Regional Economic Activity and Stock Returns

Esad Smajlbegovic[*]
University of Mannheim

First Version: October 2013
This Version: November 2013

Abstract

This study analyzes the impact of regional economic conditions on stock returns. I identify all U.S. states that are economically relevant for a firm through textual analysis of annual reports and construct a novel proxy for regional economic activity. Using this proxy, I find that economic conditions of firm-relevant U.S. regions positively influence stock returns on a monthly basis. This finding is robust to short-term reversal, individual stock momentum, industry momentum, geographic dispersion and a list of standard controls. Additionally, these results indicate that information arising from all relevant states matters over and above the information content of the mere headquarter state. Furthermore, I show that forecasts on regional economic activity predict stock returns. A zero-cost trading strategy based on this new predictive variable generates a risk-adjusted return of 6.3 (8.3) percent per year using an equal-weighted (value-weighted) portfolio. Evidence indicates that forecasts of regional activity also predict firms' real operations, suggesting that economic conditions of U.S. regions capture an important cash flow component of stock returns. Finally, this study shows that information on regional economic activity is gradually incorporated into stock prices and that the return predictability is stronger among difficult-to-arbitrage firms.

Keywords: Regional Economic Activity, Geography, 10-K Filings, Stock Returns, Real Effects, Limits-to-arbitrage
JEL: G12, G14, M41, R11

---

[*] Please address all correspondence to Esad Smajlbegovic, University of Mannheim, Business School, Chair of Finance, 68131 Mannheim, Germany; *Phone*: +49 (0)621 181 1575; *Email*: esmajlbe@mail.uni-mannheim.de.

1 Introduction

Since the seminal work of Chan, Chen, and Hsieh (1985) and Chen, Roll, and Ross (1986), scientists and practitioners have been striving to understand the empirical relationship between business cycle indicators and the financial market. While asset pricing theory derives a strong link between macroeconomic factors and the equity market, empirical studies struggle to identify these factors and produce mixed evidence as to whether they drive asset returns or not.[1] Besides the extensive list of studies that employ the sensitivity to aggregate macro risk factors to price the cross-section of returns, there are, surprisingly, very few studies that incorporate macro variables directly into asset pricing models. To address this gap in the literature, I utilize the notion of geographically segmented financial markets. Instead of looking at aggregate economic conditions, the study focuses on the regions that are economically relevant for a company and the corresponding regional economic activity. This approach enables me to construct a firm-specific macroeconomic variable and to answer two essential questions: Do changes in economic activity of firm-relevant regions affect the cross-section of stock returns? And if so, how is this macroeconomic information incorporated into the stock prices? Particularly, I asses how shocks to geographic regions of the U.S. are translated into both stock returns and real quantities (i.e., profitability) of U.S. companies.

To present some intuition as to why regional economic conditions play an important role in explaining and predicting stock returns, consider the case of the former software company viaLink Corp.[2] and the tornado outbreak in May 2-8, 1999 that ravaged the Central United States. Besides the 36 direct fatalities and almost 1000 injured people[3], this natural disaster had an extreme impact on the real economy of the region causing about $1 billion in damage. viaLink company was directly affected by the natural disaster since it mainly operated in Oklahoma in the 1990s. Figure 1 displays the stock price of the firm and the overall financial market (proxied by a one dollar investment in the market portfolio) around this event. The grey shaded region highlights the seven day period of the tornado breakout, while the circle- and cross-connected lines display viaLink's stock price and the market portfolio investment value, respectively. As evident from the Figure, the aggregate market was barely affected by the disastrous event while the company's stock price continuously decreased after the breakout from $22 down to $14 per share. Note that the shock was gradually incorporated into the stock price within almost three months. To my knowledge, there were no other firm-relevant events or news within this time period. This extreme example nicely illustrates how changes in regional economic activity and stock returns are interconnected.

---

[1] Harvey, Liu, and Zhu (2013) provide an extensive list of important macro asset pricing factors

[2] viaLink Corp. was a specialist in supply chain planning that merged with Prescient Systems, Inc. in the beginning of 2005. Later it was acquired in 2009 by Park City Group, Inc., "a trusted business solutions and services provider that enables retailers and suppliers to work collaboratively as strategic partners to reduce out-of-stocks, shrink, inventory and labor while improving profits, efficiencies, and customer service." (corporate website)

[3] see, for instance, Brooks and Doswell (2002)

[Insert Figure 1]

To extend the previous example to the entire cross-section, I decompose the U.S. market into states and obtain the differences in real economic activity across the subsample. Identifying economically relevant regions for each firm allows me to match the most important regional macroeconomic figures to each stock. Relating to the previous example, a company that is mainly operating in Oklahoma is influenced by the consumer demand and general economic activity of this state, rather than by the average national economic conditions. The same applies to a company that is economically active in California and Texas. In this case, I focus on the economic conditions of each state instead of the whole country to obtain the impact of macroeconomic changes on firm fundamentals and stock returns.

