographic network.

We also examine the stock price reaction around subsequent earnings announcements in one-day and three-day windows. This test has been widely used in previous literature to separate mispricing from risk-based explanations (e.g., Sloan (1996); La Porta et al. (1997); Engelberg, McLean, and Pontiff (2018)). The idea is straightforward. Earnings announcements help investors update expectation biases. If the predictable returns are driven by changes in potential risks, we expect that the returns are evenly affected in different trading days. In contrast, if the return predictability are due to mispricing, it will be stronger when it is closer to the earnings announcement days. We observe that roughly 16% of abnormal returns of the GPRET strategy are realized in the three-day window around earnings announcement dates. This result is hard to square with the risk-based explanation.

Our paper contributes to two strands of existing literature. First, our paper is related to literature on the relationship between geographic locations and stock returns. Pirinsky and Wang (2006) document that stock returns commove between companies headquartered in the same geographic area. Garcia and Norli (2012) show that truly local firms earn significantly higher returns than geographically dispersed firms. Korniotis and Kumar (2013) and Korniotis and Kumar (2013) find that local economic conditions or business cycles have predictability for local stock returns. Our paper is also closely related to Parsons, Sabbatucci, and Titman (2020), who document a geographic lead-lag effect using information about locations of firms' headquarters. However, our paper differs from their paper in several important respects. First, Parsons, Sabbatucci, and Titman (2020) identify firms' locations using ZIP codes from Compustat, which only contains most recent information but ignores changes over time, and thus their GPRET can be inaccurate if a firm changes its headquarter location. In contrast, we retrieve the time-varying location information from firms' 10-k annual reports such that we are able to identify all relevant locations accurately and timely. In addition, as shown by Bernile, Kumar, and Sulaeman (2015) that a typical U.S. public firm has economic interests in five states beyond its corporate headquarters location, we find that our GPRET, which incorporates locations beyond firms' headquarter locations, is able to capture more value-relevant information from the distant locations in which firms' economically relevant subsidiaries are located. Therefore, another key innovation in our paper is that we show significant results for both equal- and value-weighted portfolios using our GPRET, while results using the headquarter-based GPRET like the one in Parsons, Sabbatucci, and Titman (2020) are only significant for equal-weighted portfolios.

Our paper is also related to an emerging literature on the cross-sectional return predictability. Studies in this area depart from the rational expectations framework and have utilized behavioral explanations, combined with market frictions, to rationalize predictability of returns (Nagel (2013)). Alongside various theoretical explanations of return predictability in this behavioral framework (Shleifer and Vishny (1997); Hong and Stein (1999)), a growing number of studies offer empirical explanations as to how information of linked firms may be slowly incorporated into stock prices of focal firms. Moskowitz and Grinblatt (1999) find that past industry return is associated with future stock returns after controlling for individual stock momentum. Cohen and Frazzini (2008) find that past return of firms' principal customers forecasts future returns of focal firms after controlling for industry and cross-industry momentum effects. Similarly, using data on flow of goods to and from industries, Menzly and Ozbas (2010) find that past customer and supplier industry returns predict future stocks returns. Cohen and Lou (2012) use operating segment data to show that single-segment firm returns can predict returns of multi-segment firms operating in the same industries. Lee et al. (2019) find a lead-lag relation in returns between technology-linked firms. Parsons, Sabbatucci, and Titman (2020) find that return comovement of firms headquartered in the same state extends to a lead-lag effect. Recently, Ali and Hirshleifer (2020) find that these momentum spillover effects are a unified phenomenon captured by shared analyst coverage. A basic theme of these papers is that investor are subject to limited attention and therefore not table to process information about linked firms in a timely fashion. Our paper also contributes to this framework by showing that limited investor attention causes delayed stock price responses to the value-relevant information embedded in the geographic network.

The remainder of this paper is organized as follows. Section 2 describes data and variables. Section 3 presents our main results. Section 4 explores the underlying mechanism behind the GPRET effect. Section 5 reports results on earnings announcement returns prediction. Section 6 concludes.

2.2 Data and variables

The information about locations of firms' headquarters and material subsidiaries is collected from Exhibit 21, a section within or attached to the annual 10-K report.[*] Public firms in the U.S. are required to reveal information about the locations of their headquarters and foreign or domestic material subsidiaries in Exhibit 21. Dyreng, Lindsey, and Thornock (2013) gather and compile this Exhibit 21 data using a text search program.[†] Specifically, this program identifies states in which firms' headquarters or domestic material subsidiaries locate and counts number of times that each state appears in Exhibit 21 for each company in each year. Therefore, this data set allows us to identify a time series of firm's time-varying headquarter locations as well as material subsidiaries'

[*]The Exhibit 21 filings are available at the Securities and Exchange Commission (SEC) webpage (www.sec.gov)
[†]The data is publicly available here: https://sites.google.com/site/scottdyreng/Home/data-and-code.

locations, while the headquarter variable in Compustat captures the latest information only. The sample period of this tabulated data is from 1993 to 2014. To investigate the return predictability of the geo-peers' returns, we use the CIK number of this data to match with Compustat accounting information and therefore with CRSP monthly return data.

Different from Parsons, Sabbatucci, and Titman (2020), which define geographic peers based on the locations of firms' headquarters only, our paper uses information not only about the locations of firms' headquarters but also about the locations of firms' material subsidiaries to identify geographically linked peers.* Therefore, to test our hypothesis, we define our geographic peers' return (GPRET) for each focal firm that owns at least one material subsidiaries as the weighted-average portfolio return of related state-level portfolio returns of all geo-linked firms. The frequency of each state appears in the Exhibit 21 is used as portfolio weight. States in which more headquarters or material subsidiaries operate are more economically important to the parent firm, so they get higher weights. More specifically, first, for each firm i, we compute the state-level portfolio return as the equal-weighted return of all geo-linked peers' stock returns in each state in which firm i's headquarter or material subsidiaries locate as following:

$$GPSRET_{i,j,t} = 1/n \sum_{k=1}^{n} RET_{i,j,k,t} \tag{2.1}$$

where j denotes a state in which the material subsidiaries or the headquarter of firm i locate, k denotes firm k whose headquarter or material subsidiaries that also locate in state j, $RET_{i,j,k,t}$ is the stock return of firm k at month t. We also compute a value-weighted state-level portfolio return in our empirical tests. However, we only focus on the equal-weighted method to give more weights to smaller firms because large firms are less likely subject to limited attention and high arbitrage costs, which shocks are more likely to be priced efficiently with little underreaction.

Next, we calculate the frequency-weighted portfolio return of state-level portfolio returns for each firm i as following:

*We also define geographic peers based on the locations of headquarters and the locations of material subsidiaries, respectively. We use the one based on both the headquarters and the material subsidiaries as our main variable in this paper.

$$GPRET_{i,t} = \frac{\sum_{j=1}^{m} GPSRET_{i,j,t} * NUM_{i,j,t}}{\sum_{j=1}^{m} NUM_{i,j,t}} \tag{2.2}$$

where m denotes the number of states in which firm i's headquarter or subsidiaries locate, $NUM_{i,j,t}$ is the number of times state j appears in the Exhibit 21 data for firm i at month t.

We collect monthly stock returns from the Center for Research in Security Prices (CRSP) and accounting data from Compustat. Our sample includes all firms listed on NYSE, Nasdaq, and Amex. We keep common stocks (SHRCD equal to 10 or 11) and exclude financial firms (SIC codes between 6000 and 6999) and utilities firms (two-digit SIC codes starting with 49). To reduce the effect of micro-cap firms, we exclude firms with stock prices below five dollars per share at the end of each month. We follow Shumway (1997) to adjust stock returns for delisting. If a delisting return is missing and the delisting event is performance-related, we set the delisting return as -30%. To ensure that the geographic peers' information and other account information are fully available to investors, we skip six months at the end of each year to form our portfolios. In other words, we match the geographic peer firms' information in year t-1 to the monthly stock returns from July of year t to June of year t+1.

In the regression analysis, we also control for other firm-specific characteristics. Specifically, following Fama and French (1992), we estimate the market beta (MKT_BETA) of individual stocks using monthly returns over the past 60 months. The size (SIZE) is computed as the logarithm of the market value of the firm's outstanding equity at the end of month t. The book-tom-market ratio (BM) is the logarithm of firm's book value of equity divided by its market capitalization following Fama and French (1992). Momentum (MOM) is the stock's cumulative return from the start of month t-12 to the end of month t-2 following Jegadeesh and Titman (1993). Following Jegadeesh and Titman (1993), short-term reversal (STRV) is measured as the stock return over the prior month. Gross profitability (GP) is the firm's gross profitability following Novy-Marx (2013). Asset Growth (AG) is percentage of total asset growth between two consecutive fiscal years following Cooper, Gulen, and Schill (2008). TURNOVER is the monthly turnover ratio in month t. ILLIQ is the monthly illiquidity ratio following Amihud (2002) using daily data in month

t. IVOL is the idiosyncratic volatility following Ang et al. (2006) in month t.

The final sample consists of 308215 firm-month observations from July 1994 to June 2016. Panel A of Table 2.1 reports the summary statistics for our sample firms. The number of firms varies from 420 in June 1994 to 1887 in July 1998. The average number of states that are linked to one focal firm is around 3, indicating that many public firms are operating in 3 states on average. In terms of the market capitalization, these firms account for approximately 53% of the market capitalization of the CRSP common stocks universe, suggesting that the proportion of firms that have at least one material subsidiaries is relatively high. The main variable of interest, GPRET, has a mean of 0.01% and a standard deviation of 0.05, which indicates relatively large variation across our sample.

Panel B of Table 2.1 presents the correlations (Pearson and Spearman) among the variables we use. As we can see, the correlation between GPRET and RET_{t+1} is 0.01, which implies that GPRET probably contains additional valuable information about future focal firms' returns. GPRET has low correlations with control variables with absolute values lower than 0.07, which suggests that our GPRET is not highly correlated with other prominent anomalies.

2.3 Empirical results

2.3.1 Portfolio tests

We differentiate our study from Parsons, Sabbatucci, and Titman (2020) by considering the locations of headquarters and material subsidiaries of a focal firm, rather than only headquarters, and by showing that Parsons, Sabbatucci, and Titman (2020)'s geographic lead-lag effect is a special case of our GPRET effect. To this end, we conduct three portfolio tests using returns of geo-peers defined based on firms' headquarters, material subsidiaries, and both, respectively. Specifically, at the end of each month, we sort stocks based on GPRET into deciles and calculate equal- and value-weighted returns of these decile portfolios in the next month. These portfolios are rebalanced at the end of each month. We also form a long-short portfolio that longs the top-decile stocks with highest GPRET and shorts the bottom-decile stocks with lowest GPRET. We calculate these returns by subtracting the risk-free rate (excess returns) or by using a a variety of factor models including Fama and French (1993) three-factor model (FF3), Fama and French (1993) three factors augmented with Carhart (1997) momentum factor (FFC), Fama and French (2015) five-factor model (FF5), Fama and French (2018) six-factor model, Hou, Xue, and Zhang (2015) q-factor model (HXZ), Stambaugh and Yuan (2017) mispricing-factor model (SY), and Daniel, Hirshleifer, and Sun (2020) behavioral-factor model (DHS).

Table 2.2 reports the main results of portfolios sorted on GPRET. Geo-peers, in this setting, are defined as firms whose headquarters or material subsidiaries are located in the same state as the focal firm's headquarter or material subsidiaries. Panels A and B present equal- and value-weighted portfolio results, respectively. One notable point in Table 2.2 is that the next-month average excess return increases monotonically from 0.64% to 1.07% per month for the equal-weighted portfolio and from 0.43% to 1.00% per month for the value-weighted portfolio, when moving from the lowest to the highest GPRET decile. The average return difference between decile 10 (high-GPRET) and decile 1 (low-GPRET) is 0.43% per month with a t-statistic of 3.00 for the equal-weighted portfolio and 0.57% per month with a t-statistic of 2.30 for the value-weighted portfolio. These results indicate that stocks in the highest GPRET generate 5.16% (6.84%) annual returns compared to stocks in the lowest GPRET decile for the equal- (value-) weighted portfolios.

In addition to the average excess returns, Table 2.2 presents the magnitude and statistical significance of the risk-adjusted returns (alphas) from a variety of asset pricing models. As shown in Table 2.2, alphas adjusted by factor models increase monotonically when moving from the lowest to the highest GPRET decile. Besides, the long-short alphas are persistent when using different factor models to adjust the portfolio returns. For the equal-weighted long-short portfolio, the largest alpha is generated by DHS, 0.46% (t=3.04), and the smallest one is by FFC, 0.38% (t=2.70). For the value-weighted long-short portfolio, HXZ produces the largest alpha of 0.65% (t=2.72) and FF3 generates the smallest alpha of 0.43% (t=1.90). All equal-weighted long-short portfolios have economically and statistically significant alphas, with t-statistics all above

2.70. Results for the value-weighted long-short portfolio are slightly weak in terms of statistical significance with t-statistics greater than 1.90, but larger in terms of economic significance as magnitude of alphas are larger. Overall, these results strongly support our hypothesis that GPRET indeed contains value-relevant information on predicting future focal firm returns.

We also examine how the returns of the geographic momentum strategy vary over time. Figure 2.1 plots the excess returns of the value-weighted long-short portfolio on a per annum basis from 1994 to 2016. The long-short portfolio's returns are negative only in five out of the 23 years. When the Internet bubble in 2000 and Subprime crisis in 2008 burst, the geographic momentum strategy performed well, with a return of 0.87% in 2000 and 4.68% in 2008, suggesting that this strategy appears to be a good hedge against market extreme downturns.

Results for the equal-weighted portfolios in Table 2.2 are consistent with results of Table 5 in Parsons, Sabbatucci, and Titman (2020), which presents results of equal-weighted portfolios sorted on the equal-weighted lagged return of firms headquartered in the same city, outside its industry. In Table 5 of Parsons, Sabbatucci, and Titman (2020), such a geographic momentum strategy generates an equal-weighted average excess return of 0.42% per month (t=3.65). The alpha differences between the high geo-peer return portfolios and the low geo-peer return portfolios are positive and statistically significant: 0.47% per month (t=4.16) adjusted by CAPM model; 0.46% per month (t=3.99) adjusted by FF3 model; 0.42% per month (t=3.40) adjusted by FFC model; and 0.42% per month (t=3.32) adjusted by FF5 model. To further compare our results with Parsons, Sabbatucci, and Titman (2020), we also conduct univariate portfolio analysis based on geo-peer return where the geographic peers are defined using information about firms' headquarters only.

Table 2.3 reports the results. For the equal-weighted portfolio, the long-short portfolio alphas are generally larger than those in Table 2.2 by about 0.1%. For instance, the FFC alpha of long-short portfolio is 0.49%, while it is 0.38% in Table 2. In sharp contrast, the value-weighted long-short portfolio alphas are generally smaller than those in the Panel B of Table 2.2. Specifically, the long-short excess return is smaller by 0.12% than that in the Panel B of Table 2.2. In addition, the results show that the GPRET effect based on headquarters is not robust to the way portfolios are weighted. The equal-weighted long-short portfolio excess returns or alphas are all significant, with the smallest alpha of 0.38% (t=2.01) from SY. In sharp contrast, the largest value-weighted long-short portfolio alpha is 0.47% with the largest t-statistics of 1.56. These results indicate that the geographic momentum effect is not robust when portfolios are value-weighted, and may also suggest the reason why value-weighted portfolio results are not presented in Parsons, Sabbatucci, and Titman (2020).

We also examine the return predictability of GPRET if geographic peers are defined based on locations of material subsidiaries only. Table 2.4 reports the portfolio results. Similar to Table 2.2 and Table 2.3, the average excess returns and alphas increase monotonically from decile 1 to decile 10 for both equal- and value-weighted portfolios. We find that for equal-weighted portfolios, the GPRET effect based on material subsidiaries in Table 2.4 is slightly weaker than the GPRET effect based on headquarters in Table 2.3 except alphas from Stambaugh and Yuan (2017) mispricing-factor model and Daniel, Hirshleifer, and Sun (2020) behavioral-factor model. When the portfolios are value-weighted, the GPRET effect based on material subsidiaries is much more stronger than the GPRET effect based on headquarters regarding magnitude of alphas. Compared with the headquarter-based GPRET, the long-short excess returns or alphas produced by subsidiary-based GPRET are larger by a range from 0.07% under FFC to 0.3% under SY, and are all statistically significant. This suggests that the GPRET effect based on material subsidiaries is robust when portfolios are value-weighted. Finally, when comparing portfolio results in Table 2.4 with results in Table 2.2, we find that in general the value-weighted portfolio performs slightly better when we combine the locations of headquarters and material subsidiaries together to define geographic peers than the value-weighted portfolio using locations of material subsidiaries with respect to economic significance and statistical significance.

In light of these results, the geographic momentum in Parsons, Sabbatucci, and Titman (2020) can be seen as a special case of our geographic momentum effect as a value-weighted portfolio based on our GPRET can generate significant alphas, indicating that firms' material subsidiaries contain more valuable information in predicting future returns.

We present the factor loadings on each of the factor models used in Table 2.2 of the geographic momentum portfolio in Panels A and B of Table 2.5. The equal- and value-weighted long-short portfolios are neutral with respect to most of the factors, as the loadings on them are insignificant. The only exception is the loading on size factor (SMB), which is positive and highly significant. This suggests that the geographic momentum strategy will perform well when small firms do well. The long-short portfolio also has negative factor loadings of -0.09 (t=-2.31, equal-weighted) and -0.27 (t=-2.77, value-weighted) for MGMT, and loadings of -0.11 (t=-2.14, equal-weighted) and -0.26 (t=-2.43, value-weighted) for FIN, suggesting that the geographic momentum strategy is highly loaded on mispricing or behavioral factors that are more related to managers' decisions, as the MGMT factor of Stambaugh and Yuan (2017) arises from six anomalies that represent quantities that firm managements can affect directly, and the FIN factor of Daniel, Hirshleifer, and Sun (2020) exploits information in managers' decisions to issue or repurchase equity in response to persistent mispricing. But even after controlling for these exposures, the geographic momentum strategy still generates significant monthly alphas.

2.3.2 Firm size and the effect of GPRET

Size is related to an extensive set of return predictors. Thus, the return predictive power of our GPRET may be driven by the small stocks even though the value-weighted abnormal returns are significant in Table 2.2. We further examine our results by controlling for size. Specifically, at the end of each month, we first sort all stocks into quintiles based on the market capitalization. Within each size quintile, we further sort stocks into quintiles based on the GPRET. All 25 portfolios are rebalanced monthly, and next-month Fama and French (2018) fix-factor alphas are calculated for each of these 25 portfolios. Table 2.6 reports the results for the 25 portfolios: equal-weighted alphas in Panel A and value-weighted alphas in Panel B. We also report the alpha for each size quintile of the long-short GPRET portfolio. The results show that the GPRET effect exists in all five size quintiles. The effect is weaker among larger firms (quintile 4 and quintile 5) than among smaller firms (quintile 1 and quintile 2) for both equal- and value-weighted portfolios, although the effect is not monotonic with respect to size. This suggests that the GPRET effect is not entirely driven by the small size effect.

2.3.3 Fama-MacBeth regression

We test the return predictability of the GPRET by conducting Fama and MacBeth (1973) regressions of one-month-ahead stocks returns on GPRET and various firm-specific control variables. We winsorize all variables at the 1% and 99% levels and standardize all independent variables to zero mean and standard deviation of one. We run the regressions every calendar month and adjust the standard errors for heteroscedasticity and autocorrelation following Newey and West (1987).

Table 2.7 reports the time-series averages of the slope coefficients and the Newey and West (1987) t-statistics from the Fama and MacBeth (1973) regressions. In Column 1 where all the control variables are included, consistent with the portfolio tests, the GPRET is positively and significantly associated with the one-month-ahead return, with an average slope of 0.09 (t=3.72). In addition, the average slopes of these control variables are also consistent with prior literature: market beta, size, short-term reversal, asset growth, idiosyncratic volatility, and illiquidity ratio are negatively correlated with future returns, and book-to-market, momentum, and gross profitability are positively correlated with future stock returns.

In Column 2, we control for industry fixed effects by including Fama and French (1997) 48 industry dummies in the regression as controls. The average slope on GPRET is 0.09 with a Newey and West (1987) t-statistic of 3.90, suggesting that controlling for industry fixed effects has almost no influence on the economic significance of the cross-sectional relation between the GPRET and future stock returns. In Column 3, we further control for industry fixed effects by using industry-adjusted return (difference between a focal firm's return and the value-weighted portfolio return of firms in the same industry based on Fama and French (1997) 48 industry classifications) as the dependent variable. We purge out any predictability arising from industry momentum by subtracting industry return from the focal firm return. The average slope of GPRET again remains positive, 0.09, and highly significant with a Newey and West (1987) t-statistic of 4.20.

Previous literature suggests that firm characteristics may provide better predictability for future returns (Daniel and Titman (1997)). Thus, following Daniel and Titman (1997), we use focal stocks' returns less characteristics-matched benchmark portfolio returns (which are formed from 5-by-5-by-5 independent triple sorts on size, book-to-market, and momentum) as our dependent variable in the regression. In this specification, the average slope of GPRET in Column 4 is positive and highly significant, and has a similar magnitude to those in Columns 1-3. Overall, these results indicate that GPRET provides incremental information about future stock returns relative to various firm-specific characteristics.

2.3.4 Longer-term performance

Previous literature documents that the return predictability is driven by the slow diffusion of value-related information as a result of underreaction (Lee et al. (2019); Cohen and Frazzini (2008); Hong and Stein (1999); Hong, Lim, and Stein (2000)). Thus, we further examine whether the return predictive power of GPRET exists over the next twelve months after portfolio formation. The results are presented in Table 2.8. During the first four months, the average slopes of GPRET remain positive and significant controlling for all firm characteristics. The predictive power of GPRET on future stock returns diminishes as one moves further away from the portfolio formation month and becomes insignificant after the fifth month. These results suggest that the negative cross-sectional relation between GPRET and future stock returns is not just a one-month affair and the underreaction to such informatin persists several months after portfolio formation as a consequence of the slow information incorporation process.

2.3.5 Robustness

To further distinguish our geographic momentum effect from other cross-sectional momentum effects, we conduct several Fama and MacBeth (1973) regressions by controlling for well-known economic links. We first include a similar geographic momentum (GPRET_HQ) documented by

Parsons, Sabbatucci, and Titman (2020), which use headquarter locations to define geographic peers rather than locations of both headquarters and material subsidiaries. Moskowitz and Grinblatt (1999) show that industry momentum is a strong return predictor that subsumes the return predictability of individual stock momentum. To measure focal firm's industry return (INDRET), we classify stocks into industries based on Fama and French (1997) 48 industry classifications, and we compute focal firm's industry return as the value-weighted average return of all other stocks in the same industry. Following Ali and Hirshleifer (2020), the CF return of focal firm is calculated as the weighted average return of stocks that are connected through shared analyst coverage, and the number of analyst covering both stocks is used as weights. Analyst coverage is defined as the number of analysts who issue at least one FY1 or FY2 earnings forecast for the stock over the past 12 months. Customer (Supplier) industry return is the weighted average return of all the industries that buy from (supply to) that stock's industry following Menzly and Ozbas (2010). The portfolio weights are measured by the flow of goods to and from industries extracted from Bureau of Economic Analysis (BEA) Input-Output data (at the summary industry level). Single-segment return is the weighted average return of single-segment firms operating in the same segments as the conglomerate firm following Cohen and Lou (2012). We use the percentage of sales belonging to each segment within the conglomerate as the portfolio weights. We compute a technology momentum return (Tech. linked RET) as the weighted average return of technologically similar firms following Lee et al. (2019). The portfolio weights are the technological closeness of firm pairs. The sample period of Tech. linked RET is from July 1994 to June 2012 due to the availability of Kogan et al. (2017) patent data. Following Burt, Hrdlicka, and Harford (2020), we define a common board momentum (Board linked RET) as the weighted average return of stocks linked through common board members. Finally, we also calculate a common institutional investors momentum (Inst. linked RET) as the weighted avereage return of stocks connected through common institutional investors with the focal firm following Gao, Moulton, and Ng (2017). We include the same control variables as in Table 2.7, but we do not report their coefficients for brevity. All independent variables are normalized to zero mean and standard deviation of one after winsorization at the 1% and 99% levels. Newey and West (1987) adjusted t-statistics are shown below the coefficient estimates in parentheses.

Table 2.9 presents the results of these regressions. In Column 1, we include GPRET_ALL and GPRET_HQ in the regression. GPRET_ALL is the our geo-peer return which are defined based on locations of both headquarters and material subsidiaries. Consistent with the results in previous section, both average slopes of GPRET_ALL and GPRET_HQ are postive and statistically significant. However, the average slope of GPRET_HQ is much lower than the one of GPRET_ALL and is just significant at 10% significance level, suggesting that our geo-peer return (GPRET_ALL) indeed provides additional valuable information relative to the geo-peer return (GPRET_HQ) by Parsons, Sabbatucci, and Titman (2020). In Column 2, we add industry return in the regression, and results show that both geo-peer returns and industry return are individually strong predictors of future returns, consistent with the results in Table 2.7. Column 3 includes GPRET_ALL, GPRET_HQ, and CF return in the regression. The average slope of GPRET is unchanged, 0.10 and highly significant with a Newey and West (1987) t-statistic of 3.41 after controlling for CF return. Consistent with the results of Table 5 in Ali and Hirshleifer (2020) where the average slope of the geographic momentum based on the locations of headquarters only decreases substantially and becomes insignificant once CF return is included in the regression, the average slope of GPRET_HQ is only 0.02 and insignificant. This suggests that our geo-peer return is a better return predictor compared to the one in Parsons, Sabbatucci, and Titman (2020).

The next two columns examine return predictability of GPRET when we include controls related to focal firm's supply chain. Columns 4 and 5 show that the average slopes of GPRET_ALL become slightly smaller but still remain statistically significant after controlling for customer and supplier industry returns. In contrast, the GPRET_HQ is not significant. On the other hand, while both of customer and supplier returns are strong predictors in their own samples, their average slopes become insignificant once we include the GPRET_ALL in the regressions.

Column 6 examines the return predictive power of GPRET controlling for firm's pseudo-conglomerate return. It shows that the return predictability of GPRET_ALL still remains significant but decreases slightly after including single-segment return. In Column 7, the return predictive power of GPRET also becomes smaller and marginally significant once we include the technology momentum return in the regression. In both cases, the average slope of GPRET_HQ is still insignificant. In Column 8 and Column 9, we add common board momentum return (Board linked RET) and common institutional investor momentum (Inst. linked RET) in the regression. The result shows that the return predictive power of GPRET_ALL remain significant while the average slopea of GPRET_HQ are still insignificant.

Overall, the cross-sectional regression tests show that the our GPRET_ALL is a robust return predictor controlling for many other cross-firm momentum effects. In addition, the GPRET_ALL provides additional information relative to the GPRET_HQ in predicting future stocks returns. This implies that geographic locations of material subsidiaries are equally important in defining geographic peers of the focal firm.

2.4 Underlying mechanism

Having established that the geographic momentum is a strong return predictor, we further explore the mechanisms affecting the information incorporation process. In this section, we perform double sort analysis to test the cross-sectional sensitivity of the GPRET effect to various firm-specific characteristics associated with the investors' inattention and the costs of arbitrage. Specifically, we first group all firms into terciles based on each firm-specific characteristic X. Then, we independently sort firms into quintiles based on the GPRET. These portfolios are value-weighted and rebalanced on a monthly basis. We form a high-minus-low GPRET portfolio within each X tercile. We calculate these returns by subtracting the risk-free rate (excess returns) or by using Fama and French (2015) five-factor and Fama and French (2018) six-factor models. We also conduct a test on whether the long-short alphas differ between the two extreme X terciles.