To identify the economically relevant states for each firm, I construct a firm-specific measure that assigns each of the 50 states of the U.S. a weight between zero and one. For a given company, I obtain the economic relevance measure of a U.S. state by parsing through the company's annual reports and counting the number of citations for the given state and year. The economic relevance is defined as the citation share of the state, i.e., the number of counts of the corresponding state divided by the total number of state counts.[4] For the economic activity of the states, I employ the State Coincident Indexes and the State Leading Indexes developed by Crone and Clayton-Matthews (2005) as measures of the current and future economic situation of the U.S. states. Weighting the monthly state activity indexes by the corresponding citation shares produces the contemporaneous and predicted regional economic activity proxy, CREA and PREA, respectively. Since these proxies are constructed for each firm on a monthly frequency, it allows for conducting a rich cross-sectional analysis of stock returns purely based on the level of the macroeconomic variables.

By employing the novel proxies, the paper evaluates the following four key hypotheses. First, I test in a cross-sectional analysis whether there is a contemporaneous link between regional economic activity and stock returns. Second, using forecasts of the state activity indexes, I implement a monthly updated trading strategy and assess whether publicly available forecasts of regional macroeconomic activity predict the cross-section of individual stock returns. By analyzing the operating performance in the third section, I strive to understand whether the price movements induced by the proxy are based on cash flow news or purely explained by the change of local risk aversion and demand for risky assets (Korniotis and Kumar, 2013). Last, I examine the role of limits to arbitrage (Shleifer and Vishny, 1997) in explaining the diffusion of regional activity information into stock prices and the resulting cross-predictability of returns.

---

[4] Note that I do not directly search for U.S. cities in the 10-K filings but going manually through a random sample of annual reports, I observe that city names are followed by a comma and the corresponding U.S. state. Alternatively, I also count the number of USPS postal code abbreviations for each state. This alternative algorithm increases the noise in the economic relevance proxy but the main findings of this study remain unaffected.

I test the first hypothesis by employing the well-known framework of Fama and MacBeth (1973) and regressing the contemporaneous stock return on the past month's state activity proxy. I find that the relevant states' economic conditions drive the cross-sectional return differences. Besides the standard control variables, I account for variables that may proxy for alternative explanations of the main effect such as industry momentum, geographical dispersion, and the economic conditions of the headquarter state. However, none of these factors subsumes the CREA effect. In a second test, I sort the stocks into deciles according to the state activity proxy orthognolized by other prominent firm characteristics. The top ten percent state activity portfolio significantly outperforms the bottom portfolio by, on average, more than 60 basis points per month. This spread cannot be explained by common risk factors.

Furthermore, I find that forecasts of regional economic activity predict the stock returns in the cross-section. Particularly, firms located in states with relatively high (low) economic activity forecasts are associated with increasing (decreasing) returns in the following months. Specifically, using the PREA proxy lagged by two months, I construct monthly-updated long-short portfolios and generate an annual average excess return of 6.3 percent and 8.3 percent for the equal- and value-weighted portfolio, respectively. The prediction is based on state-specific economic indicators, which are publicly available at the time of portfolio formation.

Moreover, I show that the stock market reaction on regional economic information is based on the change of firms' real performance measured by sales-on-assets and return-on-equity, respectively. Namely, a forecast of the regional economic activity growth rate predicts the firm performance of the next quarter in the same direction. Without neglecting the role of potential changes in risk aversion of local investors induced by regional economic conditions and the resulting change in local discount rates (Korniotis and Kumar, 2013), I provide evidence that the cross-predictability documented in this study is based on news about future cash flows. Furthermore, the cumulative long-run performance of the long-short portfolio is strictly positive and displays no significant reversal over the holding period. This finding supports the hypothesis that cash flow news are indeed the main driver of the regional activity-induced stock price changes.

Last, to explain the profitability of the zero-cost portfolio, I employ various firm and stock characteristics and find that mainly difficult-to-arbitrage related characteristics, e.g. volatility, illiquidity and market capitalization, are associated with the cross-predictability of stock returns. This finding is in line with the theory of Shleifer and Vishny (1997) suggesting that certain investors' limits to exploit arbitrage opportunities may generate temporary return predictability.

The new empirical findings of this paper contribute to the existing literature in several ways. There are a number of studies that incorporate macroeconomic factors in asset pricing models to explain the cross-section of stock returns. A common procedure is the estimation of the sensitivity of stock returns on the changes of aggregate business cycle variables. However, these estimations have led to mixed evidence thus far. For instance, while Chen, Roll, and Ross (1986) find that interest rate, expected and unexpected inflation, and industrial production help in pricing size portfolios, Shanken and Weinstein (2006) show that the previous findings are very sensitive to changes in the estimation method and solely industrial production growth significantly prices size-sorted portfolios.

Throughout the last decade, inflation expectation (e.g. Chan, Chen, and Hsieh, 1985; Chen, Roll, and Ross, 1986; Ferson and Harvey, 1991), consumption (e.g. Breeden, Gibbons, and Litzenberger, 1989; Lettau and Ludvigson, 2001; Parker and Julliard, 2005; Yogo, 2006; Darrat, Li, and Park, 2011), income (e.g. Campbell, 1996; Jagannathan and Wang, 1996; Eiling, 2013), foreign exchange rates (e.g. Ferson and Harvey, 1993; Bartov and Bodnar, 1994), etc. have also served as reasonable macro factors. However, in the asset pricing literature there exist very few studies that assess the role of firm-specific macro factors measured in levels or growth rates instead of sensitivities to aggregate measures. Li, Richardson, and Tuna (2012) investigate how GDP growth forecasts of the countries a firm is exposed to affect firm performance and stock returns. To the best of my knowledge, there is no study applying a similar method to assess the exposure to inter-country and regional economic changes and its implications for stock returns. By exploiting the regional heterogeneity within the U.S., I contribute to the asset pricing literature by constructing a novel macro factor measured at the level of individual firms.

In similar vein, I also add to the emerging literature on geographically segmented financial markets. For instance, Becker (2007) provides evidence on segmented U.S. bank loan markets and its effect on economic activity. Additionally, Hong, Kubik, and Stein (2008) find that stock prices are decreasing in the ratio of the aggregate book value of firms in its region over the aggregate risk tolerance of investors in its region ("only-game-in-town" effect). In the context of asset pricing, Korniotis (2008) uses regional income growth with habit-formation to explain the cross-section of expected stock returns. My study is most related to the recent study of Korniotis and Kumar (2013) who find that state-level stock returns are predictable by the local business cycle. Using quarterly data on past regional economic conditions, they find that future stock prices of firms headquartered in the same state increase (decrease) when the unemployment rates are higher (lower) and housing collateral ratios are lower (higher). They provide evidence that the effect is driven by the change in local risk aversion and coordinated local trading. Furthermore, they show that their business cycle indicators have no impact on future firm fundamentals measured on a quarterly frequency, which indicates that the return predictability is purely based on demand shifts for risky assets rather than changes in firms' cash flows. In contrast to the approach pursued by Korniotis and Kumar (2013), I identify all economically relevant states for each company instead of the single headquarter state and combine this data set with comprehensive state indexes developed by Crone and Clayton-Matthews (2005). Interestingly, this novel firm-specific proxy uncovers a clear, positive link between firms' operating performance and the economic activity of their relevant geographic regions. Consequently, this change in future cash flows is translated into stock prices.

One of the oldest, yet still interesting, topic in finance is the cross-predictability of stock returns. Besides a number of theoretical explanations for stock return predictability (e.g. Shleifer and Vishny, 1997; Hong and Stein, 1999, 2007), a growing literature illustrates different empirical patterns of return predictability and provides explanations as to how information translates into asset prices. Economic links between customers and suppliers (Cohen and Frazzini, 2008; Menzly and Ozbas, 2010), complicated industry information for conglomerates (Cohen and Lou, 2012), predictable innovation ability (Cohen, Diether, and Malloy, 2013; Hirshleifer, Hsu, and Li, 2013) and exposure to foreign countries (Li, Richardson, and Tuna, 2012; Huang, 2012; Nguyen, 2012) are just few examples of how publicly available information may predict the cross-section of individual stock returns. A very recent study by Addoum, Kumar, and Law (2013) exploits geographically distributed information on firm performance and finds that a firm's earnings and cash flows are predictable based on the performance of other firms located in regions that are economically relevant for the firm. Consequently, this performance predictability is gradually translated into stock returns. However, to my knowledge, regional macroeconomic information has yet to be used to predict the cross-section of stock returns. Specifically, the proxy for the forecast of regional economic conditions is an innovation to this literature strand. Moreover, this new proxy predicts both firm profitability and stock returns, and constitutes a potential explanation for the findings of Addoum, Kumar, and Law (2013). Furthermore, the evidence of return predictability documented in my study is stronger among difficult-to-arbitrage firms, which provides empirical support for the theoretical framework of Shleifer and Vishny (1997).