2.4.1 Investors' limited attention

With regard to the predictability of the GPRET, one possible source could be investors' inattention. If investors' limited attention plays a significant role here, we should expect that the return predictive power should be more pronounced for firms that receive less investor attention. Following prior literature, we choose media coverage (Fang and Peress (2009); Twedt (2016)) and analyst coverage (Hong, Lim, and Stein (2000); Jiang, Qian, and Yao (2016)) as the proxies for limited attention. Twedt (2016)) use media coverage as a proxy for investor attention as investors tend to pay attention to media news. Jiang, Qian, and Yao (2016) report that abnormal returns are concentrated in firms that receive less attention in terms of analyst coverage.

To construct the media coverage, we use news data from Thomson Reuters News Analytics. The media coverage is defined as the number of news articles covering the stock in a certain month. Because this measure is highly correlated with size, we orthogonalize this raw number with respect to size to tease out the confounding size effect. Specifically, in each month we regress the logarithm of the number of news on the logarithm of size and extract the regression residual as the measure of media coverage. Similarly, analyst coverage is computed as the residual analyst coverage orthogonalized with respect to the firm size following Hong, Lim, and Stein (2000). The raw number of analyst coverage is defined number of analysts that issued at least one annual earnings forecast during the past 12-month period, using the IBES detail historical file. If the number of analyst coverage or the number of media coverage are missing, we set them to zero.

Panel A of Table 2.10 shows that the return predictability of the GPRET is stronger for portfolios that consist of stocks receiving less investor attention. Specifically, for subsamples split by residual media coverage, the long-short GPRET portfolio in the low residual media coverage group earns an average excess return of 0.68% (t=2.42), while the one in the high residual media coverage earns just 0.01% (t=0.03). Similarly, the long-short GPRET portfolio earns significant Fama and French (2015) five-factor and Fama and French (2018) six-factor alphas in the low residual media coverage group but insignificant alphas in the high residual media coverage group.

However, the differences of the long-short GPRET portfolios between the high and the low media coverage groups are not significant. With respective to subsamples split by the residual analyst coverage, the long-short GPRET portfolio also earns positive and significant excess returns and alphas in low residual analyst coverage group, and the one in high residual analyst coverage group earns negative profits. In addition, the differences of the long-short GPRET portfolios between the high and the low residual analyst coverage groups are statistically significant. Therefore, consistent with the limited attention argument, firms that attract less investor attention exhibit more pronounced predictable returns.

2.4.2 Limits to arbitrage

The evidence so far suggests that the GPRET effect is mostly consistent with mispricing. Therefore, we should expect the return spread is the largest for those stocks that most difficult to value or the most difficult to arbitrage (Shleifer and Vishny (1997)). The findings in Table 2.6 show that the GPRET effect is stronger for small firms than for large firms, consistent with the limits to arbitrage. We further explore how the GPRET effect varies with other measures of limits to arbitrage. Following the literature, we use idiosyncratic volatility (Ang et al. (2006)) (IVOL) and Amihud's illiquidity (Amihud (2002)) (ILLIQ) as the proxies of limits to arbitrage.

Panel B of Table 2.10 presents the results. In general, the long-short GPRET portfolio earns positive and significant excess returns and alphas in subsamples with high levels of limits to arbitrage measured by IVOL and ILLIQ. More importantly, the differences in the long-short GPRET strategy between the high and the low IVOL (ILLIQ) groups are statistically significant. Therefore, the results are consistent with our hypothesis that the geographic momentum effect is stronger for difficult-to-arbitrage stocks.

2.5 Earnings announcement returns prediction

Results so far are more consistent with a mispricing interpretation of the GPRET effect. To further examine whether the return predictability of GPRET is consistent with the mispricing argument, we conduct Fama and MacBeth (1973) regressions using the cumulative abnormal returns around the earnings announcement date as the dependent variable to examine stock returns around earnings announcements after portfolio formation. This is a widely used method to examine whether anomalies are the result of biased expectations (Sloan (1996); La Porta et al. (1997); Engelberg, McLean, and Pontiff (2018)). If the GPRET effect is driven by changes in potential risks, the mean returns on earnings announcement days (EADs) should be similar to the mean returns on non-EADs. In contrast, if the return predictability of the GPRET is due to mispricing, the EAD returns will higher (lower) than the non-EAD returns for high-GPRET (low-GPRET) stock as investors are surprised by the subsequent unanticipated good (bad) news.

We obtain quarterly earnings announcement dates from the I/B/E/S and the quarterly Compustat database. Following DellaVigna and Pollet (2009), we keep the earlier of the two dates when dates from Compustat and I/B/E/S are not in accordance. We calculate one- and three-day cumulative abnormal returns (CAR) during announcement periods, denoted by CAR(t) and CAR(t-1, t+1). To get robust results, we use three methods to calculate CAR, including CRSPVW-adjusted CAR, DGTW-adjusted CAR and SBM-adjusted CAR. The CRSPVW-adjusted CAR is calculated as the sum of the daily stock returns minus returns on the CRSP value-weighted portfolio. The DGTW-adjusted CAR is the sum of the daily stock returns minus returns of the characteristics-matched portfolio following Daniel and Titman (1997). SBM-adjusted CAR is the difference between the buy-and-hold return of the stock and that of a size and book-to-market (B/M) matched portfolio over the same window following Hirshleifer, Lim, and Teoh (2009). We then run Fama and MacBeth (1973) regressions of these CARs on the GPRET and various firm-specific characteristics that are observed in last month before the earnings announcement date. All independent variables are normalized to zero mean and standard deviation of one after winsorization at the 1% and 99% levels. Newey and West (1987) adjusted t-statistics are shown below the coefficient estimates in parentheses.

Table 2.11 reports the regression results. Consistent with the mispricing explanation, returns to the GPRET strategy are much larger during future earnings news releases as the GPRET is positively and significantly associated with future CARs. For instance, the average slope coefficient on the GPRET is 0.08 (t= 2.83) when using CRSPVW-adjusted CAR(t-1, t+1) as the dependent variable. On average, the standardized difference in the GPRET between decile 1 and decile 10 is around 3.51. This suggests that the return spread between two extreme GPRET portfolios during three-day window is around 0.28% (3.51*0.08), compared to a monthly abnormal return spread of approximately 0.60% including all trading days. This indicates that roughly 16% of abnormal returns of the GPRET strategy are achieved in the three-day window around EADs. The significant higher returns during earnings announcement dates are consistent with the literature. Engelberg, McLean, and Pontiff (2018) find that daily anomaly returns are six times higher on EADs and three times higher in the three days around EADs, relative to non-EADs. This suggests that around one-tenth to one-sixth of anomaly returns are realized around EADs. The results provide support to the mispricing explanation that investors do not fully incorporate the GPRET information into their earnings forecasts and are therefore surprised when earnings news are released.

2.6 Conclusion

Using a novel data set to identify time-varying locations of firms' headquarters and material subsidiaries, we establish that geographically linked firms' returns predict focal firm returns. This geographic momentum effect is robust regardless of whether the strategy is equal-weighted or value-weighted. It is also robust to a variety of firm-specific characteristics including: market beta, firm size, book-to-market, momentum, short-term return reversal, gross profitability, asset growth, turnover ratio, idiosyncratic volatility, and illiquidity ratio. It is distinct from, and cannot be explained by other cross-sectional momentum effects such as industry momentum, customer momentum, customer or supplier industry momentum, standalone-conglomerate momentum, technological momentum, and connected-firm momentum.

We find that the return predictability of GPRET is more pronounced among focal firms that receive less attention and that are more costly to arbitrage. These findings are broadly consistent with the view that psychological biases, or information processing constraints, are contributing to the return predictability effect of GPRET. We also show that roughly 16% of abnormal returns of the GPRET strategy are realized in the three-day window around earnings announcement dates, suggesting that our findings are more consistent with a mispricing explanation.

Overall, our main contribution is that we document a more comprehensive geographic linkage among public firms, which utilizes information about locations of headquarters as well as economically relevant subsidiaries. Therefore, our papers highlights the economic importance of the multidimensional nature of a firm's geographic locations, which is largely ignored by the prior studies.

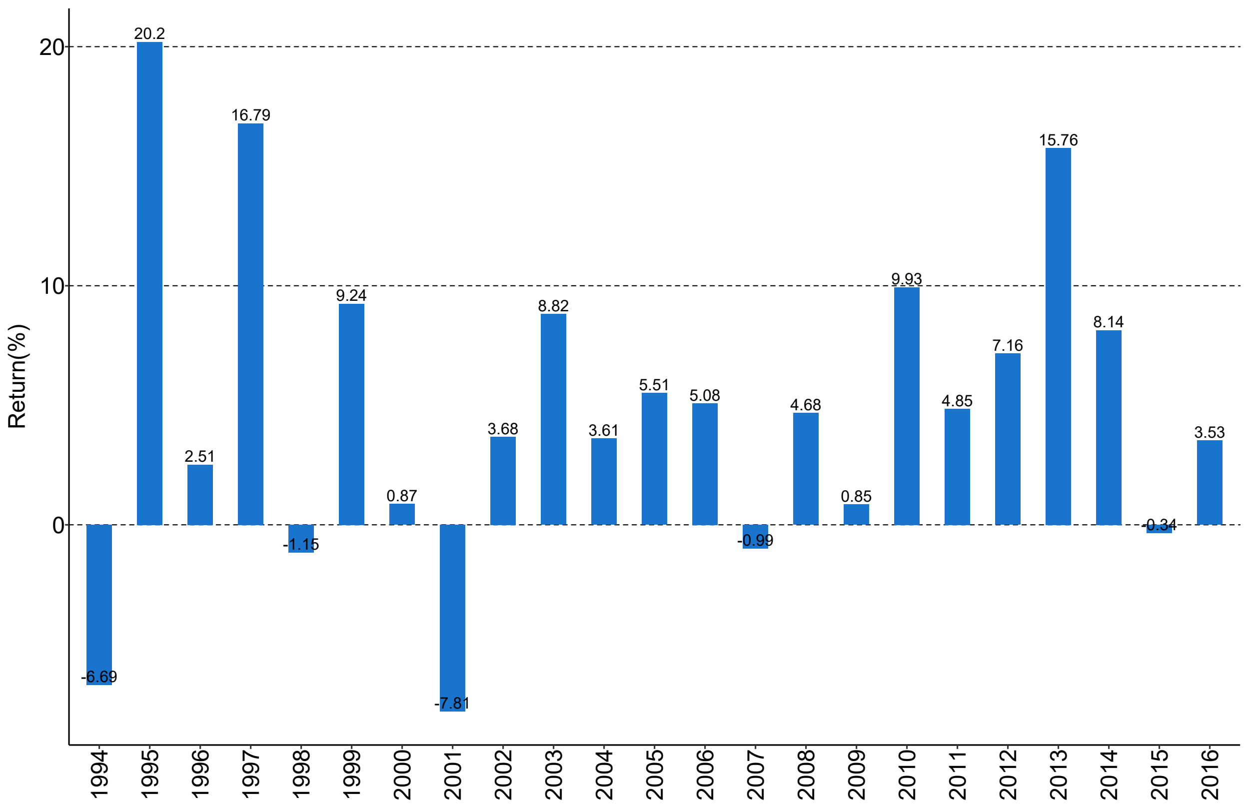


Fig. 2.1 High-minus-low GPRET portfolio returns by year

This figure depicts the annual value-weighted returns of the long-short portfolio based on GPRET from 1994 to 2016. At the end of each month, all firms are sorted into decile portfolios based on GPRET, and next month's value-weighted return of each portfolio is calculated. The portfolios are rebalanced every calendar month to maintain value weights. There are only six months in 1994 and 2016.

Table 2.1 Summary statistics

This table reports summary statistics for the key variables in the sample. The sample consists of all common stocks listed on NYSE, Nasdaq, and Amex. Non-financial and non-utility stocks with stock price greater than $5 at portfolio formation are included. All variables in the sample are winsorized at the 1% level for both tails to mitigate the effect of outliers. The mean, standard deviation (SD), minimum, median, and maximum of each variable are presented in Panel A, and their pairwise correlations are presented in Panel B. The overall sample period is from June 1994 to May 2016.

Panel A: Descriptive statistics

	Mean	Sd	Min	Median	Max
Sample statistics					
# of firms	1351	420	304	1482	1887
# of linked states	3	4	1	1	48
% of Value of CRSP	52.63	12.47	12.66	56.05	65.59
Key variables					
GPRET	0.01	0.05	-0.19	0.02	0.21
MKT_BETA	1.09	0.77	-0.99	0.96	4.82
SIZE	6.19	2.06	0.69	6.19	11.74
BM	-0.60	0.85	-3.70	-0.55	2.56
MOM	1.12	0.56	0.04	1.06	9.10
STRV	0.01	0.15	-0.71	0.01	2.06
GP	0.29	0.26	-0.49	0.24	1.25
AG	1.15	0.42	0.31	1.06	5.45
TO	0.14	0.16	0.00	0.09	1.59
IVOL	0.02	0.02	0.00	0.02	0.24
ILLIQ	1.13	4.72	0.00	0.01	36.12
RET_{t+1}	0.01	0.15	-0.71	0.01	2.06

Panel B: Pearson (Spearman) correlations below (above) the diagonal

	GPRET	MKT_BETA	SIZE	BM	MOM	STRV	GP	AG	TO	IVOL	ILLIQ	RET_{t+1}
GPRET	NA	-0.01	-0.06	0.01	0.00	0.04	-0.00	-0.01	-0.02	0.03	0.06	0.01
MKT_BETA	0.00	NA	0.05	-0.11	-0.06	-0.02	0.14	-0.02	0.37	0.27	-0.10	-0.01
SIZE	-0.07	0.03	NA	-0.37	0.24	0.10	0.04	0.17	0.42	-0.50	-0.95	0.04
BM	0.01	-0.11	-0.35	NA	-0.01	0.00	-0.31	-0.21	-0.27	0.06	0.37	0.00
MOM	0.00	-0.02	0.19	-0.01	NA	0.03	0.02	-0.01	0.05	-0.21	-0.22	0.04
STRV	0.03	-0.01	0.07	0.01	0.02	NA	0.01	0.00	0.04	-0.03	-0.07	-0.02
GP	-0.00	0.09	0.01	-0.28	0.04	0.01	NA	-0.03	0.15	0.11	-0.07	0.01
AG	-0.00	0.02	0.09	-0.13	-0.05	-0.02	-0.08	NA	0.12	-0.08	-0.17	-0.00
TO	-0.00	0.34	0.24	-0.20	0.09	0.04	0.10	0.12	NA	0.14	-0.59	0.00
IVOL	0.03	0.22	-0.50	0.06	-0.18	0.03	0.06	-0.02	0.24	NA	0.46	-0.06
ILLIQ	0.02	-0.10	-0.42	0.18	-0.13	-0.04	-0.01	-0.08	-0.18	0.38	NA	-0.04
RET_{t+1}	0.01	-0.01	0.01	0.01	0.03	-0.02	0.01	-0.02	-0.02	-0.04	-0.01	NA

Table 2.2 Univariate portfolio analysis: Geo-peers based on locations of headquarters and material subsidiaries

This table reports average monthly excess returns and alphas of portfolios sorted on the return of a portfolio of their geo-peers (GPRET). Geo-peers are defined as firms whose headquarters or material subsidiaries are located in the same state as the focal firm's headquarter or material subsidiaries. At the end of each month from 1994:06 to 2016:05, individual stocks are sorted into deciles based on geo-peer returns, where portfolio 1 (P1) contains stocks with the lowest geo-peer returns and portfolio 10 (P10) contains stocks with the highest geo-peer returns. All stocks are equally (value) weighted within a given portfolio, and the portfolios are rebalanced every calendar month to maintain equal (value) weights. The long-short portfolio is a zero-cost portfolio that buys the top decile and sells the bottom decile. Asset pricing models include Fama and French (1993) three-factor model (FF3), Fama and French (1993) three factors augmented with Carhart (1997) momentum factor (FFC), Fama and French (2015) five-factor model (FF5), Fama and French (2018) six-factor model, Hou, Xue, and Zhang (2015) q-factor model (HXZ), Stambaugh and Yuan (2017) mispricing-factor model (SY), and Daniel, Hirshleifer, and Sun (2020) behavioral-factor model (DHS). Returns and alphas are expressed in percentage. Panel A reports results of equal-weighted portfolios and Panel B reports results of value-weighted portfolios. t-statistics are shown in parentheses.

	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P10-1
Panel A: Equal-Weighted											
Excess	0.64	0.71	0.75	0.64	0.74	0.66	0.87	1.06	1.05	1.07	0.43
	(1.97)	(1.95)	(2.11)	(1.82)	(2.17)	(1.98)	(2.64)	(3.17)	(3.24)	(3.00)	(3.00)
FF3	-0.13	-0.15	-0.09	-0.22	-0.09	-0.15	0.09	0.27	0.29	0.25	0.38
	(-1.23)	(-1.31)	(-0.73)	(-1.87)	(-0.87)	(-1.51)	(0.87)	(2.32)	(2.74)	(2.20)	(2.80)
FFC	-0.05	-0.08	0.02	-0.11	0.00	-0.07	0.18	0.33	0.34	0.33	0.38
	(-0.46)	(-0.71)	(0.13)	(-0.98)	(0.00)	(-0.70)	(1.78)	(2.87)	(3.23)	(2.91)	(2.70)
FF5	-0.17	-0.20	-0.22	-0.32	-0.17	-0.23	0.03	0.17	0.24	0.27	0.45
	(-1.58)	(-1.75)	(-1.69)	(-2.73)	(-1.61)	(-2.23)	(0.28)	(1.41)	(2.13)	(2.23)	(3.08)
FF6	-0.12	-0.16	-0.14	-0.24	-0.11	-0.17	0.09	0.21	0.27	0.32	0.44
	(-1.10)	(-1.37)	(-1.15)	(-2.22)	(-1.06)	(-1.74)	(0.91)	(1.85)	(2.50)	(2.72)	(3.01)
HXZ	-0.09	-0.11	-0.07	-0.21	-0.09	-0.16	0.12	0.24	0.26	0.34	0.43
	(-0.77)	(-0.92)	(-0.51)	(-1.64)	(-0.76)	(-1.37)	(1.08)	(1.86)	(2.29)	(2.77)	(2.99)
SY	-0.04	-0.11	-0.01	-0.13	-0.03	-0.14	0.11	0.22	0.30	0.41	0.45
	(-0.33)	(-0.87)	(-0.08)	(-0.99)	(-0.21)	(-1.21)	(0.95)	(1.64)	(2.51)	(3.07)	(3.01)
DHS	0.13	0.16	0.30	0.06	0.22	0.09	0.35	0.41	0.47	0.59	0.46
	(0.87)	(0.83)	(1.52)	(0.35)	(1.34)	(0.57)	(2.07)	(2.38)	(2.76)	(3.13)	(3.04)
Panel B: Value-Weighted											
Excess	0.43	0.65	0.62	0.78	0.54	0.78	0.60	0.64	0.92	1.00	0.57
	(1.61)	(1.87)	(1.79)	(2.52)	(1.85)	(2.64)	(2.06)	(2.15)	(3.00)	(2.65)	(2.30)
FF3	-0.13	-0.11	-0.18	0.11	-0.09	0.15	0.02	0.04	0.32	0.30	0.43
	(-1.27)	(-0.72)	(-1.35)	(0.78)	(-0.69)	(1.14)	(0.14)	(0.24)	(1.95)	(1.52)	(1.90)
FFC	-0.09	-0.09	-0.11	0.12	-0.05	0.19	0.05	0.01	0.29	0.35	0.44
	(-0.89)	(-0.59)	(-0.81)	(0.84)	(-0.37)	(1.43)	(0.36)	(0.07)	(1.75)	(1.73)	(1.90)
FF5	-0.17	-0.11	-0.31	-0.04	-0.21	0.07	-0.15	-0.14	0.19	0.43	0.60
	(-1.58)	(-0.63)	(-2.22)	(-0.27)	(-1.55)	(0.53)	(-0.97)	(-0.90)	(1.10)	(2.07)	(2.53)
FF6	-0.15	-0.09	-0.25	-0.02	-0.18	0.11	-0.12	-0.15	0.18	0.45	0.60
	(-1.33)	(-0.55)	(-1.85)	(-0.16)	(-1.30)	(0.76)	(-0.76)	(-0.95)	(1.03)	(2.17)	(2.49)
HXZ	-0.09	-0.07	-0.22	0.03	-0.17	0.14	-0.03	-0.04	0.17	0.55	0.65
	(-0.86)	(-0.43)	(-1.56)	(0.18)	(-1.23)	(1.04)	(-0.19)	(-0.25)	(0.99)	(2.69)	(2.72)
SY	-0.10	-0.16	-0.10	0.09	-0.05	0.11	-0.05	-0.17	0.15	0.44	0.54
	(-0.86)	(-0.95)	(-0.67)	(0.59)	(-0.34)	(0.81)	(-0.33)	(-1.03)	(0.89)	(2.04)	(2.16)
DHS	-0.16	-0.03	0.08	0.15	-0.13	0.14	-0.03	-0.06	0.31	0.44	0.61
	(-1.45)	(-0.16)	(0.49)	(0.99)	(-0.91)	(0.97)	(-0.20)	(-0.36)	(1.79)	(2.07)	(2.46)

Table 2.3 Univariate portfolio analysis: Geo-peers based on locations of headquarters

This table reports average monthly excess returns and alphas of portfolios sorted on the return of a portfolio of their geo-peers (GPRET). Geo-peers are defined as firms whose headquarters are located in the same state as the focal firm's headquarter. At the end of each month from 1994:06 to 2016:05, individual stocks are sorted into deciles based on geo-peer returns, where portfolio 1 (P1) contains stocks with the lowest geo-peer returns and portfolio 10 (P10) contains stocks with the highest geo-peer returns. All stocks are equally (value) weighted within a given portfolio, and the portfolios are rebalanced every calendar month to maintain equal (value) weights. The long-short portfolio is a zero-cost portfolio that buys the top decile and sells the bottom decile. Asset pricing models include Fama and French (1993) three-factor model (FF3), Fama and French (1993) three factors augmented with Carhart (1997) momentum factor (FFC), Fama and French (2015) five-factor model (FF5), Fama and French (2018) six-factor model, Hou, Xue, and Zhang (2015) q-factor model (HXZ), Stambaugh and Yuan (2017) mispricing-factor model (SY), and Daniel, Hirshleifer, and Sun (2020) behavioral-factor model (DHS). Returns and alphas are expressed in percentage. Panel A reports results of equal-weighted portfolios and Panel B reports results of value-weighted portfolios. t-statistics are shown in parentheses.

	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P10-1
Panel A: Equal-Weighted											
Excess	0.60	0.73	0.74	0.71	0.81	0.83	0.86	0.78	1.00	1.04	0.44
	(1.64)	(2.02)	(2.06)	(2.10)	(2.39)	(2.52)	(2.58)	(2.32)	(2.99)	(3.01)	(2.52)
FF3	-0.24	-0.10	-0.11	-0.10	-0.02	0.03	0.06	-0.01	0.23	0.25	0.49
	(-1.92)	(-0.75)	(-0.91)	(-0.91)	(-0.22)	(0.37)	(0.60)	(-0.09)	(2.20)	(2.02)	(2.77)
FFC	-0.17	0.00	-0.02	-0.02	0.07	0.13	0.12	0.08	0.29	0.32	0.49
	(-1.35)	(0.03)	(-0.21)	(-0.19)	(0.74)	(1.42)	(1.15)	(0.78)	(2.81)	(2.63)	(2.72)
FF5	-0.30	-0.21	-0.20	-0.22	-0.10	0.01	0.02	-0.07	0.22	0.24	0.54
	(-2.26)	(-1.50)	(-1.62)	(-2.01)	(-1.02)	(0.05)	(0.22)	(-0.63)	(1.98)	(1.86)	(2.89)
FF6	-0.25	-0.13	-0.14	-0.16	-0.03	0.07	0.06	-0.01	0.26	0.29	0.54
	(-1.91)	(-0.99)	(-1.17)	(-1.54)	(-0.39)	(0.75)	(0.60)	(-0.06)	(2.41)	(2.28)	(2.85)
HXZ	-0.21	-0.06	-0.08	-0.10	-0.00	0.07	0.03	0.01	0.24	0.32	0.54
	(-1.55)	(-0.38)	(-0.63)	(-0.80)	(-0.04)	(0.73)	(0.31)	(0.12)	(2.21)	(2.47)	(2.89)
SY	-0.11	0.05	-0.06	-0.08	0.02	0.08	0.07	0.03	0.24	0.28	0.38
	(-0.77)	(0.34)	(-0.41)	(-0.63)	(0.15)	(0.75)	(0.60)	(0.22)	(2.00)	(1.93)	(2.01)
DHS	0.10	0.24	0.16	0.17	0.24	0.29	0.31	0.24	0.49	0.53	0.43
	(0.51)	(1.29)	(0.86)	(1.01)	(1.56)	(1.81)	(1.86)	(1.41)	(2.91)	(2.71)	(2.30)
Panel B: Value-Weighted											
Excess	0.65	0.83	0.81	0.54	0.83	0.69	0.65	0.60	0.74	1.00	0.35
	(1.83)	(2.61)	(2.72)	(1.93)	(3.09)	(2.36)	(2.29)	(2.21)	(2.29)	(2.80)	(1.24)
FF3	-0.08	0.18	0.18	-0.02	0.27	0.08	0.05	0.06	0.17	0.34	0.42
	(-0.47)	(1.07)	(1.28)	(-0.16)	(2.13)	(0.60)	(0.36)	(0.45)	(0.87)	(1.55)	(1.47)
FFC	-0.09	0.21	0.21	0.04	0.27	0.05	0.07	0.09	0.15	0.31	0.41
	(-0.52)	(1.27)	(1.46)	(0.31)	(2.08)	(0.37)	(0.51)	(0.59)	(0.76)	(1.40)	(1.39)
FF5	-0.16	0.09	0.10	-0.09	0.10	0.02	0.01	-0.04	0.06	0.31	0.47
	(-0.86)	(0.51)	(0.67)	(-0.60)	(0.75)	(0.14)	(0.10)	(-0.28)	(0.28)	(1.35)	(1.56)
FF6	-0.17	0.11	0.12	-0.04	0.11	0.00	0.03	-0.02	0.05	0.29	0.46
	(-0.87)	(0.67)	(0.83)	(-0.26)	(0.81)	(0.03)	(0.21)	(-0.14)	(0.24)	(1.27)	(1.51)
HXZ	-0.11	0.17	0.17	-0.01	0.16	0.06	0.02	0.06	0.07	0.36	0.47
	(-0.58)	(0.96)	(1.12)	(-0.07)	(1.22)	(0.44)	(0.11)	(0.43)	(0.36)	(1.58)	(1.56)
SY	-0.06	0.27	0.16	0.03	0.14	-0.03	0.07	0.02	-0.08	0.12	0.18
	(-0.34)	(1.52)	(1.06)	(0.19)	(0.99)	(-0.20)	(0.47)	(0.10)	(-0.36)	(0.51)	(0.59)
DHS	-0.09	0.35	0.08	0.07	0.16	0.08	-0.11	0.06	0.07	0.25	0.34
	(-0.48)	(2.01)	(0.53)	(0.46)	(1.15)	(0.53)	(-0.81)	(0.40)	(0.32)	(1.09)	(1.13)

Table 2.4 Univariate portfolio analysis: Geo-peers based on locations of material subsidiaries

This table reports average monthly excess returns and alphas of portfolios sorted on the return of a portfolio of their geo-peers (GPRET). Geo-peers are defined as firms whose material subsidiaries are located in the same state as the focal firm's material subsidiaries. At the end of each month from 1994:06 to 2016:05, individual stocks are sorted into deciles based on geo-peer returns, where portfolio 1 (P1) contains stocks with the lowest geo-peer returns and portfolio 10 (P10) contains stocks with the highest geo-peer returns. All stocks are equally (value) weighted within a given portfolio, and the portfolios are rebalanced every calendar month to maintain equal (value) weights. The long-short portfolio is a zero-cost portfolio that buys the top decile and sells the bottom decile. Asset pricing models include Fama and French (1993) three-factor model (FF3), Fama and French (1993) three factors augmented with Carhart (1997) momentum factor (FFC), Fama and French (2015) five-factor model (FF5), Fama and French (2018) six-factor model, Hou, Xue, and Zhang (2015) q-factor model (HXZ), Stambaugh and Yuan (2017) mispricing-factor model (SY), and Daniel, Hirshleifer, and Sun (2020) behavioral-factor model (DHS). Returns and alphas are expressed in percentage. Panel A reports results of equal-weighted portfolios and Panel B reports results of value-weighted portfolios. t-statistics are shown in parentheses.

	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P10-1
Panel A: Equal-Weighted											
Excess	0.65	0.70	0.73	0.66	0.74	0.66	0.87	1.06	1.05	1.07	0.42
	(1.79)	(1.98)	(2.15)	(1.95)	(2.16)	(1.98)	(2.64)	(3.17)	(3.24)	(3.00)	(2.84)
FF3	-0.19	-0.13	-0.08	-0.18	-0.09	-0.15	0.09	0.27	0.29	0.25	0.44
	(-1.62)	(-1.05)	(-0.75)	(-1.61)	(-0.91)	(-1.51)	(0.87)	(2.32)	(2.74)	(2.20)	(2.96)
FFC	-0.11	-0.04	0.01	-0.07	-0.00	-0.07	0.18	0.33	0.34	0.33	0.44
	(-0.97)	(-0.33)	(0.06)	(-0.68)	(-0.02)	(-0.70)	(1.78)	(2.87)	(3.23)	(2.91)	(2.90)
FF5	-0.19	-0.26	-0.18	-0.29	-0.17	-0.23	0.03	0.17	0.24	0.27	0.47
	(-1.57)	(-2.09)	(-1.50)	(-2.54)	(-1.63)	(-2.23)	(0.28)	(1.41)	(2.13)	(2.23)	(2.93)
FF6	-0.14	-0.19	-0.11	-0.21	-0.11	-0.17	0.09	0.21	0.27	0.32	0.46
	(-1.17)	(-1.63)	(-0.98)	(-2.01)	(-1.08)	(-1.74)	(0.91)	(1.85)	(2.50)	(2.72)	(2.90)
HXZ	-0.13	-0.10	-0.06	-0.20	-0.09	-0.16	0.12	0.24	0.26	0.34	0.47
	(-1.01)	(-0.74)	(-0.46)	(-1.57)	(-0.79)	(-1.37)	(1.08)	(1.86)	(2.29)	(2.77)	(2.99)
SY	-0.09	-0.08	-0.02	-0.10	-0.03	-0.14	0.11	0.22	0.30	0.41	0.50
	(-0.66)	(-0.61)	(-0.14)	(-0.79)	(-0.25)	(-1.21)	(0.95)	(1.64)	(2.51)	(3.07)	(3.07)
DHS	0.15	0.19	0.24	0.08	0.21	0.09	0.35	0.41	0.47	0.59	0.44
	(0.76)	(1.00)	(1.42)	(0.51)	(1.29)	(0.57)	(2.07)	(2.38)	(2.76)	(3.13)	(2.77)
Panel B: Value-Weighted											
Excess	0.46	0.67	0.67	0.53	0.58	0.78	0.60	0.64	0.92	1.00	0.54
	(1.37)	(2.34)	(2.30)	(1.89)	(1.97)	(2.64)	(2.06)	(2.15)	(3.00)	(2.65)	(2.04)
FF3	-0.21	0.08	0.06	-0.08	-0.05	0.15	0.02	0.04	0.32	0.30	0.51
	(-1.15)	(0.54)	(0.40)	(-0.62)	(-0.37)	(1.14)	(0.14)	(0.24)	(1.95)	(1.52)	(1.94)
FFC	-0.13	0.06	0.09	-0.07	-0.02	0.19	0.05	0.01	0.29	0.35	0.48
	(-0.72)	(0.42)	(0.61)	(-0.58)	(-0.15)	(1.43)	(0.36)	(0.07)	(1.75)	(1.73)	(1.79)
FF5	-0.25	-0.08	-0.05	-0.16	-0.18	0.07	-0.15	-0.14	0.19	0.43	0.68
	(-1.32)	(-0.49)	(-0.32)	(-1.20)	(-1.27)	(0.53)	(-0.97)	(-0.90)	(1.10)	(2.07)	(2.47)
FF6	-0.20	-0.08	-0.02	-0.15	-0.15	0.11	-0.12	-0.15	0.18	0.45	0.65
	(-1.03)	(-0.51)	(-0.14)	(-1.13)	(-1.08)	(0.76)	(-0.76)	(-0.95)	(1.03)	(2.17)	(2.34)
HXZ	-0.08	-0.06	0.01	-0.12	-0.13	0.14	-0.03	-0.04	0.17	0.55	0.64
	(-0.44)	(-0.37)	(0.07)	(-0.95)	(-0.88)	(1.04)	(-0.19)	(-0.25)	(0.99)	(2.69)	(2.29)
SY	-0.04	-0.04	0.06	-0.13	-0.01	0.11	-0.05	-0.17	0.15	0.44	0.48
	(-0.23)	(-0.24)	(0.37)	(-0.96)	(-0.06)	(0.81)	(-0.33)	(-1.03)	(0.89)	(2.04)	(1.68)
DHS	-0.12	-0.01	0.05	-0.12	-0.05	0.14	-0.03	-0.06	0.31	0.44	0.57
	(-0.63)	(-0.06)	(0.35)	(-0.90)	(-0.36)	(0.97)	(-0.20)	(-0.36)	(1.79)	(2.07)	(2.04)

Table 2.5 Factor loadings

This table reports the risk factor loadings of the long-short GPRET portfolio under various asset pricing factor models. At the end of each month from 1994:06 to 2016:05, individual stocks are sorted into deciles based on geo-peer returns, where portfolio 1 (P1) contains stocks with the lowest geo-peer returns and portfolio 10 (P10) contains stocks with the highest geo-peer returns. All stocks are equally (value) weighted within a given portfolio, and the portfolios are rebalanced every calendar month to maintain equal (value) weights. The long-short portfolio is a zero-cost portfolio that buys the top decile and sells the bottom decile. Asset pricing models include Fama and French (1993) three-factor model (FF3), Fama and French (1993) three factors augmented with Carhart (1997) momentum factor (FFC), Fama and French (2015) five-factor model (FF5), Fama and French (2018) six-factor model, Hou, Xue, and Zhang (2015) q-factor model (HXZ), Stambaugh and Yuan (2017) mispricing-factor model (SY), and Daniel, Hirshleifer, and Sun (2020) behavioral-factor model (DHS). Returns and alphas are expressed in percentage. Panel A reports results of equal-weighted portfolios and Panel B reports result of value-weighted portfolios. t-statistics are shown in parentheses.

Panel A: Equal-Weighted

Model	MKT	SMB	HML	MOM	RMW	CMA	IA	ROE	MGMT	PERF	PEAD	FIN	R^2
FF3	0.02	0.25	-0.05										0.13
	(1.00)	(4.78)	(-1.71)										
FFC	0.03	0.23	-0.04	0.01									0.13
	(1.04)	(4.35)	(-0.46)	(0.38)									
FF5	-0.01	0.22	0.01		-0.09	-0.05							0.13
	(-0.08)	(3.33)	(0.31)		(-1.44)	(-0.81)							
FF6	-0.00	0.22	0.02	0.02	-0.10	-0.06							0.14
	(-0.08)	(3.38)	(0.35)	(1.03)	(-1.21)	(-0.59)							
HXZ	-0.00	0.22					-0.05	-0.09					0.13
	(-0.09)	(3.43)					(-0.68)	(-1.27)					
SY	-0.01	0.24							-0.09	-0.04			0.12
	(-0.13)	(3.24)							(-2.31)	(-1.32)			
DHS	0.01										0.08	-0.11	0.06
	(0.27)										(1.19)	(-2.14)	

Panel B: Value-Weighted

Model	MKT	SMB	HML	MOM	RMW	CMA	IA	ROE	MGMT	PERF	PEAD	FIN	R^2
FF3	0.19	0.32	-0.23										0.19
	(1.31)	(4.36)	(-2.10)										
FFC	0.19	0.34	-0.23	-0.01									0.19
	(1.24)	(4.24)	(-1.89)	(-0.28)									
FF5	0.11	0.27	-0.07		-0.20	-0.24							0.21
	(1.43)	(3.59)	(-0.63)		(-1.57)	(-1.54)							
FF6	0.11	0.28	-0.06	0.01	-0.21	-0.24							0.21
	(1.48)	(3.13)	(-0.46)	(0.29)	(-1.73)	(-1.68)							
HXZ	0.09	0.25					-0.36	-0.24					0.20
	(1.22)	(3.23)					(-1.99)	(-2.53)					
SY	0.14	0.31							-0.27	-0.00			0.18
	(1.71)	(3.31)							(-2.77)	(-0.17)			
DHS	0.12										0.12	-0.26	0.16
	(1.59)										(1.11)	(-2.43)	

Table 2.6 Bivariate sorts on size and GRET

This table reports the results on how the GPRET effect varies with firm size. At the end of each month from 1994:06 to 2016:05, we sort all the stocks into quintiles based on the market capitalization. Within each size quintile, we further sort stocks into quintiles based on the GPRET. We report the Fama and French (2018) six-factor alphas of the 25 portfolios. We also report, for each size quintile, the long-short GPRET portfolio alpha. Panel A reports results of equal-weighted portfolios and Panel B reports results of value-weighted portfolios. t-statistics are shown in parentheses.

Panel A: Equal-Weighted					
Size Quintile	Small Firms	2	3	4	Large Firms
Short	-0.38	-0.13	-0.22	-0.26	-0.15
	(-1.94)	(-0.80)	(-1.34)	(-1.52)	(-1.32)
P2	0.04	-0.33	-0.36	-0.18	0.01
	(0.19)	(-2.08)	(-2.52)	(-1.30)	(0.06)
P3	-0.08	-0.26	-0.13	-0.08	0.03
	(-0.38)	(-1.75)	(-0.93)	(-0.60)	(0.31)
P4	0.17	0.39	0.16	-0.14	-0.16
	(0.82)	(2.28)	(1.18)	(-1.14)	(-1.26)
Long	0.22	0.33	0.32	0.18	0.33
	(1.20)	(2.11)	(2.18)	(1.15)	(1.82)
L/S	0.60	0.46	0.53	0.45	0.48
	(2.58)	(2.22)	(2.44)	(2.17)	(2.10)
Panel B: Value-Weighted					
Size Quintile	Small Firms	2	3	4	Large Firms
Short	-0.21	-0.19	-0.30	-0.24	-0.06
	(-1.18)	(-1.25)	(-2.02)	(-1.48)	(-0.50)
P2	0.05	-0.22	-0.29	-0.21	-0.10
	(0.25)	(-1.36)	(-1.95)	(-1.58)	(-0.86)
P3	0.09	-0.30	-0.15	-0.10	-0.04
	(0.51)	(-2.07)	(-1.09)	(-0.68)	(-0.33)
P4	0.20	0.44	0.16	-0.14	-0.02
	(1.09)	(2.72)	(1.13)	(-1.12)	(-0.15)
Long	0.50	0.34	0.35	0.16	0.27
	(2.48)	(2.07)	(2.33)	(1.09)	(2.11)
L/S	0.71	0.52	0.65	0.39	0.32
	(3.17)	(2.70)	(3.11)	(2.19)	(1.97)

Table 2.7 Fama-MacBeth regressions

This table reports the results of Fama and MacBeth (1973) return forecasting regressions. The dependent variable is either the focal firm's monthly excess return RET (the first two columns), the firm's excess return over its value-weighted industry return based on Fama and French (1997) 48 industry classifications RET-INDRET (Column 3), or firm's excess return over its value-weighted Daniel et al. (1997) characteristics benchmark return. The independent variables include geo-peer return (GPRET), market beta (MKT_BETA), natural logarithm of firm's market capitalization (Log(SIZE)), natural logarithm of book-to-market ratio (Log(BM)), price momentum (MOM), short-term reversal (STREV), gross profitability (GP), asset growth (AG), firm's turnover ratio (TO), idiosyncratic volatility (IVOL), and Amihud (2002)'s illiquidity measure (ILLIQ). All independent variables except industry dummies are normalized to zero mean and standard deviation of one after winsorization at the 1% and 99% levels. Newey and West (1987) adjusted t-statistics are shown below the coefficient estimates in parentheses. Coefficients marked with *, **, and *** are significant at the 10%, the 5%, and the 1% level, respectively.

Dep.Var.*100	$RET_{i,t+1}$	$RET_{i,t+1}$	$RET_{i,t+1}$ -$INDRET_{i,t+1}$	DGTW-adj. $RET_{i,t+1}$
GPRET	0.09***	0.08***	0.09***	0.08***
	(3.72)	(3.90)	(4.20)	(3.34)
MKT_BETA	-0.04	0.00	-0.00	-0.02
	(-0.37)	(0.02)	(-0.04)	(-0.27)
SIZE	-0.07	-0.06	-0.03	-0.03
	(-0.97)	(-0.85)	(-0.56)	(-0.73)
BM	0.10	0.14**	0.11**	-0.01
	(1.50)	(2.55)	(2.30)	(-0.11)
MOM	0.13	0.09	0.11	0.07
	(1.40)	(0.97)	(1.36)	(0.94)
STRV	-0.33***	-0.37***	-0.35***	-0.31***
	(-5.09)	(-5.89)	(-5.96)	(-5.09)
GP	0.09*	0.13***	0.06*	0.07*
	(1.83)	(3.75)	(1.77)	(1.70)
AG	-0.11***	-0.11***	-0.10***	-0.10***
	(-3.25)	(-3.59)	(-3.39)	(-3.26)
TO	0.00	0.02	-0.01	0.01
	(0.05)	(0.32)	(-0.23)	(0.14)
IVOL	-0.12**	-0.11**	-0.10*	-0.11**
	(-2.21)	(-2.16)	(-1.90)	(-2.28)
ILLIQ	-0.11***	-0.10**	-0.09**	-0.11***
	(-2.63)	(-2.32)	(-2.20)	(-3.12)
Intercept	0.83**	0.52	-0.10	-0.21***
	(2.41)	(1.22)	(-0.81)	(-3.24)
Industry FEs	No	Yes	No	No
N	308215	308214	308214	307428
Adj. R^2	0.072	0.116	0.045	0.041

Table 2.8 Fama-Macbeth regression: predicting long term future returns

This table reports the results of Fama and MacBeth (1973) return forecasting regression. The dependent variable is the future excess return that covers each month over the next 12 months. The independent variables include geo-peer return (GPRET), market beta (MKT_BETA), natural logarithm of firm's market capitalization (Log(SIZE)), natural logarithm of book-to-market ratio (Log(BM)), price momentum (MOM), short-term reversal (STREV), gross profitability (GP), asset growth (AG), firm's turnover ratio (TO), idiosyncratic volatility (IVOL), and Amihud (2002)'s illiquidity measure (ILLIQ). All independent variables are normalized to zero mean and standard deviation of one after winsorization at the 1% and 99% levels. Newey and West (1987) adjusted t-statistics are shown below the coefficient estimates in parentheses. Coefficients marked with *, **, and *** are significant at the 10%, the 5%, and the 1% level, respectively.

Month	t+1	t+2	t+3	t+4	t+5	t+6	t+7	t+8	t+9	t+10	t+11	t+12
GPRET	0.09***	0.08***	0.07***	0.07**	0.04	0.01	0.05**	0.04	0.03	-0.00	0.04	-0.00
	(3.72)	(3.41)	(3.42)	(2.40)	(1.41)	(0.20)	(2.04)	(1.41)	(1.13)	(-0.13)	(1.40)	(-0.07)
MKT_BETA	-0.04	-0.01	-0.04	-0.03	-0.01	-0.02	0.02	0.02	0.02	0.04	0.06	0.05
	(-0.37)	(-0.32)	(-0.36)	(-0.37)	(-0.03)	(-0.21)	(0.17)	(0.18)	(0.25)	(0.48)	(0.70)	(0.54)
SIZE	-0.07	-0.07	-0.04	-0.01	-0.01	-0.01	-0.02	-0.02	0.04	0.02	-0.00	0.00
	(-0.97)	(-0.68)	(-0.42)	(-0.11)	(-0.14)	(-0.09)	(-0.31)	(-0.26)	(0.61)	(0.29)	(-0.05)	(0.02)
BM	0.10	0.07	0.07	0.06	0.06	0.06	0.06	0.08	0.08	0.06	0.04	0.04
	(1.50)	(1.36)	(1.11)	(1.03)	(1.02)	(0.99)	(1.00)	(1.23)	(1.25)	(0.95)	(0.63)	(0.66)
MOM	0.13	0.14*	0.07	0.06	0.04	0.00	-0.04	-0.02	-0.05	-0.01	-0.09	-0.05
	(1.39)	(1.70)	(1.33)	(1.09)	(1.03)	(0.04)	(-0.52)	(-0.34)	(-0.78)	(-0.25)	(-1.38)	(-0.77)
STRV	-0.33***	-0.08	0.09*	-0.02	0.07	0.11**	0.05	0.02	0.04	0.03	0.12**	0.05
	(-5.08)	(-1.44)	(1.76)	(-0.65)	(1.11)	(2.31)	(1.13)	(0.34)	(0.85)	(0.62)	(2.45)	(0.91)
GP	0.09*	0.09*	0.08	0.07*	0.10**	0.10**	0.10**	0.09*	0.09*	0.09*	0.07	0.07
	(1.83)	(1.88)	(1.77)	(1.83)	(1.99)	(2.12)	(2.06)	(1.84)	(1.95)	(1.87)	(1.52)	(1.51)
AG	-0.11***	-0.08**	-0.07*	-0.08**	-0.09***	-0.09***	-0.08**	-0.08***	-0.08***	-0.08**	-0.09***	-0.08***
	(-3.26)	(-2.34)	(-1.95)	(-2.49)	(-2.63)	(-2.94)	(-2.55)	(-2.92)	(-2.73)	(-2.35)	(-2.99)	(-2.79)
TO	0.00	-0.03	-0.03	-0.03	-0.00	-0.03	-0.01	-0.01	-0.01	-0.03	0.01	-0.03
	(0.05)	(-0.80)	(-0.56)	(-0.34)	(-0.23)	(-0.49)	(-0.21)	(-0.13)	(-0.09)	(-0.42)	(0.13)	(-0.40)
IVOL	-0.12**	-0.17***	-0.09	-0.10	-0.16***	-0.03	-0.09*	-0.11**	-0.05	-0.09*	-0.10**	-0.06
	(-2.21)	(-2.96)	(-1.39)	(-1.46)	(-3.31)	(-0.61)	(-1.84)	(-2.34)	(-1.09)	(-1.83)	(-2.05)	(-1.14)
ILLIQ	-0.11***	-0.04	-0.04	-0.06*	-0.03	-0.06	-0.04	-0.03	0.00	-0.02	-0.02	-0.04
	(-2.63)	(-0.88)	(-0.79)	(-1.69)	(-0.78)	(-1.53)	(-1.17)	(-0.77)	(0.05)	(-0.80)	(-0.47)	(-1.18)
Intercept	0.83**	0.84**	0.84**	0.83**	0.82**	0.90**	0.90**	0.89**	0.87**	0.87**	0.89**	0.87**
	(2.41)	(2.45)	(2.42)	(2.40)	(2.36)	(2.57)	(2.57)	(2.49)	(2.42)	(2.41)	(2.44)	(2.41)
N	308215	306682	305179	303692	302227	300796	299353	297930	296510	295081	293657	292231
Adj. R^2	0.072	0.069	0.067	0.065	0.064	0.064	0.062	0.062	0.060	0.059	0.059	0.059

Table 2.9 Controlling for other economic links

This table reports the results of Fama and MacBeth (1973) return forecasting regressions controlling for other economic links documented by prior literature. The dependent variable is the focal firm's monthly excess return. GPRET_ALL is the weighted average return of firms whose headquarters or material subsidiaries are located in the same state as the focal firm's headquarter or material subsidiaries. GPRET_HQ is the weighted average return of firms whose headquarters are located in the same state as the focal firm's headquarter. INDRET is the value-weighted average return of all other stocks in the same Fama and French (1997) 48 industry. CF RET is the weighted average return of stocks that are connected through shared analyst coverage following Ali and Hirshleifer (2020). Customer (Supplier) industry RET is the weighted average return of all the industries that buy from (supply to) that stock's industry following Menzly and Ozbas (2010). The portfolio weights are the flow of goods to and from industries extracted from Bureau of Economic Analysis (BEA) Input-Output data (at the summary industry level). Single-segment RET is the weighted average return of single-segment firms operating in the same segments as the conglomerate firm following Cohen and Lou (2012). Tech. linked RET is the weighted average return of technologically similar firms following Lee et al. (2019). The portfolio weights are the technological closeness of firm pairs. The sample period of Tech. linked RET is from July 1994 to June 2012 due to the availability of Kogan et al. (2017) patent data. Board linked RET is the weighted averaged return of firms that share common board members wit the focal firm following Burt, Hrdlicka, and Harford (2020). The sample period of Board linked RET is from 2000 due to the availability of BoardEx data. Inst. linked RET is the weighted average return of firms connected through common institutional ownership following Gao, Moulton, and Ng (2017). Control variables are the same as the ones in Table 2.7. All independent variables are normalized to zero mean and standard deviation of one after winsorization at the 1% and 99% levels. Newey and West (1987) adjusted t-statistics are shown below the coefficient estimates in parentheses. Coefficients marked with *, **, and *** are significant at the 10%, the 5%, and the 1% level, respectively.

Dep.Var.*100	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
GPRET_ALL	0.10***	0.10***	0.10***	0.09***	0.09***	0.09**	0.08**	0.14***	0.08**
	(3.89)	(3.65)	(3.41)	(3.02)	(3.07)	(2.20)	(2.52)	(3.02)	(2.58)
GPRET_HQ	0.05*	0.04*	0.02	0.00	0.01	0.02	-0.01	-0.01	0.02
	(1.88)	(1.71)	(0.66)	(0.11)	(0.19)	(0.51)	(-0.25)	(-0.14)	(0.78)
INDRET		0.12**							
		(2.51)							
CF RET			0.35***						
			(4.10)						
Cus. ind. RET				0.05					
				(1.49)					
Sup. ind. RET					0.04				
					(1.07)				
Single-seg. RET						0.09			
						(1.60)			
Tech. linked RET							0.12**		
							(2.01)		
Board linked RET								0.07	
								(1.39)	
Inst. Linked RET									0.06
									(1.13)
Intercept	0.83**	0.83**	0.82**	0.81**	0.81**	0.90***	0.77**	1.08	0.83**
	(2.42)	(2.42)	(2.37)	(2.50)	(2.49)	(2.60)	(1.99)	(0.95)	(2.42)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	288834	288832	265455	161949	161949	99019	163365	156131	261922
Adj. R^2	0.073	0.076	0.084	0.089	0.090	0.100	0.091	0.099	0.074

Table 2.10 Limited attention and limits to arbitrage

This table reports results under various levels of attention and limits to arbitrage to evaluate the sensitivity of geographic momentum to proxies of limited attention and arbitrage costs. We first sort stocks into terciles based on the proxies of attention and limits to arbitrage variable X, including residual media coverage, residual analyst coverage, idiosyncratic volatility, and Amihud's illiquidity. We then independently sort stocks into quintiles based on the GPRET. We report excess returns and alphas of the lowest (Short leg) and the highest (Long leg) GPRET portfolios in the lowest and the highest X groups, respectively. The "Diff" reports the excess returns and alphas of the lowest and the highest GPRET portfolios as well as the long-short GPRET portfolios. All portfolios are value-weighted and rebalanced monthly. Asset pricing models include: Fama and French (2015) five-factor model and Fama and French (2018) six-factor model. t-statistics are shown in parentheses.

	Short	Long	L/S	FF5	FF6
Panel A: Subsamples split by proxies of investor attention					
High Residual Media Coverage	0.64	0.65	0.01	0.10	0.12
	(1.84)	(1.52)	(0.03)	(0.28)	(0.32)
Low Residual Media Coverage	0.55	1.23	0.68	0.76	0.73
	(1.45)	(2.88)	(2.42)	(2.02)	(2.35)
Diff	0.09	-0.57	-0.67	-0.65	-0.61
	(0.28)	(-1.66)	(-1.54)	(-1.59)	(-1.61)
High Residual Analyst Coverage	0.91	0.61	-0.30	-0.30	-0.25
	(2.44)	(1.41)	(-0.86)	(-0.79)	(-0.65)
Low Residual Analyst Coverage	0.36	0.89	0.53	0.63	0.66
	(0.95)	(2.19)	(2.15)	(1.95)	(2.05)
Diff	0.56	-0.27	-0.83	-0.94	-0.92
	(1.66)	(-0.67)	(-1.88)	(-1.93)	(-1.88)
Panel B: Subsamples split by proxies of limits to arbitrage					
High IVOL	-0.09	0.73	0.81	0.97	0.97
	(-0.14)	(1.14)	(2.08)	(1.77)	(1.76)
Low IVOL	1.04	0.74	-0.30	-0.29	-0.25
	(3.53)	(2.18)	(-1.04)	(-0.94)	(-0.81)
Diff	-1.13	-0.01	1.11	1.25	1.22
	(-2.03)	(-0.03)	(1.99)	(1.98)	(1.91)
High ILLIQ	0.07	0.79	0.72	0.92	0.95
	(0.19)	(2.15)	(2.67)	(3.24)	(3.31)
Low ILLIQ	0.73	0.54	-0.20	-0.13	-0.11
	(2.14)	(1.31)	(-0.65)	(-0.42)	(-0.33)
Diff	-0.66	0.26	0.92	1.06	1.06
	(-1.92)	(0.68)	(2.31)	(2.52)	(2.49)

Table 2.11 Earnings announcement returns prediction

This table reports the results of quarterly Fama and MacBeth (1973) return forecasting regressions. The dependent variable is cumulative abnormal returns (CAR) around the earnings announcement date in one-day or three-day window. The cumulative abnormal return is computed as either the sum of daily stock returns minus returns on the CRSP value-weighted portfolio (CRSPVW-adj. CAR), the sum of daily stock returns minus returns of the characteristics-matched portfolio following Daniel et al. (1997) (DGTW-adj. CAR), or the buy-and-hold stock return adjusted by the buy-and-hold return of a size and book-to-market(B/M) matched portfolio following Hirshleifer, Lim, and Teoh (2009) (SBM-adj. CAR). The independent variables include geo-peer return (GPRET), market beta (MKT_BETA), natural logarithm of firm's market capitalization (Log(SIZE)), natural logarithm of book-to-market ratio (Log(BM)), price momentum (MOM), short-term reversal (STREV), gross profitability (GP), asset growth (AG), firm's turnover ratio (TO), idiosyncratic volatility (IVOL), and Amihud (2002)'s illiquidity measure (ILLIQ). All independent variables are normalized to zero mean and standard deviation of one after winsorization at the 1% and 99% levels. Newey and West (1987) adjusted t-statistics are shown below the coefficient estimates in parentheses. Coefficients marked with *, **, and *** are significant at the 10%, the 5%, and the 1% level, respectively.

	CRSPVW-adj. CAR		DGTW-adj. CAR		SBM-adj. CAR	
	CAR_t	$CAR_{t-1,t+1}$	CAR_t	$CAR_{t-1,t+1}$	CAR_t	$CAR_{t-1,t+1}$
GPRET	0.04**	0.08***	0.04**	0.07***	0.04**	0.07***
	(2.17)	(2.83)	(2.25)	(2.79)	(2.22)	(2.77)
MKT_BETA	0.01	-0.02	0.01	-0.02	0.01	-0.03
	(0.65)	(-0.46)	(0.52)	(-0.74)	(0.65)	(-0.74)
SIZE	-0.03	0.05	-0.04**	0.01	-0.04**	0.02
	(-1.55)	(1.42)	(-2.52)	(0.27)	(-2.21)	(0.47)
BM	0.05*	0.11**	0.04*	0.08*	0.05*	0.09*
	(1.91)	(2.07)	(1.70)	(1.74)	(1.84)	(1.96)
MOM	-0.01	0.05	0.00	0.07*	-0.01	0.05
	(-0.41)	(1.09)	(0.10)	(1.93)	(-0.39)	(1.10)
STRV	-0.03	-0.08**	-0.02	-0.07**	-0.02	-0.07**
	(-1.58)	(-2.16)	(-1.28)	(-2.08)	(-1.37)	(-2.00)
GP	0.07***	0.15***	0.06***	0.14***	0.06***	0.14***
	(2.98)	(3.65)	(2.98)	(3.85)	(2.86)	(3.57)
AG	-0.03**	-0.02	-0.03*	-0.02	-0.03*	-0.02
	(-2.11)	(-0.81)	(-1.84)	(-0.52)	(-1.96)	(-0.54)
TO	0.01	-0.02	0.00	-0.02	0.01	-0.02
	(0.44)	(-0.47)	(0.04)	(-0.54)	(0.28)	(-0.41)
IVOL	-0.07***	-0.10***	-0.07***	-0.11***	-0.07***	-0.13***
	(-4.51)	(-2.77)	(-4.79)	(-3.16)	(-4.79)	(-3.50)
ILLIQ	0.04*	0.06**	0.04*	0.06**	0.03*	0.05*
	(1.70)	(2.08)	(1.72)	(2.30)	(1.69)	(1.66)
Intercept	0.16***	0.29***	0.15***	0.28***	0.17***	0.28***
	(7.87)	(6.21)	(8.49)	(7.60)	(9.21)	(7.78)
N	106572	106572	106422	106422	106422	106422
Adj. R^2	0.006	0.01	0.005	0.007	0.005	0.008

Chapter 3

Another Presidential Puzzle? Presidential Economic Approval Rating and the Cross-Section of Stock Returns*

We construct a monthly Presidential Economic Approval Rating (PEAR) index from 1981 to 2019, by averaging ratings on president's handling of the economy across various national polls. In the cross-section, stocks with high betas to changes in the PEAR index significantly under-perform those with low betas by 0.90% per month in the future, on a risk adjusted basis. The low-PEAR-beta premium persists up to one year, and is present in various sub-samples (based on industries, presidential cycles, transitions, and tenures) and even in other G7 countries. It is also robust to different risk adjustment models and controls for other related return predictors. Since the PEAR index is negatively correlated with measures of aggregate risk aversion, a simple risk model would predict the low PEAR-beta stocks to earn lower (not higher) expected returns. Contrary to the sentiment-induced overpricing, the premium does not come primarily from the short leg following high sentiment periods. Instead, the premium could be driven by a novel sentiment towards presidential alignment.

*This is a joint work with Zhi Da, Dashan Huang, and Liyao Wang

3.1 Introduction

The well-known presidential puzzle refers to the striking empirical fact that stock market returns are much higher under Democratic presidencies than Republican ones. Since first noted by Huang (1985) and Hensel and Ziemba (1995) and carefully documented by Santa-Clara and Valkanov (2003), the pattern remains robust. It is only recently that Pastor and Veronesi (2020) provide an ingenious solution to this puzzle. Their model of political cycles predicts that when risk aversion and therefore equity risk premium are high, agents elect Democrats, explaining the subsequent higher stock market returns during Democratic presidencies.

In this paper, we document a different presidential puzzle in the cross-section of individual stocks. We start by constructing a monthly Presidential Economic Approval Rating (PEAR) index from 1981 to 2019, by averaging approval ratings on president's handling of the economy across various national polls. The monthly index is plotted in Figure 3.1, together with the general Gallup presidential job approval rating. The two ratings are clearly positively correlated (with a correlation of 64%), yet they also diverge from time to time. Notable examples include the Gulf war, the September 11 terrorist attack, and president Trump's initial tenure. We obtain much stronger results using PEAR instead of the job approval rating, consistent with the phrase "the economy, stupid," popularized during Bill Clinton's successful 1992 presidential campaign. Indeed, PEAR seems to correlate with the political cycles well in that a low (high) PEAR predicts a Democratic (Republican) president in the future, suggesting that PEAR is inversely related to the aggregate risk aversion modeled in Pastor and Veronesi (2020). We confirm such an inverse relation using four different measures of aggregate risk aversion.

Surprisingly, in the cross-section, stocks with high betas to changes in PEAR significantly under-perform those with low betas by 0.9% per month in the future, on a risk adjusted basis. A simple extension of a risk-based model of the aggregate stock market, say Pastor and Veronesi

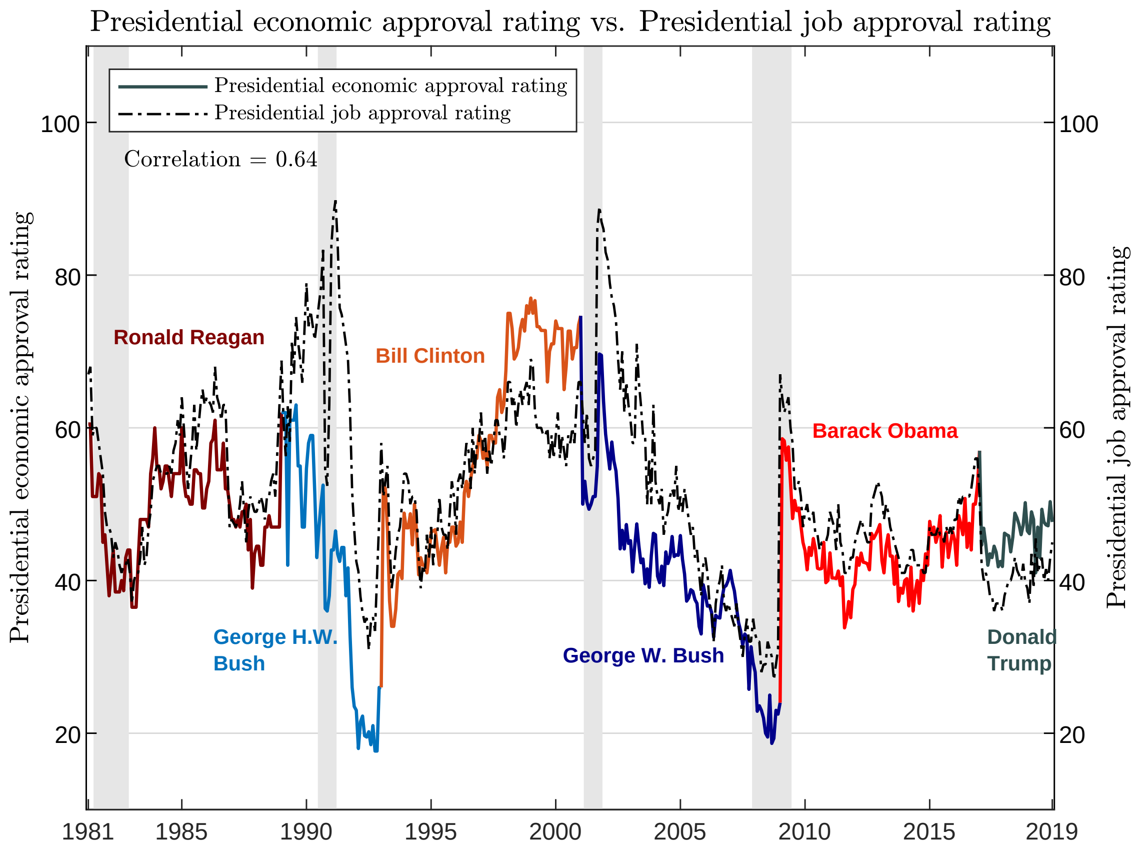


Fig. 3.1 Presidential economic approval rating (PEAR)

This figure depicts the presidential economic approval rating (PEAR) from April 1981 to December 2019, which is based on 1,713 polls conducted by 21 polling organizations and collected by the Roper iPoll at the Roper Center for Public Opinion. It takes the average value if there are multiple polls conducted by different polling organizations in one month. The Gallup presidential job approval rating is also plotted for comparison.

(2020), to the cross-section, would predict the opposite. Since high PEAR-beta stocks do worse precisely when aggregate risk aversion increases (or when PEAR decreases), they are therefore more risky and should earn higher returns on average.

The low-PEAR-beta premium is extremely robust. It survives various factor-based and characteristic-based risk adjustment models. It is not driven by any particular sub-sample periods. For example, it is present during the tenure of each of the six presidents in our sample. It is present in each of the four years of the president's term. It is positive and significant during both Democratic and Republican presidents, or after removing the presidential transition periods (six months surrounding the change of a president). The premium is even larger among large and liquid stocks and it persists for more than a year after portfolio formation. It is robust to different backward rolling windows used to estimate the beta and different methods for computing innovations in PEAR. Finally, it even shows up in other G7 countries and is significant in Canada, Germany, Japan and UK as well.

Existing literature provides evidence that different industries may have differential exposures to presidential policies and government spending [see, e.g., Belo, Gala, and Li (2013) and Addoum and Kumar (2016), among others], which may result in predictable variations in industry portfolio returns across political cycles. The low-PEAR-beta premium is not driven by such an industry-level return predictability, as it is equally strong when we examine industry-demeaned returns.

In Fama-MacBeth cross-sectional regressions, we control for a comprehensive set of potential return predictors which we group into three categories. The first category includes alternative measures of beta, such as the market beta, the beta on the macroeconomic uncertainty of Jurado, Ludvigson, and Ng (2015), and the beta on the Baker and Wurgler (2006) sentiment index (Chen, Han, and Pan, 2020). The second category includes variables related to government and politics. They are the political alignment index (Kim, Pantzalis, and Park, 2012), political sensitivity (Addoum and Kumar, 2016), political connectedness (Cooper, Gulen, and Ovtchinnikov, 2010), and government spending exposure (Belo, Gala, and Li, 2013). The third category includes other firm characteristics such as size, book-to-market, momentum, short-term reversal, idiosyncratic volatility, illiquidity, and distress. Neither of these return predictors is highly correlated with the PEAR beta. Not surprisingly, we find that the coefficient on the PEAR beta remains negative and significant, even after simultaneously including all the control variables and industry fixed effects. Its magnitude is about half of its counterpart in an univariate regression, suggesting that all the other variables, even when combined, explain at most half of the low-PEAR-beta premium.

While the low-PEAR-beta premium cannot be explained by exposure to time-varying risk aversion, can it reflect exposure to other macroeconomic risk factors? We examine a large set of macro variables, including industrial production growth, unexpected inflation, change in expected inflation, term premium, default premium, total factor productivity growth, labor income growth, capital share growth (Lettau, Ludvigson, and Ma, 2019), consumption growth, ultimate consumption growth (Parker and Julliard, 2005), consumption-wealth ratio, aggregate market volatility, VIX, variance risk premium, and GDP growth. The correlations between the change in PEAR and the macro variables are low in general. Even the highest correlation (in absolute term) is only 0.14 (with the capital share growth). As a result, the PEAR beta is not highly correlated with the betas on these macro variables either. In other words, the low-PEAR-beta premium does not seem to capture exposures to these additional risk factors. Including these macro betas in the Fama-MacBeth cross-sectional regressions hardly changes the coefficient on the PEAR beta, consistent with the findings in Shen, Yu, and Zhao (2017) that the exposure to macroeconomic risks generally does not explain the cross-sectional variation in average stock returns very well.

Stocks with positive PEAR betas experience higher returns when the presidential economic approval rating improves. To the extent that PEAR may indicate consumer confidence (De Boef and Kellstedt, 2004), high PEAR-beta stocks could suffer from sentiment-induced overpricing, explaining their subsequent low returns when their overpricing gets corrected. Indeed, Stambaugh, Yu, and Yuan (2015b) find the long-short anomaly return spread to be much stronger following high levels of sentiment. They also find this pattern to be especially true for the short legs of various anomaly strategies, consistent with short-sale impediments. Unfortunately, such sentiment-induced overpricing does not seem to fully explain the low-PEAR-beta premium. We examine four measures of investor sentiment: (1) Baker and Wurgler's (2006) sentiment index, (2) Michigan consumer sentiment index, (3) AAII bull-bear index, and (4) the PEAR index itself. We find significantly higher low-minus-high beta return spreads following high levels of sentiment, only when the PEAR index is used. However, we do not find any evidence that the short-leg (high PEAR-beta stocks) alpha is higher following high levels of sentiment. In fact, in all cases, the long-leg has a higher alpha (in absolute term) than the short-leg does, inconsistent with the notion that short sale constraints may explain the low-PEAR-beta premium.

Intuitively, the PEAR beta could measure a firm's alignment to the economic policies of the current president or the presidential party. The business of a positive PEAR-beta firm must align well with the current presidential economic policies, so its stock price moves in tandem with the policies' approval rating. Could such a "presidential alignment" lead to a government bailout during bad times? If so, a high PEAR-beta stock could be a hedge for downside risk and thus will earn a lower expected return. Empirically, corporate bailouts are relatively rare and tend to happen to mega firms or firms in the finance sector (Faccio, Masulis, and McConnell, 2006). Yet, our sample excludes finance companies and the high-PEAR-beta stocks are not large-cap stocks either. Additional evidence does not support such a "hedging" story either. During bad times, as indicated by NBER recession dates, high-PEAR-beta firms earn even lower returns than low-PEAR-beta firms, inconsistent with the notion of a bailout. In addition, the PEAR beta has a low correlation with the measure of financial distress (Campbell, Hilscher, and Szilagyi, 2008). Controlling for the distress risk does not alter the low-PEAR-beta premium.

Instead, it is possible that the premium could be driven by a novel sentiment towards presidential alignment. Due to such a sentiment, high PEAR-beta firms, which are aligned with the current presidential economic policies, are overpriced currently. The sentiment is not fully appreciated by the market and as a result, the mispricing takes a long time to correct, even if it is on the long side. The correction results in the low-PEAR-beta premium.

We document several pieces of supporting (though not conclusive) evidence for this conjecture. First, if we compute the PEAR beta using only months in the five-year rolling window when a president from a different party were in power, the low-PEAR-beta premium ceases to be significant, highlighting the importance of alignment to the current presidential economic policies. Second, consistent with the mispricing interpretation, we find the PEAR beta to negatively predict analyst forecast errors, future revisions in their long term growth (LTG) forecasts and stock recommendations. In addition, the PEAR beta negatively predicts future earnings announcement returns. The evidence suggests that both analysts and investors are initially too optimistic (pessimistic) in valuing high- (low-) PEAR-beta stocks. Third, we find the low-PEAR-beta premium to be higher and to accrue faster during presidential party transition, consistent with the notion that changing presidential parties facilitate a faster correction to mispricing caused by sentiment to presidential alignment. Finally, stocks in the high and low PEAR-beta deciles tend to receive lower media coverage, potentially contributing to limited investor attention and slow correction to mispricing.

The rest of the paper proceeds as follows. Section 3.2 introduces data and key variables. Section 3.3 presents main results. Section 3.4 excludes existing interpretations. Section 3.5 proposes a new explanation. Section 3.6 concludes.

3.2 Data and Key Variables

This section describes the data on PEAR and other key variables used in this paper.

3.2.1 The PEAR index

To measure public opinion on the president's handling of the economy, we construct a presidential economic approval rating (PEAR) index by using various national polls. Unlike the Gallup presidential job approval rating index that captures the extent to which people approve or disapprove of the way the current president is handling the economy, foreign affair, health policy, etc, we focus on the responses to an economy-specific question: "Do you approve or disapprove of the way (name of president) is handling the economy?", which is closely related to the conceptualization of "confidence in the president's economic stewardship". The data are from Roper iPoll at the Roper Center for Public Opinion.*

Specifically, we collect 2,100 polls in total from 46 organizations over the period from April 1981 to December 2019.† We do not consider several polls irregularly conducted between 1977

*The wording of this question is basically the same across polling organizations, while the predefined responses to the question can be sightly different. Specifically, most polling questions simply ask if a respondent approves or disapproves of the president, while very few questions break out approval or disapproval into subcategories to indicate whether the respondent "strongly" or "somewhat" approves or disapproves of the president. We follow the standard treatment in polling and sum up the percentages of both "strongly" and "somewhat" approve choices as the ratio of approval rating overall.

†Some polls may be conducted by one organization but sponsored by another organization. For example, since 1981, ABC News and The Washington Post, both separately and together, have commissioned polls on this issue.

and 1981. We exclude those organizations conducting less than 5 polls in our sample. We also exclude polls that are conducted in one month but released in the next month (some of them are even lagged for more than one month), so that the public opinion is captured in a timely fashion. In doing so, we are left with 1,713 polls from 21 polling organizations. Hence, each month we have about 3.7 polls on average. Table A1 presents the summary statistics of each polling organization used in the construction of the PEAR index.

From each poll, we obtain an approval rating, a percentage number indicating the proportion of respondents who approve of the way of the president handling the economy. We construct the PEAR index by simply averaging approval ratings available in each month. In our sample period, there are inconsecutive 50 months with missing data. We fill these missing entries with the previous month values to ensure that the PEAR index is a real-time series.

Figure 3.1 plots the time series dynamics of the PEAR index. It also plots the general Gallup presidential job approval rating index for comparison. The two ratings are clearly positively correlated (with a correlation of 64%), yet they also diverge from time to time. Notable examples include the Gulf war, the September 11 terrorist attack, and president Trump's initial tenure. In Section 3.3, we show that using PEAR generates much stronger results than the general job approval rating, consistent with the phrase "the economy, stupid," popularized during Bill Clinton's successful 1992 presidential campaign.

Table 3.1 reports the summary statistics of PEAR and six other sentiment and politics-related indexes, including Baker and Wurgler (2006) (orthogonalized) investor sentiment, Michigan consumer sentiment, presidential job approval rating, (equally-weighted) aggregate political risk and sentiment (Hassan et al., 2019), and political uncertainty [measured by the economic policy uncertainty of Baker, Bloom, and Davis (2016)]. All the time series are at the monthly frequency and over the 1981:04–2019:12 period, except for the quarterly aggregate political risk and sentiment being over 2002Q1–2019Q4, and political uncertainty being over 1985:01–2019:12.

These surveys are conducted by themselves and other organizations, including Chilton Research Services, Taylor Nelson Sofres Intersearch, Langer Research Associates, etc. To ensure data consistency, we classify theses polls as conducted by ABC News, The Washington Post, or both.

Panel A of Table 3.1 presents the mean, median, min, max, volatility, AR(1), and AR(12), where AR(1) and AR(12) refer to the first- and 12th-order autocorrelations. The PEAR index ranges from 17.67 to 77, with a mean of 46.9, suggesting that on average less than half of respondents consent to the way how the president is handling of the economy. Two extreme examples are George H.W. Bush and George W. Bush, whose ratings drop to below 20 at the end of their tenures. In contrast, the presidential job approval rating is generally higher than PEAR, with a mean of 51.65. This pattern is especially pronounced during the presidency of George H.W. Bush and George W. Bush. For example, after the Gulf war, President George H.W. Bush has a job approval rating around 90, but a lugubrious economic approval rating, which is below 50.

To examine the relationships between PEAR and the six other sentiment and politics-related variables, Panel B of Table 3.1 reports their level and change correlations. PEAR is highly positively correlated with Michigan consumer sentiment and presidential job approval rating, with level correlations of 0.63 and 0.64, and change correlations of 0.14 and 0.23, thereby suggesting that these three indexes capture some common low frequent movements, say the presidential cycles, but they capture different salient events at the monthly frequency. Another interesting observation is that PEAR is not highly correlated with political sentiment and political uncertainty, especially with their changes.

3.2.2 PEAR beta

We use PEAR beta to measure how the stock price of a firm responds to the change of PEAR. For each stock and each month from June 1981, we run the following time series regression with a 60-month rolling window, requiring at least 24 observations,

$$R_{i,t} = \alpha + \beta_{i,0}\Delta\text{PEAR}_t + \beta_{i,1}\Delta\text{PEAR}_{t-1} + \varepsilon_{i,t}, \tag{3.1}$$

where $R_{i,t}$ is the excess return of stock i in month t, and ΔPEAR_t is the change of PEAR from month $t-1$ to month t. The regression includes the lagged change of PEAR to accommodate the non-synchronicity between the timing of the survey and that of return measurement. Following Dimson (1979), PEAR beta, β_{PEAR}, is defined as

$$\beta_i = \beta_{i,0} + \beta_{i,1}, \tag{3.2}$$

where we abbreviate the time subscript for brevity. This specification is commonly used in the literature. For example, Fama and French (1992), Antoniou, Doukas, and Subrahmanyam (2016), and Liu, Stambaugh, and Yuan (2018) estimate an individual stock's market beta in the same way as (3.2), regressing monthly returns on the current and lagged market returns using a 60-month rolling window and then summing up the coefficients.

Since we require at least 24 months of non-missing observations for each stock to run the regression, we use an expanding window over the 1981:06–1983:05 period and a fixed 60-month rolling window after June 1983. Thus, our empirical analysis spans the 1983:06–2019:12 period, 439 months in total.

3.2.3 Other variables

We obtain monthly stock returns from the Center for Research in Security Prices (CRSP) and quarterly and annual accounting data from Compustat. Our data sample includes all common stocks listed on the NYSE, Amex, and Nasdaq exchanges. Financial and Utility firms are excluded from our analysis. In addition, we exclude stocks with a price per share less than $1 and stocks with missing returns. We adjust stock returns for delisting to avoid survivorship bias following Shumway (1997).

We estimate market beta (β_{MKT}) and sentiment beta (β_{BW}) the same as β_{PEAR}, and uncertainty-beta (β_{UNC}) as Bali, Brown, and Tang (2017). We calculate firm size (SIZE) as the logarithm of the product of price per share and the number of shares outstanding (in millions of dollars). The logarithm of book to market ratio (BM) is calculated as the book value of shareholder equity plus deferred taxes and investment tax credit (if available) minus the book value of preferred stocks at the end of the last fiscal year, $t-1$, scaled by the market value at the end of December of year $t-1$.* Firms with negative book values are excluded from the analysis. We follow Fama and French (1992) and match the annual BM information in year $t-1$ to monthly returns from July of year t to June of year $t+1$.

We define momentum (MOM) as the cumulative return of a stock over a 11-month window ending one month before the portfolio formation. Short-term reversal (STR) is defined as the stock return over the prior month. Following Ang, Hodrick, Xing, and Zhang (2006), the monthly idiosyncratic volatility (IVOL) is the standard deviation of the stock's daily idiosyncratic returns relative to the Fama and French (1993) three-factor model over the prior month. We measure the illiquidity (ILLIQ) of a stock as the ratio of the daily absolute stock return to the daily dollar trading volume averaged in the prior month, which is further is scaled by 10^6 (Amihud, 2002). A stock is required to have at least 15 valid daily returns to calculate the IVOL and ILLIQ. The mispricing score (MISP) is from Stambaugh, Yu, and Yuan (2015b), which is a rank variable constructed by 11 anomalies. The higher the score, the more likely the stock is overvalued. MISP ends in 2016 and we extend it to 2019 by ourselves.

We consider four politics-related variables. Following Kim, Pantzalis, and Park (2012), we use the state-level political alignment index (PAI) of each state's leading politicians with the ruling (presidential) party to proxy for local firms' proximity to political power. We use political sensitivity (PS) to capture how a firm's stock price fluctuates over the presidential cycles (Addoum and Kumar, 2016). We define political connectedness (PC) as a dummy variable as to whether a firm makes a contribution to the PAC (regardless of party affiliation) in the last 5 years following Cooper, Gulen, and Ovtchinnikov (2010) and Addoum and Kumar (2016). As in Belo, Gala, and Li (2013), we calculate the industry-level government spending exposure (GSE) as the proportion of the industry's total output (3-digit SIC) being purchased by the government sector for final use to capture the impact of political cycles on asset prices. Table A2 details the construction of these

*Depending on availability, the stockholders' equity, common equity plus the carrying value of preferred stock, or total assets minus total liabilities in that order is used as shareholders' equity. Similarly, we use redemption, liquidation, or par value in that order depending on availability to estimate the book value of preferred stocks.

variables.

Table 3.2 reports the autocorrelations and pairwise correlations of the key variables used in this paper. In Panel A, the first- and 12th- autocorrelations of PEAR beta are 0.82 and 0.34, suggesting that the effect of PEAR on the cross-section of stock returns, if there is any, is persistent and decays dramatically after one year. This implication is empirically confirmed in Section 3.3.1.

Panel B of Table 3.2 shows that PEAR beta has low correlations with all other variables. The absolute values are all smaller than 0.10, except for market beta, which is 0.16. For example, PEAR beta has negligible correlations with the four politics-related variables (PAI, PS, PC, and GSE) and MISP. This result implies that the PEAR-beta effect on stock returns is unlikely to be explained by these variables or the economic mechanisms behind them.

3.3 Empirical Results

In this section, we conduct portfolio analyses and Fama-MacBeth regressions to assess the predictive power of PEAR beta over future stock returns. We perform a number of tests to show that our results are robust qualitatively and quantitatively.

3.3.1 Average and risk adjusted returns of PEAR-beta decile portfolios

At the end of each month from June 1983 to November 2019, we form decile portfolios by sorting firms into ten groups based on their PEAR betas, where decile 1 (10) contains stocks with low (high) PEAR betas. We value-weight these portfolios and rebalance them at a monthly frequency. The PEAR-beta speard portfolio refers to the strategy that buys stocks in decile 1 and sells stocks in decile 10.

Panel A of Table 3.3 reports the portfolio sorting results. The first row presents the average PEAR betas of the decile portfolios, which increase from -1.30 for decile 1 to 2.32 for decile 10. The range of PEAR beta is much larger than the market-beta portfolios (from 0.59 to 1.51) in Liu, Stambaugh, and Yuan (2018) and the uncertainty-beta portfolios (from -0.62 to 0.72) in Bali, Brown, and Tang (2017). In the second row, the monthly average excess returns of the PEAR-beta portfolios decrease from 1.23% for decile 1 to 0.11% for decile 10, with the difference between the low and high PEAR-beta portfolios equal to 1.11% (t-value = 4.13).

We calculate the risk adjusted returns of the PEAR-beta portfolios with five factor models and the Daniel et al. (1997) characteristics model (DGTW). The five factor models include the Fama and French (2015) five-factor model (FF5), the Hou, Xue, and Zhang (2015) q-factor model (HXZ), the Stambaugh and Yuan (2017) mispricing-factor model (SY), the Daniel, Hirshleifer, and Sun (2020) behavioral-factor model (DHS), and the FF5 model augmented by the betting-against-beta factor (BAB) of Frazzini and Pedersen (2014).

The rest rows (row 3 to row 8) of Panel A make three observations. First, although the six models we use represent the most recent development of asset pricing, they cannot explain the PEAR-beta portfolios well. The abnormal return of the PEAR-beta spread portfolio ranges from 0.66% with the DHS model to 0.90% with the FF5 model, suggesting that at least 60 percent of the average return of the PEAR-beta spread portfolio is not explained by existing asset pricing factors or firm characteristics. Second, unlike the well-known anomalies in Stambaugh, Yu, and Yuan (2015b), the performance of the PEAR-beta spread portfolio is mainly from the long-leg. The low PEAR-beta portfolio is generally undervalued, whereas the high PEAR-beta portfolio is fairly priced. The only exception is the DGTW model, where the low and high PEAR-beta portfolios are both mispriced. For this reason, we interchangeably use the terms of the PEAR-beta spread portfolio and the low-PEAR-beta premium. Third, the low-PEAR-beta premium is unrelated to the market-beta anomaly, and its abnormal return is unchanged even the BAB factor is controlled.

Panel B of Table 3.3 reports the results of portfolios sorted by industry demeaned PEAR betas, where 48 industries are classified following Fama and French (1997). If the low-PEAR-beta premium is an industry-level phenomenon, such as Belo, Gala, and Li (2013) and Addoum and Kumar (2016), the average PEAR betas of the decile portfolios after industry demeaning should have a small spread, and the low-PEAR-beta premium should become negligible.

The results in Panel B show that the industry effect contributes a small fraction of the low-

PEAR-beta premium. The average PEAR betas increase from -1.68 for decile 1 to 1.78 for decile 10, with the difference quantitatively close to the case without industry demeaning (-3.46 vs. -3.62). The average returns of the PEAR-beta portfolios decrease from 0.98% for decile 1 to 0.13% for decile 10, with the difference equal to 0.85% (t-value $= 4.15$). This value suggests that the industry effect explains only a quarter of the low-PEAR-beta premium (1.11%).

When turning to the risk adjusted return, the low-PEAR-beta premium also slightly decreases. It ranges from 0.53% for the DHS model to 0.79% for the FF5 model. All the values are statistically significant and economically sizeable. Because the low-PEAR-beta premium cannot be explained by all the asset pricing models, for simplicity, we use the FF5 as the benchmark for calculating the risk adjusted returns in the sequel.

To explore how much an investor can make if he consecutively trades for the low-PEAR-beta premium, Figure 3.2 plots the log cumulative FF5 alphas of the PEAR-beta spread portfolio. In our sample period from June 1983 to December 2019, the investor can make a risk adjusted profit of $31.82, which does not suffer from large drawdowns. Thus, trading the PEAR-beta spread portfolio can greatly expand an investor's investment opportunities in our sample period.

In this paper, we rebalance the PEAR-beta portfolios at the monthly frequency. A natural question is how long the low-PEAR-beta premium persists. Figure 3.3 presents the average returns of the PEAR-beta spread portfolio up to 36 months after formation. With 1.96 as the critical value for significance, the figure shows that the low-PEAR-beta premium is persistent and generally significant up to 12 months after formation. Moreover, the premium does not display a reversal pattern. This result is comparable with the uncertainty-beta premium documented in Bali, Brown, and Tang (2017), which is persistent and significant up to 11 months.

In sum, this subsection shows that high PEAR-beta stocks under-perform low PEAR-beta stocks in the future in terms of average and risk adjusted returns, which we label as the low-PEAR-beta premium. A strategy trading for this premium generates statistically and economically significant profits.

3.3.2 Robustness

This subsection performs a battery of robustness checks to show that the low-PEAR-beta premium is not specific to a sub-sample or a sub-period, and is robust to different estimation methods.

Performance over political cycles

The well-known presidential puzzle refers to the striking time series fact that stock market returns are much higher under Democratic presidencies than Republican ones. While our low-PEAR-beta premium is a cross-sectional phenomenon, one may be still curious if it is also stronger under Democratic presidencies.

We split the sample into two sub-periods, Democratic and Republican. A month is defined to be Democratic if the president is a Democrat in that month. Since the inauguration of a new president is always around the 20th of January, we assume February is the commencement of the four-year term as a new president. In doing so, we have identified 192 months as Democratic and 247 months as Republican. Panel A of Table 3.4 reports the average and risk adjusted returns of the PEAR-beta spread portfolio in these two sub-periods. The average return is 1.25% (t-value = 3.06) under Democratic presidencies and 1.01% (t-value = 2.81) under Republican presidencies, with the difference (0.24%) insignificant from zero (t-value = 0.44). The risk adjusted returns are 1.36% (t-value = 3.86) and 0.55% (t-value = 1.81) under Democratic and Republican presidencies, respectively. In this case, the difference is 0.81% and marginally significant (t-value = 1.67).[*]

Figure 3.4 goes one step further by plotting the average and risk adjusted returns of the PEAR-beta spread portfolio within each president tenure. Our sample covers six presidents, two Democrats and four Republicans. The figure shows that while the low-PEAR-beta premium is stronger during Democratic presidencies, it is also strong during a Republican presidency. Indeed, in the four-year term of president George H.W. Bush, the PEAR-beta spread portfolio has an average return of 1.82% and an FF5 alpha of 1.46% per month, which is comparable, and even

[*]If we restrict the polling agents to those with at least 20 polls, the risk adjusted returns of the PEAR-beta spread portfolio would be 1.53% (t-value = 3.08) and 0.61% (t-value = 2.29) during Democratic and Republican presidencies, with the difference equal to 0.92% (t-value = 1.83).

slightly better than, with the president Bill Clinton term (1.49% and 1.55%). Of course, the worst performance is also from the Republican presidency, Donald J. Trump, which has mediocre average and risk adjusted returns (0.19% and 0.34%)

Performance over presidential transition and non-transition periods

A stock with a positive PEAR beta during a Republican presidency can be thought of as a Republican firm. Its stock price is likely to suffer when a Democratic president is elected. If the market is efficient, then most of the price adjustment should occur during the presidential period. Empirically, Addoum and Kumar (2016) and Meeuwis et al. (2020) find that investors rebalance their portfolios dramatically around president elections, because of political climate change or political disagreement. To explore if such presidential transitions explain our finding, we perform two tests in this subsection.

First, we split the sample into transition and non-transition periods. A transition period consists of six months before and after a new president's inauguration. With six presidents, we have five transitions, covering 65 months in total. Panel B of Table 3.4 shows that the average return in the transition period is higher than that in the non-transition period (2.25% vs. 0.91%), but the risk adjusted returns are statistically indifferent in these two sub-periods, with the difference equal to -0.48 (t-value $= -0.75$). The result is similar if we use November of the election year as the event month as in Brogaard et al. (2020).

Next, we examine how the PEAR-beta spread portfolio performs across the four years of a president tenure. In the literature, Belo, Gala, and Li (2013) show that the government spending exposure has stronger power in predicting future stock returns in years 2 and 3 of a president tenure. In contrast, Addoum and Kumar (2016) find that stock prices are more sensitive to the political climate change in the first and fourth years. Figure 3.5 shows that the low-PEAR-beta premium is different from Belo, Gala, and Li (2013) and Addoum and Kumar (2016). Its performance, especially after risk adjustment, is stronger in the first and second years during a president term. Importantly, the low-PEAR-beta premium is present in each of the four years.

Performance over NBER recessions and expansions

As shown in Pastor and Veronesi (2020), financial crises or economic recessions are more likely to happen during a Republican president's term, which raises an interesting question that whether the low-PEAR-beta premium is weaker during economic recessions, given the time series presidential puzzle.

When splitting the smaple period into NBER-dated economic recessions and expansions, we find that the low-PEAR-beta premium is stronger in NBER recessions. Specifically, the average return and FF5 alpha are 2.86% and 1.63% in recessions, whereas the counterparts in NBER expansions are 0.95% and 0.83%. This result is reported in Panel C of Table 3.4, and has two immediate implications. First, although the low-PEAR-beta premium is stronger under the Democratic presidencies, it can be even stronger over economic downturns during a Republican presidency. Second, high PEAR-beta firms do not perform better than those with low PEAR betas, suggesting in turn that they do not benefit from the Republican president or party policies.

Performance among different firms

Limits-to-arbitrage or transaction costs are an important determinant of mispricing, and plague the existing asset pricing models (Fama and French, 2015; Hou, Xue, and Zhang, 2015). In this subsection, we examine how the low-PEAR-beta premium performs among firms with low and high limits-to-arbitrage.

We consider three measures of limits-to-arbitrage, IVOL (Ang et al., 2006), illiquidity (Amihud, 2002), and firm size. For each measure, at the end of each month, we sort firms into two subgroups and construct a PEAR-beta spread portfolio within each subgroup. Panel D of Table 3.4 reports the results with IVOL. Surprisingly, the low-PEAR-beta premium is even stronger among low IVOL stocks, although statistically insignificant. Its FF5 alpha is 0.96% (t-value = 3.81) among low IVOL stocks, and 0.69% (t-value = 2.17) among high IVOL stocks. This empirical pattern continues to hold when we measure limits-to-arbitrage with Amihud's (2002) illiquidity or firm size (Panels E and F). The FF5 alphas of the low-PEAR-beta premiums are 0.95% and

0.15% among liquid and illiquid stocks, and 0.98% (t-value = 3.20) and 0.40% (t-value = 2.02) among big and small firms, respectively. These finding imply that the low-PEAR-beta premium is economically significant as it goes beyond transaction costs. Thus, it is different from most of anomalies that are concentrated among small firms (Hou, Xue, and Zhang, 2015).

Alternative PEAR-beta estimates

In this paper, we estimate the PEAR beta with (3.2). Because the PEAR beta and market beta have a correlation of 0.16 (Table 3.2), one natural question is what happens if we control for the market return when estimating the PEAR beta.

To answer the above question, we include the market return in regression (3.1) to estimate the PEAR beta, redo the single portfolio sorting as Table 3.3, and report the results in Panel G of Table 3.4. In this case, the average and risk adjusted returns of the PEAR-beta spread portfolio are 0.92% (t-value = 4.00) and 0.93% (t-value = 4.07), which are close to the case without controlling for the market beta (1.11% and 0.90%). Untabulated results show that including the lagged market return in regression (3.1) does not affect the low-PEAR-beta premium either.

We also examine the robustness to different rolling windows used to estimate the PEAR beta, four and eight years (coinciding with one or two presidential terms). The results are quantitatively similar to the baseline results with a five-year rolling window. Thus, the low-PEAR-beta premium is robust to alternative estimation methods.

Alternative PEAR indexes

In this subsection, we show that the low-PEAR-beta premium is robust to three variations to the construction of the PEAR index.

First, in the main analyses, we use the change of PEAR to calculate the PEAR beta, and implicitly assume that it is independent over time, which may not be true empirically. To address this concern, we assume that the change of PEAR follows an AR(1) process and use the residual to estimate the PEAR beta. Panel H of Table 3.4 shows that, with this variation, the average and risk adjusted returns of the PEAR-beta spread portfolio are 0.78% (t-value = 2.81) and 0.68% (t-value = 2.85), which are slightly weaker than the baseline results. A caveat here is that the AR(1) estimation uses the full sample and thus introduces a forward-looking bias. We thus prefer estimating the PEAR beta using the simple changes.

Second, as shown in Table 3.1, the presidential job approval rating index is highly, positively correlated with PEAR. So one interesting question is whether this alternative index can generate similar results. Panel H of Table 3.4 reports the average return and FF5 alpha are 0.65% (t-value = 2.76) and 0.42% (t-value = 1.73), respectively. These values are smaller than those using PEAR, suggesting that PEAR is more relevant for the financial market. This finding is consistent with the phrase "the economy, stupid," popularized during Bill Clinton's successful 1992 presidential campaign.

Lastly, we consider polls from top 6 polling organizations, which conduct the most surveys in our sample period. In this case, since there are many missing values, especially in the early years, we fill in the missing values using the dyad ratios algorithm of Stimson (1999). This algorithm is popular in political science, and it uses smoothing and interpolation to deal with irregular, non-balanced, and sparse panel data. By using this new index, the PEAR-beta spread portfolio has an average return of 1.10% and an abnormal return of 1.11%. There are two possible reasons for this better performance. First, because of interpolation, the PEAR index by using the dyad ratios algorithm is not a real time series, the outperformance may therefore reflect a look-ahead bias. Second, the top 6 polling agents are more sophisticated and provide more accurate polls than other agents.

To conclude, this subsection shows that the low-PEAR-beta premium is largely robust to alternative methods for constructing the PEAR index.

3.3.3 International evidence

In this subsection, we conduct an out-of-sample test by showing that the low-PEAR-beta premium continues to hold in other G7 countries. That is, the US PEAR index also affects the stock returns of other G7 countries.

Specifically, we collect firm-level stock returns and marketcaps of Canada, France, Germany, Italy, Japan, and UK from Datastream, and use similar filters as Griffin, Kelly, and Nardari (2010) and Hou, Karolyi, and Kho (2011).* Because the results using US dollar and local currencies are similar, we report the results with local currencies in Table 3.5. Same as the baseline, all portfolios are valued-weighted and rebalanced at the monthly frequency. The sample period is 1983:05–2019:12. The FF5 factor data are from Schmidt et al. (2019) and only available after 1990. As such, the risk adjusted returns of the PEAR-beta spread portfolios are based on the 1991:07–2019:12 period for all countries, except from Japan that starts from 1990:07.

Overall, Table 3.5 shows that the low-PEAR-beta premium exists in all the other G7 countries. The average and risk adjusted returns of the PEAR-beta spread portfolios are positive in all the countries. In particular, the PEAR-beta spread portfolios have significant average and risk adjusted returns in Canada, Japan, and UK. A mixed result comes from Germany, which has an insignificant average return but a significant FF5 alpha. One possible reason is that the average return and alpha are measured in different sample periods, and the low-PEAR-beta premium is weaker before 1991. Among the six countries, the strongest performance comes from Canada, partially because investors in Canada are more economically and geographically linked to the US.

3.3.4 Fama-MacBeth regressions

So far we have tested the significance of the PEAR beta as a determinant of the cross-section of future returns at the portfolio level. This portfolio-level analysis has non-parametric merit in the sense that we do not impose a functional form on the relation between the PEAR beta and future returns. However, it also has two disadvantages. First, it gives up a large amount of information in the cross-section via aggregation. Second, it is hard to control for multiple effects or factors

*In particular, we require that firms selected for each country are domestically incorporated based on their home country information (GEOGC); We eliminate non-common stocks such as preferred stocks, warrants, REITs, and ADRs. If a stock has multiple share classes, only the primary class is included. To filter out suspicious stock returns, we set returns to missing for stocks with returns higher than 300%. Specifically, if R_t or R_{t-1} is greater than 300%, and $(1+R_t)\times(1+R_{t-1})-1<50\%$, then both R_t and R_{t-1} are set to missing. We also treat the monthly returns as missing that fall outside the 0.1% to 99.9% range in each country.

simultaneously using portfolio-level analysis. To address these concerns, in this subsection we run Fama-MacBeth cross-sectional regressions of firms' one-month-ahead excess returns on their PEAR betas and various firm- and industry-specific characteristics to gauge the incremental return predictive power of the PEAR beta.

In Fama-MacBeth cross-sectional regressions, we control for a comprehensive set of potential return predictors which we group into three categories. The first category includes alternative measures of beta, such as the market beta, the beta on the Jurado, Ludvigson, and Ng (2015) macroeconomic uncertainty index (Bali, Brown, and Tang, 2017), and the beta on the Baker and Wurgler (2006) sentiment index (Chen, Han, and Pan, 2020). The second category includes variables related to government and politics. They are the political alignment index (Kim, Pantzalis, and Park, 2012), political sensitivity (Addoum and Kumar, 2016), political connectedness (Cooper, Gulen, and Ovtchinnikov, 2010), and government spending exposure (Belo, Gala, and Li, 2013). The third category includes other firm characteristics such as size, book-to-market, momentum, short-term reversal, idiosyncratic volatility, and illiquidity.

Table 3.6 reports the results. In column 1, the univariate regression shows that the PEAR beta has a significantly negative coefficient of -0.17 with a t-value of -2.73. Economically, the absolute t-value is proportional to the Sharpe ratio of the PEAR-beta spread portfolio, which equals to the annualized Sharpe ratio times $\sqrt{T}$, the number of years in the sample. So the -2.73 t-value suggests that an investor can earn an annualized Sharpe ratio of 0.45 (i.e., $2.73/\sqrt{37}$) if he trades for the low-PEAR-beta premium. This value is comparable to the market Sharpe ratio of 0.51. In column 2, when we control for firm characteristics in the regression, the coefficient of PEAR beta drops to -0.12 but the t-value slightly decreases to -2.64 in magnitude, suggesting that the predictive power of PEAR beta is robust to these known firm characteristics.

In column 3, when we further include other beta predictors (i.e., β_{MKT}, β_{UNC}, and β_{BW}), the regression coefficient on PEAR beta slightly changes to -0.11 with a t-value of -3.26. Interestingly, the sentiment beta, β_{BW}, has a significantly negative regression coefficient in this case, consistent with the argument in Baker and Wurgler (2006). The market beta and uncertainty beta lose power. In column 4, we instead control for political variables (i.e, political alignment index, political sensitivity, political connectedness, and government spending exposure), and find the coefficient of PEAR beta to be -0.10 (t-value $= -2.35$). This result suggests that the interpretations underlying these politics-related variables are unlikely to completely explain the low-PEAR-beta premium.

In column 5, when we pool all the three categories of controls in the regression, the coefficient on PEAR beta remains -0.09 with a t-value of -2.43. This magnitude is about half of its counterpart in column 1, suggesting that all the controlling variables, even when combined, explain at most half of the low-PEAR-beta premium. This result is not surprising because, as we have shown in Table 3.2, PEAR beta has low correlations with these variables.

In column 6—the last column of Table 3.6—we run the Fama-MacBeth regression by controlling for the Fama-French 48 industry fixed effects. We drop the industry-level predictor political sensitivity as it is calculated based on Fama-French 48 industries. The regression coefficient of PEAR beta remains -0.08 with a t-value of -2.34. Again, the low-PEAR-beta premium is different from Belo, Gala, and Li (2013) and Addoum and Kumar (2016), and it is not an industry-level phenomenon.

In sum, a significant part of the low-PEAR-beta premium cannot be explained by existing well-known return predictors.

3.4 Interpreting the Results

In this section, we provide four potential channels to understand our main findings, and show that they are at most partially explain the low-PEAR-beta premium.

3.4.1 Risk aversion

To explain the presidential puzzle, Pastor and Veronesi (2020) develop a model of political cycles driven by time-varying risk aversion. They argue that when risk aversion is high, agents are more likely to elect Democrats that promise more redistribution. In contrast, when risk aversion is low, agents are more likely to elect Republicans to take more business risk. With risk aversion as an exogenous driver, the risk premium of the stock market is expected to be high during Democratic presidencies and low during Republican presidencies.

From Figure 3.1, the PEAR index seems to correlate well with the political cycles in that a low (high) PEAR predicts a Democratic (Republican) president in the future, suggesting that PEAR is inversely related to the aggregate risk aversion modeled in Pastor and Veronesi (2020).

More formally, we find that the PEAR index positively predicts the fraction of a Republican being the president in the following eight years, and negatively predicts the following eight-year market returns:

$$\text{Fraction of Republicans}_{t+1,t+96} = a + 0.18 \times \text{PEAR}_t + u_{t+1,t+96}, \quad t\text{-value} = 9.31, \tag{3.3}$$

$$R_{t+1,t+96} = \alpha - 38.60 \times \text{PEAR}_t + \varepsilon_{t+1,t+96}, \quad t\text{-value} = -2.26 \tag{3.4}$$

where PEAR is normalized in the regressions and the t-values are based on the Newey-West standard errors that handle the autocorrelations with a lag of 95.

In addition, we consider four different measures of aggregate risk aversion, including the unemployment rate, aggregate risk aversion from Miranda-Agrippino and Rey (2020), negative of the surplus consumption ratio from the habit model of Campbell and Cochrane (1999), and option-based risk aversion from Faccini et al. (2019). Figure 3.6 shows that PEAR is indeed negatively correlated with these four risk aversion measures, and the coefficients of regressing these measures on PEAR are always negative and significant, thereby PEAR appearing to be capturing aggregate risk aversion.

In the cross-section, however, a standard risk model would predict the opposite of the low-PEAR-beta premium. If PEAR measures the negative of risk aversion, high PEAR-beta stocks do worse precisely when aggregate risk aversion increases (or when PEAR decreases), and they are therefore more risky and should earn higher returns. Such a risk-based story is therefore inconsistent with our empirical findings that high PEAR-beta stocks under-perform the low PEAR-beta stocks in the future.

3.4.2 Macroeconomic risk

Although risk aversion does not provide a full explanation to our findings, it is possible that the low-PEAR-beta premium actually reflects exposure to other macroeconomic risk factors. We examine this possibility by studying a large set of macro variables, including industrial production growth, unexpected inflation, change in expected inflation, term premium, default premium, total factor productivity growth, labor income growth, capital share growth (Lettau, Ludvigson, and Ma, 2019), consumption growth, ultimate consumption growth (Parker and Julliard, 2005), consumption-wealth ratio, aggregate market volatility, VIX, variance risk premium, and GDP growth.

Panel A of Table 3.7 presents the correlation between the chagne of PEAR and the macro variables. Generally, the correlations are very low, and the highest one is 0.14 between the change of PEAR and the capital share growth. However, according to Lettau, Ludvigson, and Ma (2019), capital share growth is impossible to be an explanation to the low-PEAR-beta premium, because it demands a positive risk premium whereas the change of PEAR suggests a negative risk premium.

Panel B of Table 3.7 reports the correlations between PEAR beta and the marco betas. To mitigate the potential outlier effects, we also consider the rank correlations that are calculated based on the cross-sectional ranks of these betas. In this panel, the PEAR beta has the highest correlation with the VRP beta, 0.15, in absolute term. However, the VRP beta is unable to explain the low-PEAR-beta premium either, because it does not have any power in predicting future stock returns in our sample period.

Therefore, although there is always the possibility that PEAR captures a state variable related to macroeconomic risk, and that PEAR beta and its pricing dynamically vary with this state variable, it seems safe to conclude that existing rational channels are unable to fully explain the low-PEAR-beta premium.

3.4.3 Investor sentiment

Because the PEAR index is based on the responses to "Do you approve or disapprove of the way (name of president) is handling the economy?", one may interpret it as a measure of investor sentiment like the Michigan consumer sentiment index. In this way, stocks with positive PEAR betas experience higher returns when the presidential economic approval rating improves. To the extent that PEAR may indicate consumer confidence (De Boef and Kellstedt, 2004), high PEAR-beta stocks could suffer from sentiment-induced overpricing, explaining their subsequent low returns when their overpricing gets corrected. Indeed, Stambaugh, Yu, and Yuan (2012) find the long-short anomaly return spread to be much stronger following high levels of sentiment. They also find this pattern to be especially true for the short legs of the anomaly strategies, consistent with short-sale impediments.

Unfortunately, such sentiment-induced overpricing does not seem to fully explain the low-PEAR-beta premium. We examine four measure of investor sentiment: (1) Baker and Wurgler (2006) sentiment index, (2) Michigan consumer sentiment index, (3) AAII bull-bear index, and (4) the PEAR index itself. We split the sample into two subsamples based on the median values of the four sentiment measures, and examine the difference of the low-PEAR-beta premium in these two periods. In Table 3.8 we find significantly higher PEAR-beta spread portfolio returns following high levels of sentiment, only when the PEAR index is used. However, we do not find any evidence that the short-leg (high PEAR-beta stocks) alpha is higher following high levels of sentiment. In fact, in all cases, the long-leg has a higher alpha (in absolute term) than the short-leg does, inconsistent with the notion that short sale constraints may explain the low-PEAR-beta premium.

3.4.4 Hedge for downside risk?

Intuitively, the PEAR beta could measure a firm's alignment to the economic policies of the current president or the presidential party. The business of a positive PEAR-beta firm must align well with the current presidential economic policies, so its stock price moves in tandem with the policies' approval rating. It is possible that such a "presidential alignment" leads to a government bailout during bad times. If that happens, a high PEAR-beta stock could actually be a good hedge for downside risk. Could their lower future returns reflect the hedging benefits? We believe the answer is No.

Empirically, corporate bailouts are relatively rare. For stance, Faccio, Masulis, and McConnell (2006) finds that over a sample period from 1997 to 2002, of the 450 political connected firms from 35 countries, only 51 firms received bailouts. In the U.S., financial firms, especially banks, are more likely to be bailed out since these firms are deeply intertwined with the economy through debts and obligations, as evident by a list of historical bailouts in the U.S. collected by the non-profit investigative journalism group, ProPublica. However, financial firms are excluded in our analysis. For non-financial firms, only those mega firms have higher chances of receiving bailouts. These firms are less likely to enter our PEAR beta decile 1 and 10 portfolios, as firms in these extreme portfolios are smaller compared to firms in other portfolios.

Additional evidence does not support such a "hedging" story either. During bad times, as indicated by NBER recession dates, high-PEAR-beta firms earn even lower returns than low-PEAR-beta firms (see Panel C of Table 3.4), inconsistent with the notion of a bailout. In addition, the PEAR beta has a low correlation, 0.10 as shown in Panel B of Table 3.2, with the measure of financial distress (Campbell, Hilscher, and Szilagyi, 2008). Also, Table 3.6 shows that controlling for the distress risk does not alter the low-PEAR-beta premium.

In sum, this section presents four possible interpretations and finds none of them seems promising to explain the low-PEAR-beta premium, thereby suggesting us searching for a new economic mechanism, which is the focus of the next section.

3.5 PEAR Beta and Presidential Alignment

Instead, it is possible that a sentiment regarding the current presidential alignment may cause the market to misprice stocks with high- and low-PEAR-betas. As a concrete example, consider two energy companies: Renewable Energy Group Inc (NASDAQ: REGI) and Panhandle Oil & Gas Inc (NYSE: PHX). As the name implies, the first company benefits from Obama era's clean energy policy while the second company, being a traditional gas and drilling firm, benefits more from the energy policy from the Trump's administration.

Their current presidential alignments are nicely captured by their PEAR betas, as evident in Panel A of Figure 3.7. During the Obama's presidency (2014-2016), Renewable Energy has a large and positive PEAR beta and Panhandle has a negative PEAR beta. After Trump's election in 2017, their PEAR betas start to converge. After a year, they flip. Renewable Energy has a negative PEAR beta while Panhandle has a positive PEAR beta. In this example, PEAR beta becomes a self-revealed and dynamic measure of a firm's alignment with the current presidential economic policies.

Sentiment regarding the current presidential alignment may cause the market to misprice these two stocks. Specifically, during the Obama's term, Renewable Energy tends to overpriced and Panhandle underpriced. The gradual correction to such mispricing results in the low-PEAR-beta premium.

We document four pieces of supporting (though not conclusive) evidence for this conjecture. First, in each month, we split the past 60 months into two sub-samples (if applicable), one coming from the current presidential party and the other from the opposite presidential party, with a requirement of at least 12 observations in each subsample. We then estimate two betas for that firm using each subsample: a current party beta and an opposite party beta. Finally, we explore the low-PEAR-beta premium with respect to these two betas. Panel A of Table 3.9 shows that the low-PEAR-beta premium exists for both betas, but it is only significant when the current party beta is used for sorting. For example, the FF5 alpha of the PEAR-beta spread portfolio is 1.08% (t-value = 4.72) with the current party beta, whereas it is only 0.35% (t-value = 0.89) with the opposite party beta. The evidence highlights the importance of alignment to the *current* presidential economic policies.

Second, consistent with the mispricing interpretation, we find the PEAR beta to negatively predict analyst forecast errors, future revisions in their long term growth (LTG) forecasts and stock recommendations. In addition, the PEAR beta negatively predicts future earnings announcement window returns. The results can be found in Panel B of Table 3.9, which suggest that both analysts and investors are initially too optimistic (pessimistic) in valuing high (low) PEAR-beta stocks.

Third, we expect the low-PEAR-beta premium to be corrected more quickly when the presidential party changes. In the illustrative example, the PEAR betas of Renewable Energy and Panhandle quickly flip after presidential transition. Figure 3.7 confirms that PEAR betas of stocks in decile 10 and 1 converge faster after a presidential transition (Panel C) than during the full sample (Panel B). Consistently, Panel B of Table 3.4 reports a bigger low-PEAR-beta premium during such transition periods. Figure 3.8 further plots the average returns of the PEAR-beta spread portfolios formed during the transition period. The figure shows that the low-PEAR-beta premium is much higher immediately after formation, but also drops quickly and ceases to be significant after only 3 months. This result is in contrast to Figure 3.3 where the low-PEAR-beta premium is lower initially but also significant for up to 12 months. Overall, the evidence suggests that the mispricing captured by PEAR beta is corrected more quickly during the transition periods when investors' sentiment towards the previous president's economic policies weakens.

Finally, we provide evidence that limited investor attention is one reason why mispricing induced by such a presidential sentiment is persistent. We consider two proxies for investor attention, abnormal google search volume and abnormal trading volume. Same as Da, Engelberg, and Gao (2011), the abnormal google search volume is calculated as the change of google search volume in the current month relative to the average search volume in the past two months. Following Gervais, Kaniel, and Mingelgrin (2001), the abnormal trading volume is defined as the change of trading volume (turnover) in the last week of each month relative to the trading volume in the previous nine weeks. Each month, we sort firms into two subgroups based on one of the attention measures, and within each subgroup we further sort firms into deciles based on the PEAR beta. Panel C of Table 3.9 reports the FF5 alphas of these portfolios. As expected, the low-PEAR-beta premium is much stronger among firms that receive less investor attention. For example, the FF5 alphas of the PEAR-beta spread portfolios are 1.22% (t-value = 3.68) and 0.51% (t-value = 1.47) in the low and high abnormal google search volume stocks, respectively. The result by using the abnormal trading volume is also similar (1.02% vs. 0.55%).

3.6 Conclusion

In this paper, we construct a monthly presidential economic approval rating (PEAR) index from 1981 to 2019, and show that, in the cross-section, stocks with high PEAR beta significantly underperform those with low PEAR beta by 0.9% per month in the future, on a risk adjusted basis. The low-PEAR-beta premium persists up to one year and remains significant in a number of robustness tests. Contrary to the sentiment-induced overpricing, this premium does not come primarily from the short leg during high sentiment period. Since the PEAR index is negatively correlated with measures of aggregate risk aversion, a standard risk model would predict the low PEAR-beta stocks to earn lower (not higher) expected returns. The low-PEAR-beta premium remains a puzzle.

A number of topics are of interest for future research. First, developing an economic explanation with equilibrium theories is apparently desirable. Second, extending our results to other markets or asset classes will be worthwhile. Finally, given the data availability, we examine the low-PEAR-beta premium over the past four decades around. We look forward to research for extending the PEAR index to a longer period, perhaps with the help of sophisticated machine learning techniques.

Table 3.1 Summary statistics of PEAR and other related indexes

This table reports level and change correlations between the presidential economic approval rating (PEAR) and other sentiment and politics related indexes, consisting of (orthogonalized) investor sentiment (Baker and Wurgler, 2006), University of Michigan consumer sentiment, presidential job approval rating (Liu and Shaliastovich, 2021), aggregate political risk and sentiment (Hassan et al., 2019), and political uncertainty [measured by economic policy uncertainty in Baker, Bloom, and Davis (2016)]. AR(1) and AR(12) refer to the first- and 12th-order autocorrelations. All the time series are at the monthly frequency and over the 1981:04–2019:12 period, except for quarterly aggregate political risk and sentiment being 2002Q1–2019Q4, and political uncertainty being 1985:01–2019:12.

Panel A: Summary statistics

	Mean	Median	Min	Max	Volatility	AR(1)	AR(12)
PEAR	46.90	46.00	17.67	77.00	11.60	0.93	0.59
Investor sentiment	0.29	0.16	−0.89	3.20	0.62	0.97	0.33
Consumer sentiment	87.69	90.90	55.30	112.00	11.92	0.95	0.66
Presidential approval	51.65	50.00	27.00	89.80	11.72	0.93	0.48
Political risk	7.39	7.17	5.74	10.09	1.00	0.73	−0.02
Political sentiment	4.35	4.24	1.75	6.39	1.00	0.92	0.59
Political uncertainty	127.45	100.84	53.95	406.30	78.55	0.96	0.72

Panel B: Correlations

	PEAR	Investor sentiment	Consumer sentiment	Presidential approval	Political risk	Political sentiment	Political uncertainty
Correlation between levels							
PEAR sentiment	1.00						
Investor sentiment	0.22***	1.00					
Consumer sentiment	0.63***	0.26***	1.00				
Presidential approval	0.64***	0.12***	0.26***	1.00			
Political risk	0.08	−0.56***	−0.54***	0.09	1.00		
Political sentiment	0.22*	0.24*	0.59***	−0.35***	−0.23**	1.00	
Political uncertainty	−0.23***	−0.32***	−0.60***	−0.17***	0.56***	−0.17	1.00
Correlation between changes							
PEAR	1.00						
Investor sentiment	−0.09**	1.00					
Consumer sentiment	0.14***	−0.04	1.00				
Presidential approval	0.23***	−0.05	0.13***	1.00			
Political risk	0.20*	−0.09	0.16	0.22*	1.00		
Political sentiment	−0.13	−0.04	−0.14	−0.17	−0.40***	1.00	
Political uncertainty	0.03	0.03	−0.10**	0.15***	−0.02	−0.03	1.00

Table 3.2 Autocorrelations and pairwise correlations

This table reports autocorrelations and pairwise correlations of firm-specific characteristics, including PEAR-beta (β_{PEAR}), market-beta (β_{MKT}), economic uncertainty-beta (β_{UNC} , Bali, Brown, and Tang, 2017), sentiment beta (β_{BW}, Chen, Han, and Pan, 2020), political alignment index (PAI, Kim, Pantzalis, and Park, 2012), political sensitivity (PS, Addoum and Kumar, 2016), political connectedness (PC, Cooper, Gulen, and Ovtchinnikov, 2010), government spending exposure (GSE, Belo, Gala, and Li, 2013), log firm size (SIZE), log book-to-market ratio (BM), momentum, short-reversal (STR), idiosyncratic volatility (IVOL), illiquidity (ILLIQ, Amihud, 2002), failure probability (Distress, Campbell, Hilscher, and Szilagyi, 2008), and mispricing score (MISP, Stambaugh, Yu, and Yuan, 2015b). AR(1) and AR(12) refer to the first- and 12th-order autocorrelations. The sample period is 1983:06–2019:12.

	β_{PEAR}	β_{MKT}	β_{UNC}	β_{BW}	PAI	PS	PC	GSE	SIZE	BM	MOM	STR	IVOL	ILLIQ	Distress	MISP
Panel A: Autocorrelation																
AR(1)	0.82	0.84	0.79	0.80	0.92	0.85	0.97	0.68	0.83	0.90	0.76	−0.05	0.26	0.48	0.31	0.78
AR(12)	0.34	0.39	0.26	0.31	0.40	0.36	0.73	0.45	0.36	0.24	−0.17	−0.01	0.11	0.14	0.12	−0.02
Panel B: Pairwise correlation: standard (rank) correlation above (below) the diagonal																
β_{PEAR}		0.17	0.09	0.09	−0.02	−0.04	−0.04	0.04	−0.06	−0.00	−0.02	−0.00	0.08	0.01	0.07	0.05
β_{MKT}	0.16		0.04	0.20	−0.02	−0.08	−0.11	0.06	−0.08	−0.08	0.01	0.00	0.18	−0.02	0.14	0.21
β_{UNC}	0.08	0.01		0.02	0.00	−0.02	−0.01	−0.02	0.00	−0.05	−0.03	0.00	0.02	−0.00	0.02	0.02
β_{BW}	0.07	0.19	0.03		−0.02	−0.07	−0.08	−0.01	−0.10	−0.05	0.03	0.01	0.11	0.01	0.09	0.07
PAI	−0.02	−0.00	0.00	−0.02		0.06	−0.01	−0.01	−0.01	−0.01	0.01	0.00	0.01	−0.00	0.00	0.00
PS	−0.04	−0.07	−0.01	−0.08	0.06		−0.01	−0.05	0.03	0.01	0.05	0.02	−0.03	−0.01	−0.03	−0.06
PC	−0.06	−0.14	−0.03	−0.13	−0.01	−0.00		0.12	0.44	−0.07	−0.00	−0.00	−0.17	−0.06	−0.17	−0.16
GSE	0.06	0.08	−0.02	0.01	0.00	−0.09	0.02		−0.00	0.02	−0.00	−0.00	0.04	0.00	0.03	0.04
SIZE	−0.06	−0.07	0.02	−0.13	−0.01	0.03	0.46	−0.01		−0.28	0.13	0.03	−0.43	−0.23	−0.33	−0.20
BM	0.02	−0.06	−0.06	−0.03	−0.02	0.01	−0.08	−0.00	−0.31		0.00	0.02	0.02	0.10	0.01	−0.05
MOM	−0.04	−0.06	−0.06	−0.01	−0.00	0.07	0.05	−0.01	0.23	0.00		−0.01	−0.10	−0.07	−0.10	−0.23
STR	−0.01	−0.02	−0.01	−0.01	0.00	0.02	0.02	−0.01	0.09	0.01	0.02		0.21	0.01	0.10	−0.04
IVOL	0.10	0.26	0.01	0.15	0.02	−0.04	−0.29	0.06	−0.52	0.04	−0.21	−0.00		0.25	0.67	0.24
ILLIQ	0.05	0.05	−0.03	0.11	0.01	−0.03	−0.45	0.00	−0.92	0.32	−0.23	−0.07	0.52		0.19	0.02
Distress	0.10	0.27	0.01	0.15	0.02	−0.04	−0.30	0.05	−0.55	0.07	−0.27	−0.07	0.92	0.54		0.24
MISP	0.05	0.21	0.02	0.07	0.01	−0.04	−0.17	0.01	−0.18	−0.02	−0.34	−0.05	0.28	0.15	0.39	

Table 3.3 Average returns and alphas of PEAR-beta portfolios

This table reports monthly average excess returns and alphas (in %) of decile portfolios sorted by PEAR beta (β_{PEAR}), where P1 (P10) refers to the portfolio with low (high) β_{PEAR}, and L-H refers to the strategy that buys P1 and sells P10. All portfolios are value-weighted and rebalanced at a monthly frequency. Factor models include Fama and French (2015) five-factor model (FF5), Hou, Xue, and Zhang (2015) q-factor model (HXZ), Stambaugh and Yuan (2017) mispricing-factor model (SY), Daniel, Hirshleifer, and Sun (2020) behavioral-factor model (DHS), Fama and French (2015) five-factor model plus Frazzini and Pedersen (2014) betting against beta factor (FF5+BAB), and Daniel et al. (1997) characteristics-based model (DGTW). Reported in parentheses are t-values. Industry demeaned β_{PEAR} is based on the Fama-French 48 industries. The sample period is 1983:06–2019:12.

	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	L-H
Panel A: Sort on β_{PEAR}											
β_{PEAR}	−1.30	−0.45	−0.17	0.03	0.20	0.38	0.57	0.83	1.21	2.32	−3.62
Excess	1.23	0.92	0.81	0.79	0.73	0.64	0.47	0.49	0.44	0.11	1.11
	(4.11)	(4.08)	(3.75)	(3.89)	(3.34)	(2.89)	(1.91)	(1.73)	(1.37)	(0.29)	(4.13)
α_{FF5}	0.66	0.15	−0.02	0.06	0.04	−0.06	−0.11	−0.15	−0.11	−0.24	0.90
	(4.06)	(1.30)	(−0.20)	(0.88)	(0.55)	(−0.78)	(−1.11)	(−1.16)	(−0.77)	(−1.59)	(3.77)
α_{HXZ}	0.69	0.23	0.04	0.07	0.08	−0.01	0.00	−0.01	−0.01	−0.10	0.78
	(4.20)	(1.91)	(0.43)	(0.93)	(1.04)	(−0.09)	(0.00)	(−0.05)	(−0.06)	(−0.57)	(3.14)
α_{SY}	0.53	0.19	−0.01	0.03	0.08	−0.05	−0.08	−0.03	−0.15	−0.26	0.79
	(3.09)	(1.54)	(−0.11)	(0.42)	(1.00)	(−0.54)	(−0.76)	(−0.25)	(−0.94)	(−1.30)	(2.94)
α_{DHS}	0.72	0.14	0.06	0.02	−0.01	0.02	0.01	0.03	0.04	0.07	0.66
	(3.87)	(1.10)	(0.57)	(0.33)	(−0.06)	(0.24)	(0.07)	(0.22)	(0.26)	(0.34)	(2.48)
$\alpha_{\text{FF5+BAB}}$	0.73	0.18	0.00	0.03	0.03	−0.08	−0.08	−0.09	−0.06	−0.17	0.90
	(4.51)	(1.52)	(0.01)	(0.47)	(0.33)	(−0.92)	(−0.80)	(−0.67)	(−0.44)	(−1.13)	(3.74)
DGTW	0.43	0.09	0.06	0.09	0.01	0.00	−0.13	−0.16	−0.08	−0.36	0.79
	(3.15)	(1.12)	(0.74)	(1.56)	(0.11)	(0.07)	(−1.72)	(−1.75)	(−0.66)	(−2.14)	(3.97)
Panel B: Sort on industry demeaned β_{PEAR}											
β_{PEAR}	−1.68	−0.79	−0.49	−0.29	−0.12	0.04	0.21	0.43	0.76	1.78	−3.46
Excess	0.98	0.82	0.66	0.79	0.80	0.65	0.71	0.52	0.55	0.13	0.85
	(2.85)	(3.37)	(3.09)	(3.81)	(3.91)	(3.00)	(3.15)	(2.15)	(1.87)	(0.36)	(4.15)
α_{FF5}	0.54	0.18	−0.03	0.12	0.06	−0.14	0.04	−0.22	−0.09	−0.26	0.79
	(3.52)	(1.78)	(−0.39)	(1.77)	(0.95)	(−1.87)	(0.40)	(−2.24)	(−0.70)	(−1.85)	(3.95)
α_{HXZ}	0.59	0.26	0.04	0.15	0.08	−0.07	0.10	−0.09	−0.02	−0.14	0.73
	(3.73)	(2.42)	(0.40)	(2.14)	(1.28)	(−0.92)	(1.09)	(−0.93)	(−0.18)	(−0.87)	(3.54)
α_{SY}	0.41	0.28	−0.08	0.08	0.09	−0.05	0.06	−0.16	−0.03	−0.28	0.68
	(2.41)	(2.57)	(−0.93)	(1.21)	(1.27)	(−0.65)	(0.70)	(−1.61)	(−0.25)	(−1.51)	(3.12)
α_{DHS}	0.56	0.27	−0.08	0.08	0.11	−0.06	0.11	−0.07	0.05	0.03	0.53
	(3.34)	(2.42)	(−0.90)	(1.07)	(1.64)	(−0.81)	(1.16)	(−0.71)	(0.41)	(0.18)	(2.39)
$\alpha_{\text{FF5+BAB}}$	0.64	0.24	−0.01	0.14	0.05	−0.17	0.02	−0.19	−0.11	−0.21	0.85
	(4.29)	(2.32)	(−0.07)	(2.08)	(0.81)	(−2.26)	(0.21)	(−2.00)	(−0.86)	(−1.49)	(4.20)
DGTW	0.22	0.10	−0.02	0.03	0.07	−0.00	0.10	−0.16	−0.13	−0.35	0.57
	(1.44)	(1.24)	(−0.26)	(0.60)	(1.31)	(−0.03)	(1.42)	(−2.17)	(−1.29)	(−2.21)	(3.04)

Table 3.4 Subperiod/subsample analyses of PEAR-beta portfolios

This table reports the monthly average excess returns and FF5 alphas of PEAR-beta (β_{PEAR}) portfolio in different subsamples. Panel A splits the sample into Democratic and Republican presidency periods. Panel B considers president transition and non-transition periods, where transition periods are defined as six months surrounding the January of new president inauguration. Panel C splits the sample into NBER-dated recessions and expansions. Panel D splits the sample based on the idiosyncratic volatility (IVOL) (Ang et al., 2006). Panel E splits the sample into two subsamples based on illiquidity (Amihud, 2002). Panel F splits stocks into two size groups based on the median NYSE breakpoints. Panel G considers alternative β_{PEAR} estimations: estimating β_{PEAR} by including the MKT factor or using a 4-year or 8-year rolling window. Panel H considers alternative PEAR indexes, such as using the innovation of the AR(1) process of PEAR, the president job approval rating, and the index based on the polls from top 6 polling agents [the missing values are filled by using the dyad ratios algorithm of Stimson (1999)]. All portfolios are value-weighted and rebalanced at a monthly frequency. Reported in parentheses are t-values. The sample period is 1983:06–2019:12.

	Excess return	FF5 alpha	#(obs.)		Excess return	FF5 alpha	#(obs.)
Panel A: Democratic vs. Republican presidents				Panel B: Transition vs. non-transition periods			
Democratic	1.25	1.36	192	Transition	2.25	1.32	65
	(3.06)	(3.86)			(3.22)	(2.17)	
Republican	1.01	0.55	247	Non-transition	0.91	0.83	374
	(2.81)	(1.81)			(3.15)	(3.27)	
Difference	−0.24	−0.81		Difference	−1.33	−0.48	
	(−0.44)	(−1.67)			(−1.76)	(−0.75)	
Panel C: Recessions vs. expansions				Panel D: Low vs. high IVOL firms			
Recession	2.86	1.63	34	Low IVOL	1.00	0.96	
	(2.00)	(1.45)			(3.89)	(3.81)	
Expansion	0.95	0.83	405	High IVOL	1.05	0.69	
	(3.64)	(3.24)			(3.11)	(2.17)	
Difference	−1.91	−0.80		Difference	0.05	−0.27	
	(−1.31)	(−0.70)			(0.15)	(−0.81)	
Panel E: Liquid vs. illiquid firms				Panel F: Small vs. big firms			
Liquid	1.14	0.95		Small	0.71	0.40	
	(4.08)	(3.80)			(3.16)	(2.02)	
Illiquid	0.44	0.15		Big	1.18	0.98	
	(2.12)	(0.81)			(3.57)	(3.20)	
Difference	−0.70	−0.80		Difference	0.47	0.58	
	(−2.81)	(−3.07)			(1.79)	(2.14)	
Panel G: Alternative β_{PEAR} estimation				Panel H: Alternative PEAR			
Including MKT	0.92	0.93		Innovation of PEAR AR(1)	0.78	0.68	
	(4.00)	(4.07)			(2.81)	(2.85)	
4-year rolling	0.90	0.58		Presidential approval rating	0.65	0.42	
	(3.19)	(2.39)			(2.76)	(1.73)	
8-year rolling	0.98	0.88		Top 6 agents	1.10	1.11	
	(3.71)	(3.65)			(3.82)	(3.88)	

Table 3.5 International evidence

This table reports monthly average excess returns and FF5 alphas (in %) of decile portfolios based on PEAR beta in other G7 countries. Stock return and market capitalization information are from Datastream. All returns and market capitalizations are based on local currencies, risk-free rate for each country is the 90-day interbank rate, and the international Fama-French five-factor data are from Schmidt et al. (2019). P1 (P10) refers to the portfolio with low (high) PEAR beta, and L-H refers to the strategy that buys P1 and sells P10. All portfolios are value-weighted and rebalanced at a monthly frequency. Reported in parentheses are t-values. The sample period is 1983:06–2019:12 for excess returns, and 1991:07–2019:12 for FF5 alphas (1990:07–2019:12 for Japan).

	Excess return			FF5 alpha		
	P1	P10	L-H	P1	P10	L-H
Canada	0.55	−0.71	1.26	0.08	−1.15	1.23
	(1.18)	(−1.38)	(2.41)	(0.18)	(−2.25)	(1.97)
France	0.73	0.61	0.12	0.07	−0.13	0.20
	(2.58)	(1.54)	(0.38)	(0.29)	(−0.37)	(0.50)
Germany	0.41	−0.09	0.50	0.39	−0.81	1.20
	(1.43)	(−0.26)	(1.52)	(1.59)	(−2.84)	(2.91)
Italy	0.60	0.22	0.38	0.04	−0.20	0.23
	(1.77)	(0.55)	(1.07)	(0.13)	(−0.59)	(0.53)
Japan	0.43	−0.14	0.57	0.27	−0.57	0.83
	(1.36)	(−0.43)	(2.04)	(1.86)	(−2.81)	(2.95)
UK	1.23	0.46	0.77	0.33	−0.64	0.97
	(4.55)	(1.34)	(2.44)	(1.54)	(−2.44)	(2.61)

Table 3.6 Fama-Macbeth regressions

This table reports the results of Fama-MacBeth regressions of one-month-ahead stock excess returns on PEAR beta (β_{PEAR}), controlling for other firm-specific characteristics, which include log firm size (SIZE), log book-to-market ratio (BM), price momentum (MOM), short-term reversal (STR), idiosyncratic volatility (IVOL), illiquidity (ILLIQ, Amihud, 2002), failure probability (Distress, Campbell, Hilscher, and Szilagyi, 2008), β_{MKT}, β_{UNC} (Bali, Brown, and Tang, 2017), β_{BW} (Chen, Han, and Pan, 2020), political alignment index (PAI, Kim, Pantzalis, and Park, 2012), political sensitivity (PS, Addoum and Kumar, 2016), government spending exposure (GSE, Belo, Gala, and Li, 2013), and political connectedness (PC, Cooper, Gulen, and Ovtchinnikov, 2010). In Column 6, we include 48 industry dummies classified following Fama and French (1997). All independent variables except for industry dummies are winsorized at the 1st and 99th percentiles, and then normalized to zero mean and standard deviation of one. Intercepts are included in all the regressions but unreported for brevity. Newey-West t-values are reported in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively. The sample period is 1983:06–2019:12.

	DepVar.: One-month-ahead excess returns (%)					
	(1)	(2)	(3)	(4)	(5)	(6)
β_{PEAR}	-0.17^{***}	-0.12^{***}	-0.11^{***}	-0.11^{**}	-0.09^{**}	-0.09^{***}
	(-2.73)	(-2.65)	(-3.38)	(-2.41)	(-2.53)	(-2.67)
β_{MKT}			0.09		0.10	0.10
			(1.07)		(1.21)	(1.29)
β_{UNC}			-0.04		-0.04	-0.06^{*}
			(-1.24)		(-1.22)	(-1.92)
β_{BW}			-0.08^{*}		-0.06	-0.06
			(-1.78)		(-1.48)	(-1.42)
PAI				0.07^{**}	0.06^{**}	0.04^{*}
				(2.14)	(2.14)	(1.76)
PS				0.17^{***}	0.16^{***}	
				(3.23)	(3.29)	
PC				-0.15^{**}	-0.16^{***}	-0.15^{***}
				(-2.32)	(-2.66)	(-2.62)
GSE				0.01	0.01	0.04
				(0.15)	(0.17)	(1.23)
SIZE		-0.18^{***}	-0.21^{***}	-0.17^{**}	-0.20^{***}	-0.18^{***}
		(-2.82)	(-3.13)	(-2.49)	(-2.70)	(-2.68)
BM		0.21^{***}	0.21^{***}	0.24^{***}	0.22^{***}	0.27^{***}
		(2.95)	(3.07)	(3.51)	(3.41)	(5.23)
MOM		0.22^{***}	0.22^{***}	0.19^{**}	0.19^{***}	0.17^{***}
		(2.68)	(3.16)	(2.36)	(2.88)	(2.79)
STR		-0.43^{***}	-0.49^{***}	-0.54^{***}	-0.58^{***}	-0.64^{***}
		(-6.80)	(-7.20)	(-7.47)	(-7.63)	(-8.48)
IVOL		0.39^{***}	0.42^{***}	0.28^{**}	0.30^{***}	0.26^{**}
		(3.51)	(3.90)	(2.36)	(2.61)	(2.47)
ILLIQ		0.05	0.06	0.04	0.05	0.06
		(1.02)	(1.56)	(0.83)	(1.33)	(1.48)
Distress		-0.83^{***}	-0.88^{***}	-0.57^{***}	-0.60^{***}	-0.57^{***}
		(-5.65)	(-6.12)	(-3.55)	(-4.07)	(-4.06)
Industry FEs	No	No	No	No	No	Yes
#(obs.)	1,382,258	1,233,512	1,209,591	726,462	726,414	746,160
Adj. R^2	0.005	0.039	0.045	0.049	0.055	0.070

Table 3.7 The relationship between PEAR and macro variables

Panel A reports the correlations between the change in PEAR and other macro variables, and Panel B reports the correlations of their betas, where raw corr refers to the correlation without transforming the variables, and rank corr refers to the rank correlation after transforming each variable into a rank one. Macro variables include industrial production growth (IPG), unexpected inflation (UI), change in expected inflation (DEI), term premium (TERM), default premium (DEF), total factor productivity growth (TFP), labor income growth (LIG), capital share growth (CSG, Lettau, Ludvigson, and Ma, 2019), consumption growth (CG), ultimate consumption growth (UCG, Parker and Julliard, 2005), change in consumption to wealth ratio (CAY), change in aggregate market volatility (VOL), change in VIX, variance risk premium (VRP), and growth in gross domestic product (GDP). The correlations of PEAR with TFP, CS, CG, UCG, CAY, and GDP are at the quarterly frequency. β_{CS}, β_{CG}, β_{UCG}, β_{CAY}, β_{TFP}, and β_{GDP} are estimated from regressions of quarterly excess returns on the current and lagged values of the variables in the past 10 years. β_{VIX} is estimated from regressions of excess stock returns on the excess market returns and the changes in VIX using daily data in a month. Other betas are estimated using the same specification of estimating β_{PEAR}. We flip the signs of β_{UI}, β_{DEI}, β_{TERM}, β_{DEF}, β_{VOL}, β_{VIX}, and β_{VRP} so that they are capturing the correct direction of risk (i.e., high beta implies high risk). The sample period is 1983:06–2019:12, except for β_{VIX} and β_{VRP} being 1990:01–2019:12.

Panel A: Correlations between the change in PEAR and other macro variables

	IPG	UI	DEI	TERM	DEF	TFP	LIG	CSG	CG	UCG	CAY	VOL	VIX	VRP	GDP
Raw corr	0.01	−0.06	−0.05	0.01	0.03	0.10	−0.07	0.14	0.01	−0.06	0.06	0.06	0.03	0.10	0.08

Panel B: Correlations between β_{PEAR} and macro betas

	β_{IPG}	β_{UI}	β_{DEI}	β_{TERM}	β_{DEF}	β_{TFP}	β_{LIG}	β_{CSG}	β_{CG}	β_{UCG}	β_{CAY}	β_{VOL}	β_{VIX}	β_{VRP}	β_{GDP}
Raw corr	−0.06	0.08	0.07	0.01	−0.02	0.04	−0.07	0.07	0.02	−0.05	−0.06	−0.00	−0.01	−0.15	−0.02
Rank corr	−0.05	0.08	0.07	0.00	−0.03	0.05	−0.05	0.08	0.01	−0.05	−0.05	−0.01	−0.01	−0.10	−0.02

Table 3.8 Low-PEAR-beta premiums in high and low sentiment periods

This table reports the monthly average excess returns and FF5 alphas (in %) of PEAR-beta (β_{PEAR}) decile portfolios in high and low sentiment periods. We consider four indexes as the proxy for investor sentiment, including Baker and Wurgler (2006) sentiment index, Michigan consumer sentiment index, AAII bull-bear index, and PEAR itself. A month is defined as a high sentiment month if the sentiment index in the previous month is above its median. P1 and P10 refer to the low and high β_{PEAR} portfolios, and L-H refers to their difference. All portfolios are value-weighted and rebalanced at a month frequency. Reported in parentheses are t-values. The sample period is 1983:06–2019:12.

	Low sentiment	High sentiment	Difference
Panel A: Baker and Wurgler (2006) sentiment index			
P1	0.91	0.54	−0.37
	(3.73)	(2.17)	(−1.09)
P10	−0.31	−0.21	0.10
	(−1.26)	(−1.08)	(0.34)
L-H	1.22	0.75	−0.47
	(3.46)	(2.18)	(−1.00)
Panel B: Michigan consumer sentiment index			
P1	0.42	0.83	0.41
	(1.89)	(3.29)	(1.27)
P10	−0.02	−0.41	−0.38
	(−0.09)	(−2.15)	(−1.33)
L-H	0.45	1.24	0.79
	(1.24)	(3.74)	(1.70)
Panel C: AAII bull-bear index			
P1	0.65	0.64	−0.02
	(2.84)	(2.35)	(−0.05)
P10	−0.32	−0.26	0.06
	(−1.49)	(−1.08)	(0.17)
L-H	0.97	0.90	−0.07
	(2.94)	(2.31)	(−0.15)
Panel D: PEAR			
P1	0.23	1.13	0.91
	(1.22)	(3.85)	(2.76)
P10	−0.20	−0.28	−0.08
	(−1.06)	(−1.19)	(−0.28)
L-H	0.43	1.42	0.99
	(1.48)	(3.49)	(2.05)

Table 3.9 PEAR beta and presidential alignment

This table reports the results of PEAR beta reflecting a new type of sentiment. In Panel A, at the end of each month, we split the past 60 months into two sub-samples, one coming from the current president party and the other from the opposite party (with a requirement of at least 12 observations), and then estimate a current party beta and an opposite part beta for each firm accordingly. In Panel B, we run Fama-MacBeth regressions of measures of analyst reaction and cumulative abnormal return (CAR, in %) around earnings announcement days (Edays) on the past PEAR beta (β_{PEAR}) as well as firm-specific characteristics (same as model 2 in Table 3.6). The measures of analyst reaction include analyst forecast errors (AFE_{t+12}, in%), revisions in long-term growth rate forecasts ($\Delta\text{LTG}_{t+12} = \text{LTG}_{t+12} - \text{LTG}_t$, in %), and revisions in analyst recommendations ($\Delta\text{Rec}_{t+12} = \text{Rec}_{t+12} - \text{Rec}_t$, in %). The CARs' results, adjusted by Daniel et al. (1997) characteristics-matched benchmark returns, are at the quarterly frequency and based on the quarter-end month PEAR betas. In Panel C, we sort stocks into two subgroups based on their abnormal Google search volume or abnormal trading volume, and explore the low-PEAR-beta premium within each subgroup. The constructions of the abnormal Google search volume and abnormal trading volume follow Da, Engelberg, and Gao (2011) and Gervais, Kaniel, and Mingelgrin (2001), respectively. Reported in Panels A and C are the FF5 alphas. The sample period is 1983:06–2019:12, expect for analyst recommendations being 1994:12–2019:12 and abnormal Google search volume being 2004:04–2019:12.

Panel A: Low-PEAR-beta premiums constructed by current and opposite party betas

	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	L-H
Current	0.77	0.14	0.04	0.06	−0.03	−0.07	−0.13	−0.11	−0.07	−0.31	1.08
	(4.67)	(1.12)	(0.39)	(0.72)	(−0.29)	(−0.72)	(−1.11)	(−0.79)	(−0.47)	(−2.16)	(4.72)
Opposite	0.05	−0.19	0.16	0.04	−0.07	−0.03	0.07	0.03	0.08	−0.31	0.35
	(0.21)	(−1.18)	(1.00)	(0.29)	(−0.50)	(−0.20)	(0.30)	(0.12)	(0.35)	(−1.09)	(0.89)

Panel B: Analyst reactions and CARs around Edays

	AFE_{t+12}	ΔLTG_{t+12}	ΔRec_{t+12}	CAR_{q+1}	CAR_{q+2}	CAR_{q+3}
β_{PEAR}	−0.90***	−0.12***	−0.18**	−0.06**	−0.05**	−0.04
	(−3.89)	(−2.73)	(−2.21)	(−2.24)	(−2.27)	(−1.38)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Adj. R^2	0.049	0.048	0.041	0.007	0.006	0.006

Panel C: Low-PEAR-beta premiums among low and high attention stocks

	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	L-H
					Abnormal Google search volume						
Low	0.52	0.50	−0.09	0.21	0.02	−0.08	−0.12	−0.18	−0.38	−0.70	1.22
	(2.04)	(2.53)	(−0.58)	(1.51)	(0.15)	(−0.68)	(−0.82)	(−0.84)	(−1.91)	(−3.10)	(3.68)
High	0.14	0.04	0.16	0.03	0.04	0.10	−0.12	−0.01	−0.31	−0.37	0.51
	(0.55)	(0.24)	(1.15)	(0.24)	(0.31)	(0.70)	(−0.82)	(−0.09)	(−1.55)	(−1.57)	(1.47)
					Abnormal trading volume						
Low	0.44	0.04	−0.01	0.05	−0.15	−0.13	−0.16	−0.35	−0.30	−0.58	1.02
	(2.36)	(0.31)	(−0.05)	(0.53)	(−1.41)	(−1.27)	(−1.31)	(−2.32)	(−1.83)	(−3.21)	(3.63)
High	0.60	0.23	0.08	0.13	0.25	0.02	0.10	−0.17	0.13	0.06	0.55
	(3.16)	(1.66)	(0.73)	(1.11)	(2.18)	(0.20)	(0.73)	(−1.09)	(0.76)	(0.31)	(2.05)

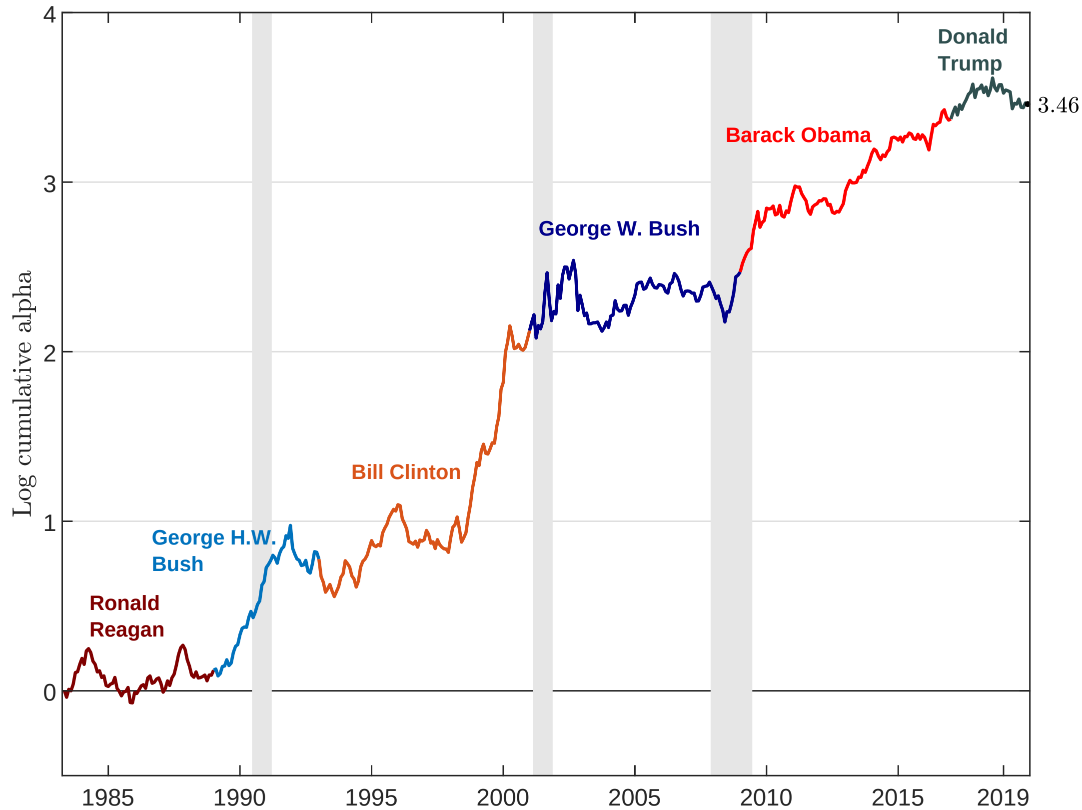


Fig. 3.2 Log cumulative alpha of the PEAR-beta spread portfolio

This figure plots the log cumulative FF5 alpha of the PEAR-beta spread portfolio. The sample period is 1983:06-2019:12.

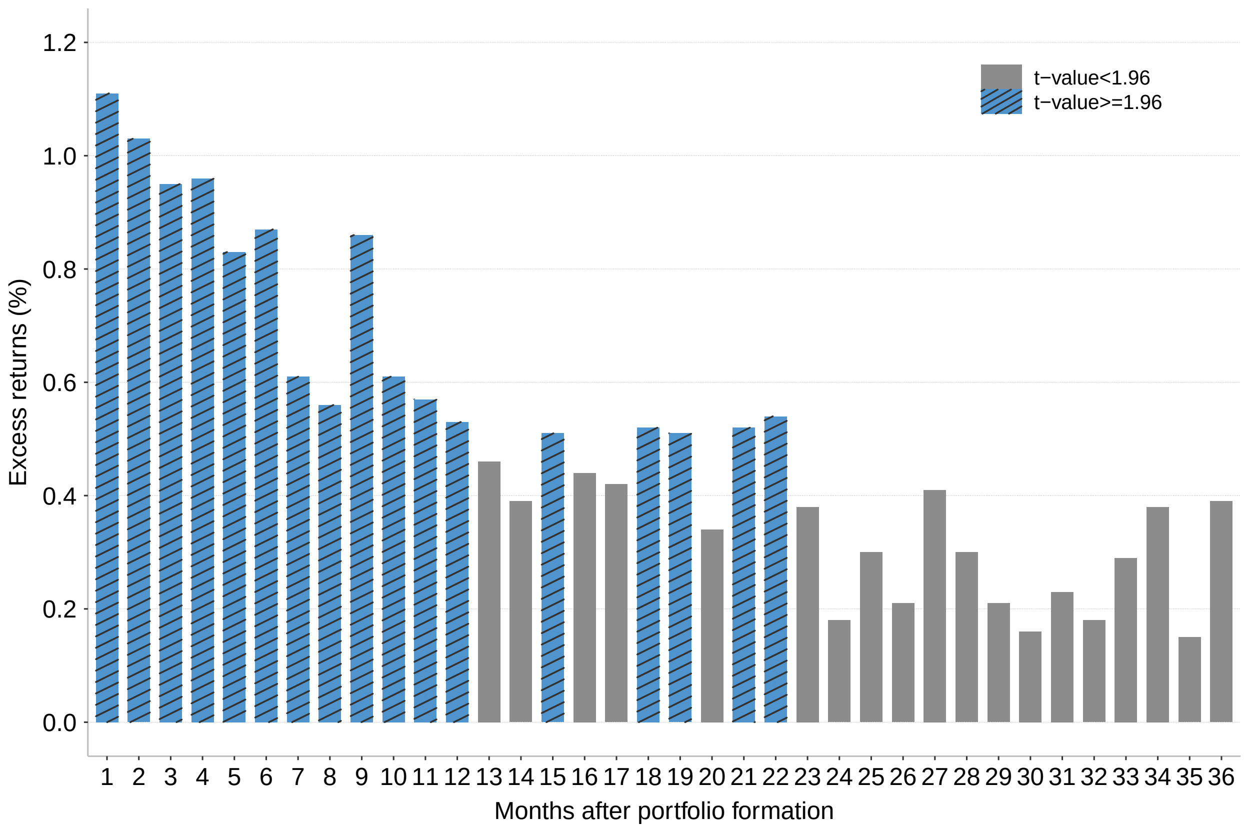


Fig. 3.3 PEAR-beta spread portfolio returns after formation

This figure plots the average excess returns of the PEAR-beta spread portfolio after formation. Grey (blue) indicates that the t-value is smaller (larger) than 1.96. The sample period is 1983:06–2019:12.

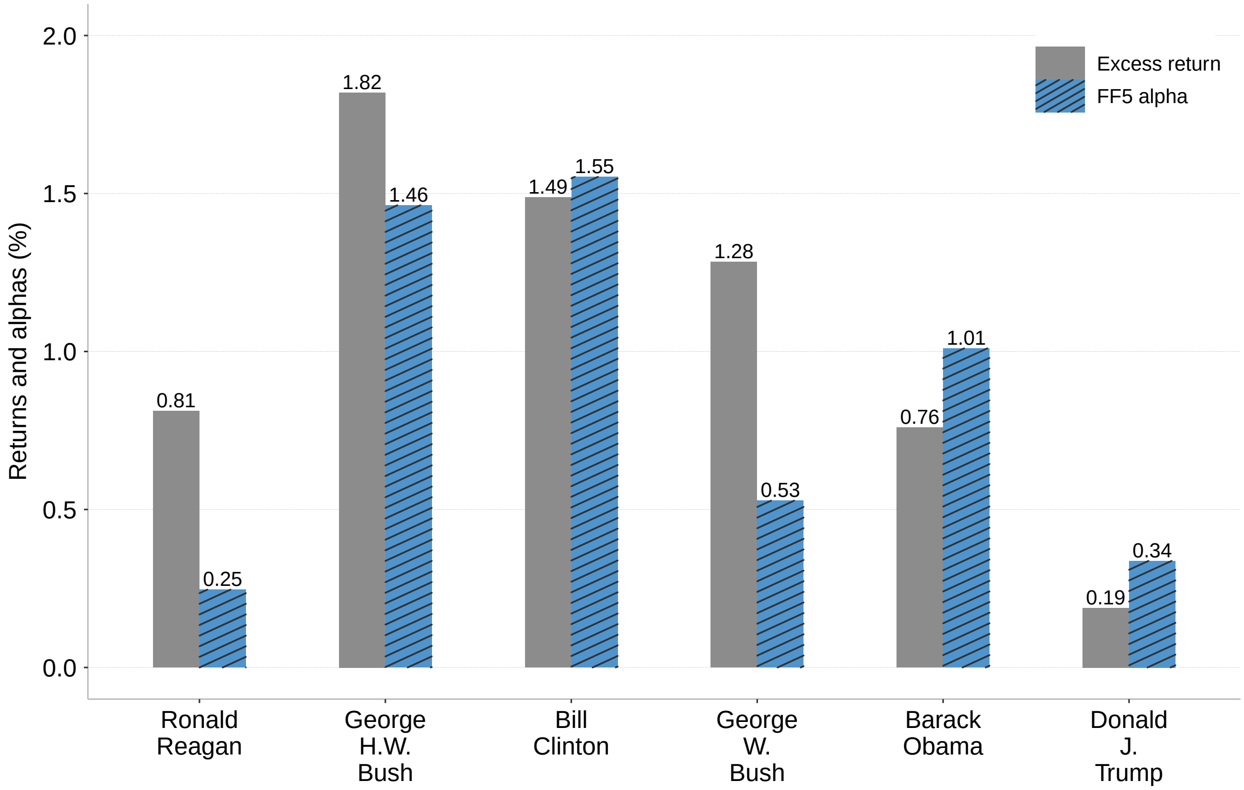


Fig. 3.4 Low-PEAR-beta premiums during different presidencies

This figure plots the monthly average excess returns of the PEAR-beta spread portfolio across different presidents. The sample period is 1983:06-2019:12.

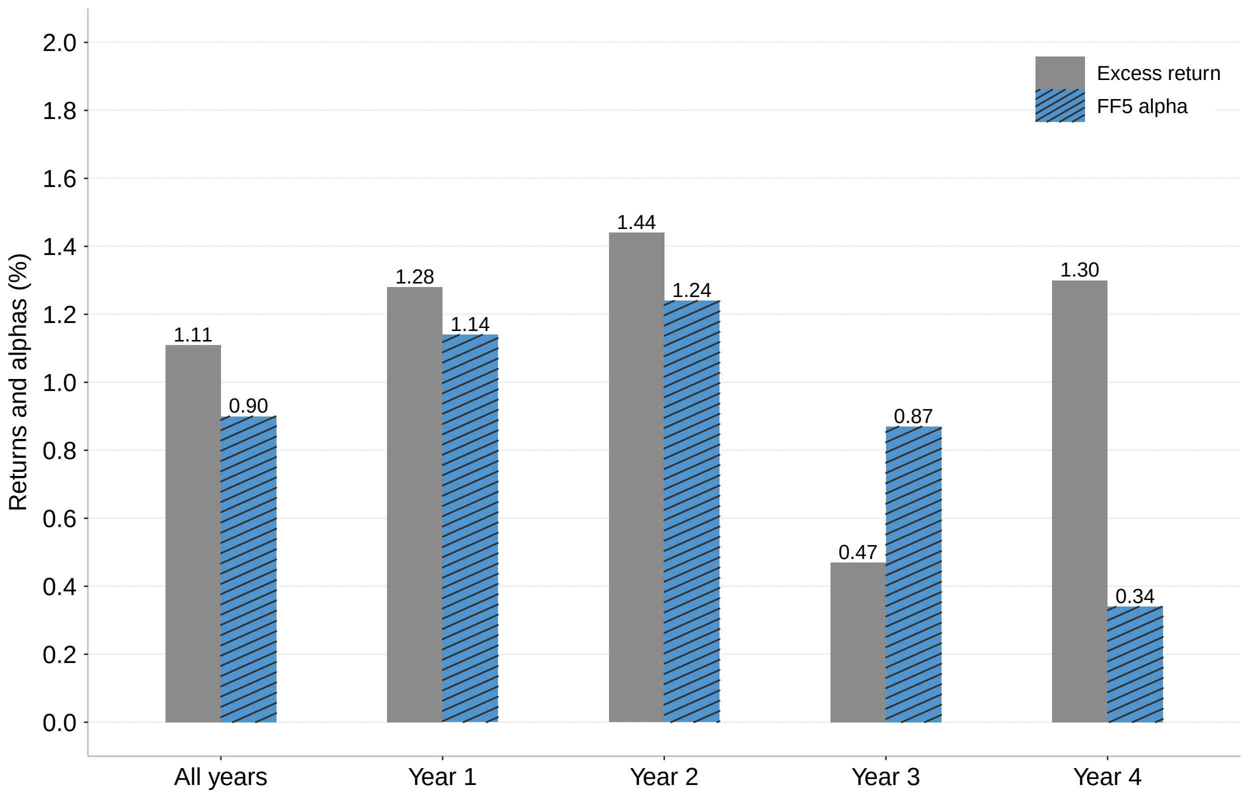


Fig. 3.5 Low-PEAR-beta premiums across years of the presidential term

This figure plots the monthly average excess returns and FF5 alphas of the PEAR-beta spread portfolio across years of the presidential term. The sample period is 1983:06-2019:12.

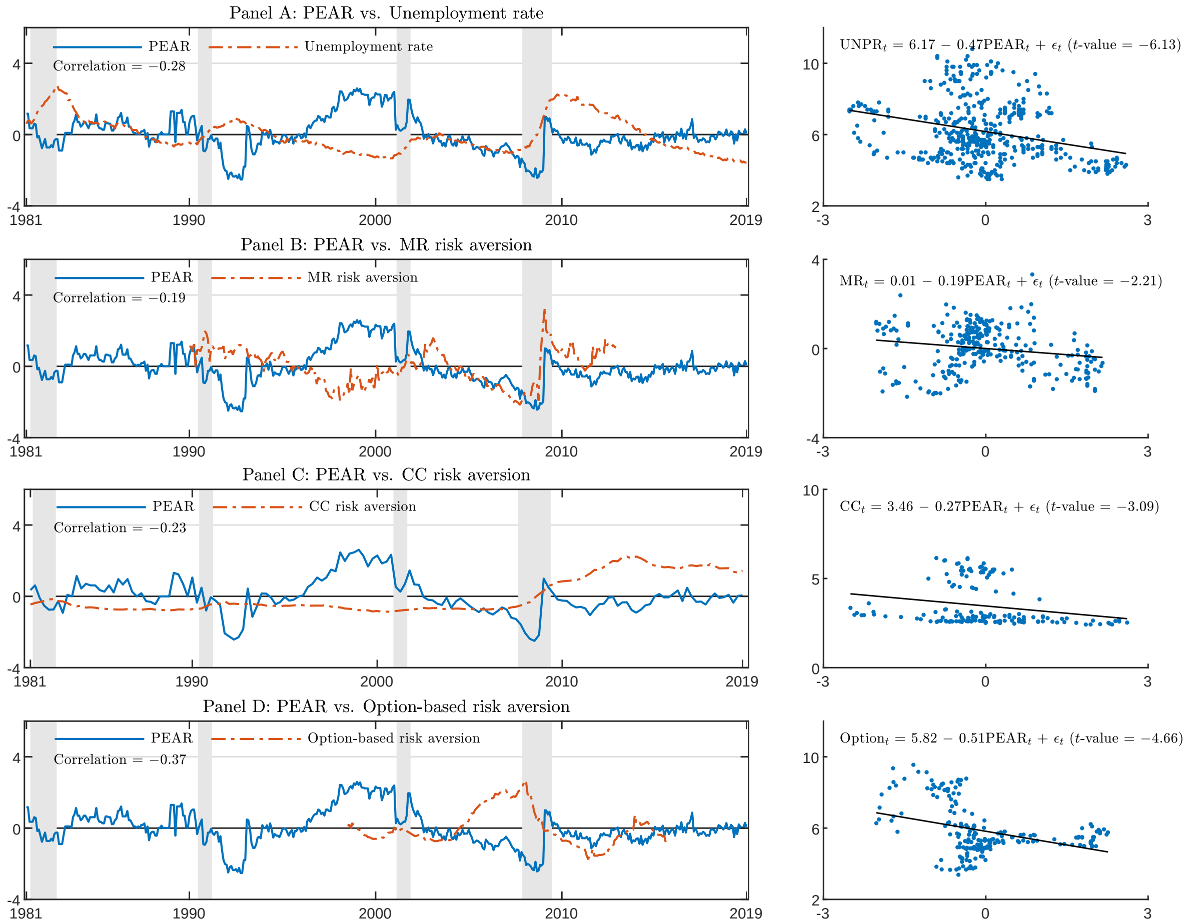


Fig. 3.6 PEAR vs. risk aversion

This figure plots the time series dynamics and scatter diagrams of PEAR and risk aversion. We consider four risk aversion measures, including unemployment rate (UNPR) (Pastor and Veronesi, 2020), aggregate risk aversion (MR, Miranda-Agrippino and Rey, 2020), negative of surplus consumption ratio (CC, Campbell and Cochrane, 1999), and option-based risk aversion (Option) (Faccini et al., 2019). The sample period is 1981:04–2019:12 for UNPR and CC, 1990:01–2012:12 for MR, and 1998:07–2015:08 for the option-based risk aversion.

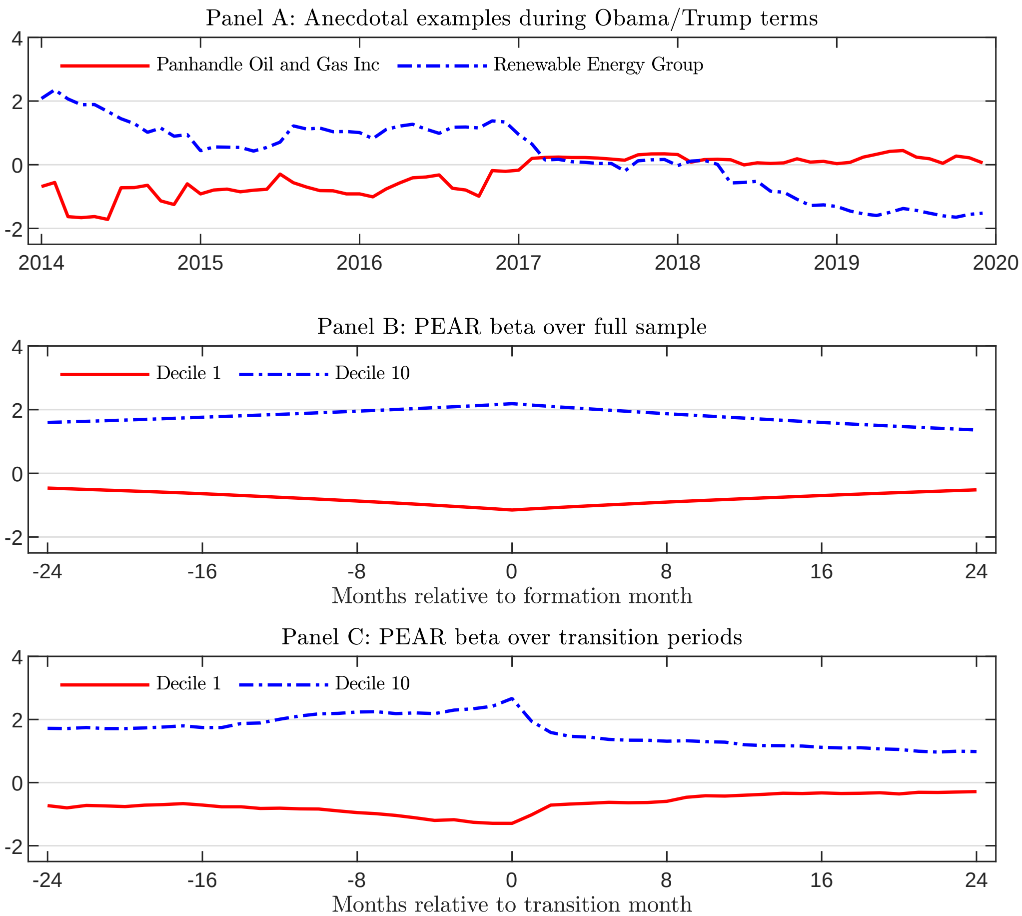


Fig. 3.7 Trend of PEAR beta

Panel A plots the PEAR betas of two anecdotal examples (Panhandle Oil and Gas Inc. vs. Renewable Energy Group) during Obama's and Trump's terms. Panel B plots the average values of PEAR beta in portfolio decile 1 and decile 10 around the formation month. Panel C plots the average values of PEAR beta in portfolio decile 1 and decile 10 around the presidential transition months. The sample period is 1983:06-2019:12.

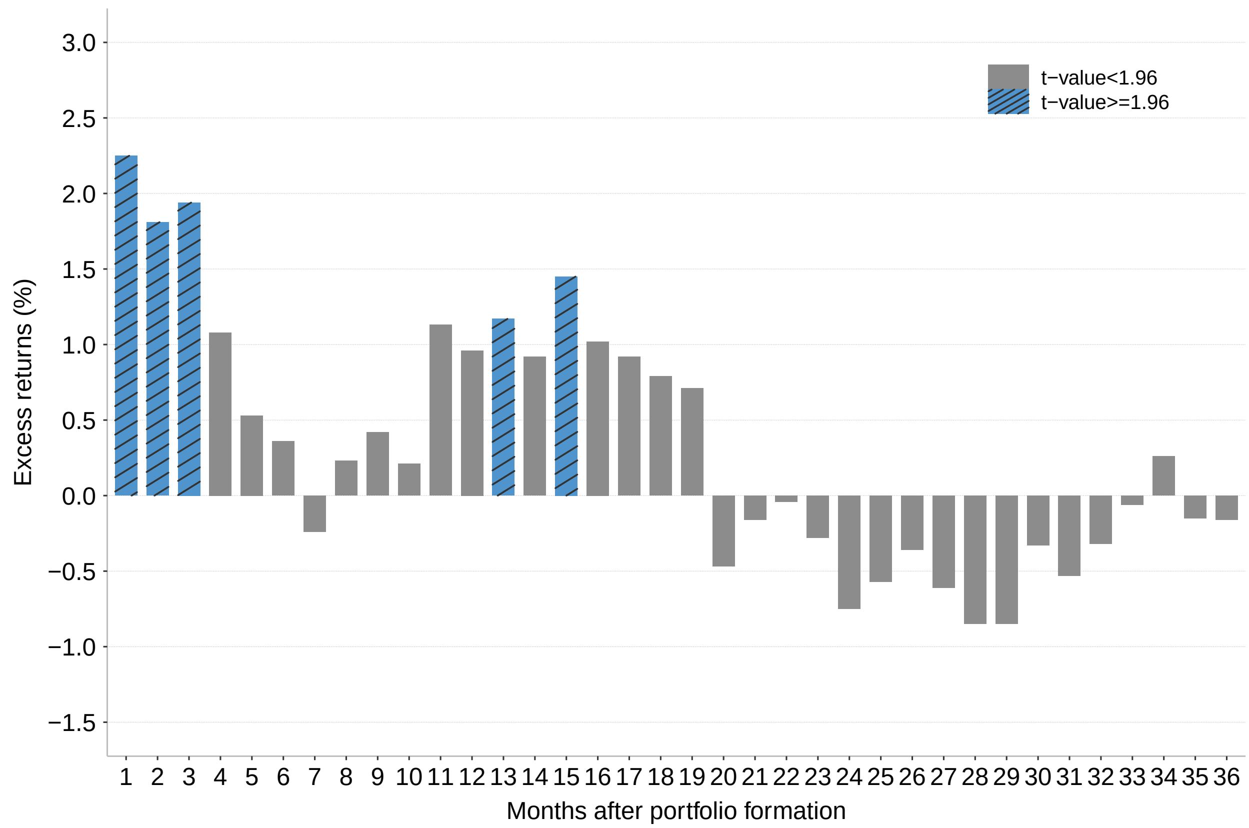


Fig. 3.8 PEAR-beta spread portfolio returns after formation: transition periods

This figure plots the average excess returns of the PEAR-beta spread portfolio after formation for the sample over a window of six months surrounding the January of new president inauguration. Grey (blue) indicates that the t-value is smaller (larger) than 1.96. The sample period is 1983:06–2019:12.

Appendix

Table A1 Data sources of PEAR

This table reports the summary statistics of the survey data used to construct our PEAR index. Reported are polling organization name, sample period, the total number of polling results, and the typical question wording of each polling organization. In total, there are 21 polling organizations with 1,713 polling results included in the sample.

Survey organization	Period	*N*	Typical question wording
ABC News	1981:09-2003:09	19	Do you approve or disapprove of the way Ronald Reagan/(George) Bush/(Bill) Clinton/(George W.) Bush is handling the nation's economy?
ABC News/Washington Post	1981:10-2019:09	206	Do you approve or disapprove of the way Reagan/(President George) Bush/(Bill) Clinton/(George W.) Bush/(Barack) Obama/(Donald) Trump is handling the economy?
American Research Group	2001:07-2019:12	210	Do you approve or disapprove of the way George W. Bush/Barack Obama/Donald Trump is handling the economy?
The Associated Press-NORC Center for Public Affairs Research	2002:11-2019:10	12	Overall, do you approve, disapprove, or neither approve nor disapprove of the way George W. Bush/Barack Obama/Donald Trump is handling the economy?
CBS News	1991:01-2019:05	150	How about the economy? Do you approve or disapprove of the way George Bush/Bill Clinton/George W. Bush/Barack Obama/Donald Trump is handling the economy?
CBS News/New York Times	1981:04-2016:07	196	Do you approve or disapprove of the way Ronald Reagan/Bill Clinton/George W. Bush/Barack Obama is handling the economy?
Consumer News and Business Channel (CNBC)	2009:12-2019:12	11	Do you generally approve or disapprove of the way Barack Obama/Donald Trump is handling the economy?
Cable News Network (CNN)	2006:05-2019:11	58	Do you approve or disapprove of the way George W. Bush/Barack Obama/Donald Trump is handling the economy?
FOX news (FOX)	2017:03-2019:09	21	Do you approve or disapprove of the way Donald Trump is handling... the economy?

Table A1 (continued)

Survey organization	Period	N	Typical question wording
Gallup Organization	1992:01-2019:11	169	Do you approve or disapprove of the way President Reagan/Bush/Bill Clinton/George W. Bush/Barack Obama/ Donald Trump is handling the economy?
Gesellschaft fr Konsumforschung (CfK)	2009:02-2018:10	43	Overall, do you approve, disapprove, or neither approve nor disapprove of the way Barack Obama/Donald Trump is handling... the economy?
Greenberg	2005:07-2011:05	11	Do you approve or disapprove of the way George (W.) Bush/Barack Obama is handling the economy?
Ipsos	2002:01-2008:07	139	And when it comes to handling the economy, do you approve or disapprove or have mixed feelings about the way George W. Bush is handling that issue?
Los Angeles Times	1981:04-2008:05	56	Do you approve or disapprove of the way Ronald Reagan/(Bill) Clinton/George W. Bush is handling the economy?
Marist College Institute for Public Opinion	2003:04-2019:09	32	Do you approve of disapprove of how President George (W.) Bush/Barack Obama/Donald Trump is handling the economy?
NBC News/Wall Street Journal	1988:07-2019:08	183	Do you generally approve or disapprove of the job Ronald Reagan/George Bush/Bill Clinton/Barack Obama/Donald Trump is doing in handling the economy?
Princeton Survey Research Associates	1994:10-2017:02	92	Do you approve or disapprove of the way Bill Clinton/George W. Bush/Barack Obama/Donald Trump is handling the economy?
Quinnipiac University Polling Institute	2002:02-2019:12	64	Do you approve or disapprove of the way George W. Bush/Barack Obama/Donald Trump is handling the economy?
The Tarrance Group	1994:01-2003:09	8	Do you approve or disapprove of the way President George W. Bush/Bill Clinton is handling the economy?
Time magazine	2004:04-2013:06	25	Do you approve or disapprove of the job President (George W.) Bush/(Barack) Obama is doing in each of these areas... handling the economy
Washington Post	1990:03-2010:03	8	Do you approve of the way (Bill) Clinton/(George W.) Bush/(Barack) Obama is handling... the economy?

Table A2 Variable definitions

This table describes the constructions of main variables used in this paper.

Variable	Description
Other betas	
CAPM beta (β_{MKT})	We estimate the market beta same as the PEAR beta using a 60-month rolling window, with the requirement of at least 24 months of data are available (Fama and French, 1992).
Sentiment beta (β_{BW})	We estimate the sentiment beta same as PEAR-beta using changes and lagged changes of the Baker and Wurgler (2006) sentiment index in a 60-month rolling window, with the requirement of at least 24 months of data are available (Chen, Han, and Pan, 2020).
UNC beta (β_{UNC})	We estimate the UNC beta using 60-month rolling regressions of excess stock returns on UNC index together with market, size, book-to-market, momentum, liquidity, investment, and profitability factors, with the requirement of at least 24 months of data are available (Bali, Brown, and Tang, 2017).
Political variables	
Political alignment index (PAI)	PAI is calculated as the degree of a state's governor, control of its legislature, and the bulk of its members in Congress aligned with the presidential party (Kim, Pantzalis, and Park, 2012).
Political sensitivity (PS)	PS is estimated using the 15-year monthly rolling regressions of Fama and French (1997) 48 industry value-weighted excess returns on market excess return and a Republican dummy (Addoum and Kumar, 2016).
Political connectedness (PC)	PC is defined as a dummy variable which equals to one if a firm makes a contribution to a PAC (regardless of party affiliation) in the last 5 years and zero otherwise (Cooper, Gulen, and Ovtchinnikov, 2010; Addoum and Kumar, 2016).
Government spending exposure (GSE)	GSE is calculated as the proportion of an industry's total output (3-digit SIC) being purchased by the government sector for final use (Belo, Gala, and Li, 2013).
Analyst variables	
Analyst earnings forecast revisions (AFE)	The difference between actual reported earnings and the consensus earnings forecast, scaled by the closing stock price in the previous month.
Revision in analyst recommendations (ΔRec)	The difference between the current consensus recommendation and its value over one previous month.
Revision in long-term growth rate forecasts (ΔLTG)	The difference between the current consensus long-term growth rate forecast and its value over one previous month.

Table A2 (continued)

Variable	Description
Other anomaly variables	
Size	The logarithm of the product of price per share and the number of shares outstanding (in millions of dollars).
Book-to-market ratio (BM)	The book value of shareholder equity plus deferred taxes and investment tax credit (if available) minus the book value of preferred stocks at the end of fiscal year $t-1$, scaled by the market value at the end of December of year $t-1$ (Fama and French, 1992).
Momentum (MOM)	The cumulative return of a stock over a 11-month window ending one month before the portfolio formation (Jegadeesh and Titman, 1993).
Short-term reversal (STR)	The return of a stock over the prior month (Jegadeesh, 1990).
Idiosyncratic volatility (IVOL)	The standard deviation of a stock's daily idiosyncratic returns relative to the Fama and French (1993) three-factor model over the prior month (Ang et al., 2006).
Illiquidity ratio (ILLIQ)	The ratio of the daily absolute stock return to the daily dollar trading volume averaged in the prior month (Amihud, 2002).
Failure probability (Distress)	Distress is defined as $-9.164-0.058*\text{PRICE}+0.075*\text{MB}-2.13*\text{CASHMTA}-0.045*\text{RSIZE}+1.41*\text{IdioRisk}-7.13*\text{EXRETAVG}+1.42*\text{TLMTA}-20.26*\text{NIMTAAVG}$, where all other variables are calculated following Campbell, Hilscher, and Szilagyi (2008).

References

Aarland, K., Davis, J. C., Henderson, J. V., Ono, Y., 2007. Spatial organization of firms: The decision to split production and administration. RAND Journal of Economics 38, 480–494.

Addoum, J. M., Kumar, A., 2016. Political sentiment and predictable returns. Review of Financial Studies 29, 3471–3518.

Ali, U., Hirshleifer, D., 2020. Shared analyst coverage: Unifying momentum spillover effects. Journal of Financial Economics 136, 649–675.

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

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

Antoniou, C., Doukas, J. A., Subrahmanyam, A., 2016. Investor sentiment, beta, and the cost of equity capital. Management Science 62, 347–367.

Asker, J., Farre-Mensa, J., Ljungqvist, A., 2014. Corporate investment and stock market listing: A puzzle? Review of Financial Studies 28, 342–390.

Bailey, W., Karolyi, G. A., Salva, C., 2006. The economic consequences of increased disclosure: Evidence from international cross-listings. Journal of Financial Economics 81, 175–213.

Bailey, W., Li, H., Mao, C. X., Zhong, R., 2003. Regulation fair disclosure and earnings information: Market, analyst, and corporate responses. Journal of Finance 58, 2487–2514.

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

Baker, S. R., Bloom, N., Davis, S. J., 2016. Measuring economic policy uncertainty. Quarterly Journal of Economics 131, 1593–1636.

Bali, T. G., Brown, S., Murray, S., Tang, Y., 2017. A lottery-demand-based explanation of the beta anomaly. Journal of Financial and Quantitative Analysis 52, 2369–2397.

Bali, T. G., Brown, S. J., Tang, Y., 2017. Is economic uncertainty priced in the cross-section of stock returns? Journal of Financial Economics 126, 471–489.

Bali, T. G., Cakici, N., Whitelaw, R. F., 2011. Maxing out: Stocks as lotteries and the cross-section of expected returns. Journal of Financial Economics 99, 427–446.

Bali, T. G., Hirshleifer, D. A., Peng, L., Tang, Y., 2018. Attention, social interaction, and investor attraction to lottery stocks. SSRN working paper .

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

Barber, B. M., Odean, T., 2007. All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors. Review of Financial Studies 21, 785–818.

Belo, F., Gala, V. D., Li, J., 2013. Government spending, political cycles, and the cross section of stock returns. Journal of Financial Economics 107, 305–324.

Bernard, V. L., Thomas, J. K., 1989. Post-earnings-announcement drift: delayed price response or risk premium? Journal of Accounting Research 27, 1–36.

Bernile, G., Kumar, A., Sulaeman, J., 2015. Home away from home: Geography of information and local investors. Review of Financial Studies 28, 2009–2049.

Breeden, D. T., 1979. An intertemporal asset pricing model with stochastic consumption and investment opportunities. Journal of Financial Economics 7, 265–296.

Brogaard, J., Dai, L., Ngo, P. T. H., Zhang, B., 2020. Global political uncertainty and asset prices. Review of Financial Studies 33, 1737–1780.

Burt, A., Hrdlicka, C., Harford, J., 2020. How much do directors influence firm value? Review of Financial Studies 33, 1818–1847.

Bushee, B. J., 2001. Do institutional investors prefer near-term earnings over long-run value? Contemporary Accounting Research 18, 207–246.

Campbell, J. Y., Cochrane, J. H., 1999. By force of habit: A consumption-based explanation of aggregate stock market behavior. Journal of Political Economy 107, 205–251.

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

Campbell, J. Y., Hilscher, J., Szilagyi, J., 2008. In search of distress risk. Journal of Finance 63, 2899–2939.

Carhart, M. M., 1997. On persistence in mutual fund performance. Journal of Finance 52, 57–82.

Chen, S., Miao, B., Shevlin, T., 2015. A new measure of disclosure quality: The level of disaggregation of accounting data in annual reports. Journal of Accounting Research 53, 1017–1054.

Chen, Y., Han, B., Pan, J., 2020. Sentiment trading and hedge fund returns. Journal of Finance, forthcoming.

Claessens, S., Djankov, S., Lang, L. H., 2000. The separation of ownership and control in east asian corporations. Journal of Financial Economics 58, 81–112.

Cohen, L., Frazzini, A., 2008. Economic links and predictable returns. Journal of Finance 63, 1977–2011.

Cohen, L., Lou, D., 2012. Complicated firms. Journal of Financial Economics 104, 383–400.

Collins, D. W., Hribar, P., 2000. Earnings-based and accrual-based market anomalies: one effect or two? Journal of Accounting and Economics 29, 101–123.

Cooper, M. J., Gulen, H., Ovtchinnikov, A. V., 2010. Corporate political contributions and stock returns. Journal of Finance 65, 687–724.

Cooper, M. J., Gulen, H., Schill, M. J., 2008. Asset growth and the cross-section of stock returns. Journal of Finance 63, 1609–1651.

Da, Z., Engelberg, J., Gao, P., 2011. In search of attention. Journal of Finance 66, 1461–1499.

Da, Z., Gurun, U. G., Warachka, M., 2014. Frog in the pan: Continuous information and momentum. Review of Financial Studies 27, 2171–2218.

Daniel, K., Grinblatt, M., Titman, S., Wermers, R., 1997. Measuring mutual fund performance with characteristic-based benchmarks. Journal of Finance 52, 1035–1058.

Daniel, K., Hirshleifer, D., Sun, L., 2020. Short-and long-horizon behavioral factors. Review of Financial Studies 33, 1673–1736.

Daniel, K., Moskowitz, T. J., 2016. Momentum crashes. Journal of Financial Economics 122, 221–247.

Daniel, K., Titman, S., 1997. Evidence on the characteristics of cross sectional variation in stock returns. Journal of Finance 52, 1–33.

De Boef, S., Kellstedt, P. M., 2004. The political (and economic) origins of consumer confidence. American Journal of Political Science 48, 633–649.

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

Dempsey, S. J., Harrison, D. M., Luchtenberg, K. F., Seiler, M. J., 2012. Financial opacity and firm performance: the readability of reit annual reports. Journal of Real Estate Finance and Economics 45, 450–470.

Dimson, E., 1979. Risk measurement when shares are subject to infrequent trading. Journal of Financial Economics 7, 197–226.

Dye, R. A., 1985. Disclosure of nonproprietary information. Journal of Accounting Research pp. 123–145.

Dyreng, S. D., Lindsey, B. P., Thornock, J. R., 2013. Exploring the role delaware plays as a domestic tax haven. Journal of Financial Economics 108, 751–772.

Engelberg, J., McLean, R. D., Pontiff, J., 2018. Anomalies and news. Journal of Finance 73, 1971–2001.

Faccini, R., Konstantinidi, E., Skiadopoulos, G., Sarantopoulou-Chiourea, S., 2019. A new predictor of U.S. real economic activity: The S&P 500 option implied risk aversion. Management Science 65, 4927–4949.

Faccio, M., Masulis, R., McConnell, J., 2006. Political connections and corporate bailouts. Journal of Finance 6, 2597–2635.

Fama, E. F., French, K. R., 1992. The cross-section of expected stock returns. Journal of Finance 47, 427–465.

Fama, E. F., French, K. R., 1993. Common risk factors in the returns on stocks and bonds. Journal of Financial Economics 33, 3–56.

Fama, E. F., French, K. R., 1997. Industry costs of equity. Journal of Financial Economics 43, 153–193.

Fama, E. F., French, K. R., 2000. Forecasting profitability and earnings. Journal of Business 73, 161–175.

Fama, E. F., French, K. R., 2008. Dissecting anomalies. Journal of Finance 63, 1653–1678.

Fama, E. F., French, K. R., 2015. A five-factor asset pricing model. Journal of Financial Economics 116, 1–22.

Fama, E. F., French, K. R., 2018. Choosing factors. Journal of Financial Economics 128, 234–252.

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

Fang, L., Peress, J., 2009. Media coverage and the cross-section of stock returns. Journal of Finance 64, 2023–2052.

Frazzini, A., Pedersen, L. H., 2014. Betting against beta. Journal of Financial Economics 111, 1–25.

Gao, G. P., Moulton, P. C., Ng, D. T., 2017. Institutional ownership and return predictability across economically unrelated stocks. Journal of Financial Intermediation 31, 45–63.

Garcia, D., Norli, Ø., 2012. Geographic dispersion and stock returns. Journal of Financial Economics 106, 547–565.

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

Giroud, X., 2013. Proximity and investment: Evidence from plant-level data. Quarterly Journal of Economics 128, 861–915.

Griffin, J. M., Kelly, P. J., Nardari, F., 2010. Do market efficiency measures yield correct inferences? A comparison of developed and emerging markets. Review of Financial Studies 23, 3225–3277.

Hassan, T. A., Hollander, S., van Lent, L., Tahoun, A., 2019. Firm-level political risk: Measurement and effects. Quarterly Journal of Economics 134, 2135–2202.

Hensel, C. R., Ziemba, W. T., 1995. United states investment returns during democratic and republican administrations, 1928–1993. Financial Analysts Journal 51, 61–69.

Hirshleifer, D., Hsu, P.-H., Li, D., 2013. Innovative efficiency and stock returns. Journal of Financial Economics 107, 632–654.

Hirshleifer, D., Hsu, P.-H., Li, D., 2018. Innovative originality, profitability, and stock returns. Review of Financial Studies 31, 2553–2605.

Hirshleifer, D., Lim, S. S., Teoh, S. H., 2009. Driven to distraction: Extraneous events and underreaction to earnings news. Journal of Finance 64, 2289–2325.

Hirshleifer, D., Teoh, S. H., 2003. Limited attention, information disclosure, and financial reporting. Journal of Accounting and Economics 36, 337–386.

Hong, H., Lim, T., Stein, J. C., 2000. Bad news travels slowly: Size, analyst coverage, and the profitability of momentum strategies. Journal of Finance 55, 265–295.

Hong, H., Stein, J. C., 1999. A unified theory of underreaction, momentum trading, and overreaction in asset markets. Journal of Finance 54, 2143–2184.

Hong, H., Stein, J. C., 2007. Disagreement and the stock market. Journal of Economic Perspectives 21, 109–128.

Hong, H., Torous, W., Valkanov, R., 2007. Do industries lead stock markets? Journal of Financial Economics 83, 367–396.

Hou, K., Karolyi, G. A., Kho, B.-C., 2011. What factors drive global stock returns? Review of Financial Studies 24, 2527–2574.

Hou, K., Mo, H., Xue, C., Zhang, L., 2020. An augmented q-factor model with expected growth. Review of Finance .

Hou, K., Moskowitz, T. J., 2005. Market frictions, price delay, and the cross-section of expected returns. Review of Financial Studies 18, 981–1020.

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

Huang, R. D., 1985. Common stock returns and presidential elections. Financial Analysts Journal 41, 58–61.

Huberman, G., Regev, T., 2001. Contagious speculation and a cure for cancer: A nonevent that made stock prices soar. Journal of Finance 56, 387–396.

Hutton, A. P., Marcus, A. J., Tehranian, H., 2009. Opaque financial reports, r2, and crash risk. Journal of Financial Economics 94, 67–86.

Jegadeesh, N., 1990. Evidence of predictable behavior of security returns. Journal of Finance 45, 881–898.

Jegadeesh, N., Titman, S., 1993. Returns to buying winners and selling losers: Implications for stock market efficiency. Journal of Finance 48, 65–91.

Jiang, Y., Qian, Y., Yao, T., 2016. R&d spillover and predictable returns. Review of Finance 20, 1769–1797.

Jurado, K., Ludvigson, S. C., Ng, S., 2015. Measuring uncertainty. American Economic Review 105, 1177–1216.

Kalemli-Ozcan, S., Sorensen, B., Villegas-Sanchez, C., Volosovych, V., Yesiltas, S., 2015. How to construct nationally representative firm level data from the orbis global database. Tech. rep., NBER.

Kim, C. F., Pantzalis, C., Park, J. C., 2012. Political geography and stock returns: The value and risk implications of proximity to political power. Journal of Financial Economics 106, 196–228.

Kogan, L., Papanikolaou, D., Seru, A., Stoffman, N., 2017. Technological innovation, resource allocation, and growth. Quarterly Journal of Economics 132, 665–712.

Korniotis, G. M., Kumar, A., 2013. State-level business cycles and local return predictability. Journal of Finance 68, 1037–1096.

Kothari, S., Laguerre, T. E., Leone, A. J., 2002. Capitalization versus expensing: Evidence on the uncertainty of future earnings from capital expenditures versus r&d outlays. Review of Accounting Studies 7, 355–382.

Kothari, S. P., Shu, S., Wysocki, P. D., 2009. Do managers withhold bad news? Journal of Accounting Research 47, 241–276.

La Porta, R., Lakonishok, J., Shleifer, A., Vishny, R., 1997. Good news for value stocks: Further evidence on market efficiency. Journal of Finance 52, 859–874.

Lee, C. M., So, E. C., 2015. Alphanomics: The informational underpinnings of market efficiency. Foundations and Trends® in Accounting 9, 59–258.

Lee, C. M., Sun, S. T., Wang, R., Zhang, R., 2019. Technological links and predictable returns. Journal of Financial Economics 132, 76–96.

Lettau, M., Ludvigson, S. C., Ma, S., 2019. Capital share risk in US asset pricing. Journal of Finance 74, 1753–1792.

Li, F., 2008. Annual report readability, current earnings, and earnings persistence. Journal of Accounting and Economics 45, 221–247.

Liu, J., Stambaugh, R. F., Yuan, Y., 2018. Absolving beta of volatility's effects. Journal of Financial Economics 128, 1–15.

Liu, Y., Shaliastovich, I., 2021. Government policy approval and exchange rates. Journal of Financial Economics, forthcoming.

Livnat, J., Mendenhall, R. R., 2006. Comparing the post–earnings announcement drift for surprises calculated from analyst and time series forecasts. Journal of Accounting Research 44, 177–205.

Meeuwis, M., Parker, J. A., Schoar, A., Simester, D., 2020. Belief disagreement and portfolio choice. NBER Working Paper .

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

Merton, R. C., 1973. An intertemporal capital asset pricing model. Econometrica 41, 867–887.

Miranda-Agrippino, S., Rey, H., 2020. U.S. monetary policy and the global financial cycle. Review of Economic Studies 87, 2754–2776.

Moskowitz, T. J., Grinblatt, M., 1999. Do industries explain momentum? Journal of Finance 54, 1249–1290.

Nagel, S., 2005. Short sales, institutional investors and the cross-section of stock returns. Journal of Financial Economics 78, 277–309.

Nagel, S., 2013. Empirical cross-sectional asset pricing. Annual Review of Financial Economics 5, 167–199.

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

Novy-Marx, R., 2013. The other side of value: The gross profitability premium. Journal of Financial Economics 108, 1–28.

Parker, J. A., Julliard, C., 2005. Consumption risk and the cross section of expected returns. Journal of Political Economy 113, 185–222.

Parsons, C. A., Sabbatucci, R., Titman, S., 2020. Geographic lead-lag effects. Review of Financial Studies .

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

Pastor, L., Veronesi, P., 2020. Political cycles and stock returns. Journal of Political Economy 128, 4011–4045.

Patell, J. M., Wolfson, M. A., 1982. Good news, bad news, and the intraday timing of corporate disclosures. The Accounting Review 57, 509–527.

Peng, L., 2005. Learning with information capacity constraints. Journal of Financial and Quantitative Analysis pp. 307–329.

Peng, L., Xiong, W., 2006. Investor attention, overconfidence and category learning. Journal of Financial Economics 80, 563–602.

Pirinsky, C., Wang, Q., 2006. Does corporate headquarters location matter for stock returns? Journal of Finance 61, 1991–2015.

Pontiff, J., 2006. Costly arbitrage and the myth of idiosyncratic risk. Journal of Accounting and Economics 42, 35–52.

Santa-Clara, P., Valkanov, R., 2003. The presidential puzzle: Political cycles and the stock market. Journal of Finance 58, 1841–1872.

Schmidt, P. S., von Arx, U., Schrimpf, A., Wagner, A. F., Ziegler, A., 2019. Common risk factors in international stock markets. Financial Markets and Portfolio Management 33, 213–241.

Shen, J., Yu, J., Zhao, S., 2017. Investor sentiment and economic forces. Journal of Monetary Economics 86, 1–21.

Shleifer, A., Vishny, R. W., 1997. The limits of arbitrage. Journal of Finance 52, 35–55.

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

Sims, C. A., 2003. Implications of rational inattention. Journal of Monetary Economics 50, 665–690.

Sloan, R. G., 1996. Do stock prices fully reflect information in accruals and cash flows about future earnings? Accounting Review pp. 289–315.

Stambaugh, R. F., Yu, J., Yuan, Y., 2012. The short of it: Investor sentiment and anomalies. Journal of Financial Economics 104, 288–302.

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

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

Stambaugh, R. F., Yuan, Y., 2017. Mispricing factors. Review of Financial Studies 30, 1270–1315.

Stimson, J., 1999. Public Opinion In America: Moods, Cycles, And Swings. Westview Press Inc; 2nd Edition.

Subramanian, R., Insley, R. G., Blackwell, R. D., 1993. Performance and readability: A comparison of annual reports of profitable and unprofitable corporations. The Journal of Business Communication (1973) 30, 49–61.

Teoh, S. H., Welch, I., Wong, T. J., 1998. Earnings management and the long-run market performance of initial public offerings. Journal of Finance 53, 1935–1974.

Twedt, B., 2016. Spreading the word: Price discovery and newswire dissemination of management earnings guidance. The Accounting Review 91, 317–346.

Verrecchia, R. E., 1983. Discretionary disclosure. Journal of Accounting and Economics 5, 179–194.

Verrecchia, R. E., 2001. Essays on disclosure. Journal of Accounting and Economics 32, 97–180.