The remainder of the paper is structured as follows: Section 2 describes the data set and the construction of the novel firm-specific proxy for contemporaneous and predicted regional economic activity. In Section 3, I analyze the link between the contemporaneous proxy and stock returns. Section 4 uses a trading strategy to demonstrate that stock returns are predictable based on the forecasts of regional economic activity. In Section 5, I analyze the regional economic activity effect in more detail by examining (1) the role of the predicted proxy in forecasting firms' operating performance, (2) the long-run performance of the corresponding trading strategy, and (3) the underlying mechanism that supports the existence of regional activity induced return predictability. Section 6 concludes.

2 Data

Before analyzing the impact of regional economic activity on stock returns, I need to determine the exact definition and level of the term *region*. On the one hand, the measure of regions should provide sufficient heterogeneity in economic conditions; on the other hand, the choice of the regional level is limited by data availability. With this trade-off in mind, defining the 50 U.S. states as regions seems most appropriate in this study, as other available data is not as useful. For example, the nine U.S. Census Divisions provide less variation in economic activity, whereas the information on firms' operations in Metropolitan Statistical Areas is practically unobservable. Therefore, the terms *region* and *state* are used interchangeably in this study. To construct the firm-specific state activity proxy, I first obtain data on the economic relevance of all 50 states for each company and combine it with data on the economic activity of the respective states.

2.1 Regional Economic Activity

Similar to Bernile, Kumar, and Sulaeman (2012), the economic relevance of each state for a firm is extracted from the 10-K annual reports stored in the Electronic Data Gathering, Analysis, and Retrieval database (EDGAR) of the U.S. Securities and Exchange Commission (SEC). All American companies trading on U.S. stock exchanges are obliged to submit the Form 10-K filings within 90 days after the end of their fiscal year.[5] Besides balance sheets, income statements, and financial footnotes, the reports contain, most importantly for my purpose, data on the location of factories, warehouses, sales and branch offices. The relevant information is typically stored within the descriptions of the business evolution for the last year, the financial conditions of the company, major properties, distribution, legal proceedings, and sometimes in extensive supplemental documents.

---

[5] Depending on companies' public float, the deadlines for accelerated filers are 60 and 75 days after the end of the fiscal year. Further information is available on SEC's Form 10-K website: `http://www.sec.gov/answers/form10k.htm`

To extract this geographic information from the 10-K filings, I employ a Perl algorithm to count the number of citations for each of the 50 U.S. states within all items of the annual reports filed between 1994 and 2010. If I find more than one annual report within one fiscal year for one firm, I consider the one with the highest state citations. In case of missing reports in year τ, I use citation counts from the lagged year $\tau - 1$ (no forward-looking bias).[6] In the next step, I define economic relevance of a U.S. state for a given firm as the citation share of the state in the firm's 10-K report. Citation share of a firm-state observation is simply the ratio between the number of citations of the U.S. state and the total number of citations of all states:

$$CitShare_{i,s,\tau} = \frac{n_{i,s,\tau}}{\sum_{s=1}^{50} n_{i,s,\tau}} \tag{1}$$

where $n_{i,s,\tau}$ is the number of state s counts in firm i's annual report in year τ. $CitShare_{i,s,\tau}$ is a firm-state-year observation that, per construction, takes a value between zero and one.

To capture a coherent picture of the contemporaneous economic activity of each state, I use the State Coincident Indexes (SCI_t) developed by Crone and Clayton-Matthews (2005). This choice is justified by two arguments. First, this variable is widely accepted as an activity index in the macroeconomic literature, denoted by the authors as the "the most *comprehensive measure* of economic activity" for all 50 States. Second, it is, to my knowledge, the only consistent index published on a *monthly basis* available for all U.S. states. Similar to Stock and Watson (1989), the coincident indexes are in fact the latent dynamic factors estimated with the Kalman filter approach. To construct the SCI, Crone and Clayton-Matthews (2005) employ state-level indicator series of nonagricultural employment, the unemployment rate, average hours worked in manufacturing, and real wage and salary disbursement. Additionally, I use the authors' State Leading Indexes (SLI_t or $\widehat{SCI_{s,t+6}}$) that forecast the six-month growth rate of the coincident indexes. To estimate the SLI, the model includes data on the past and present coincident index and other variables that lead the economy: state-level housing permits, state initial unemployment insurance claims, delivery times from the Institute for Supply Management (ISM) manufacturing survey, and the interest rate spread between the 10-year Treasury bond and the 3-month Treasury bill. With very few exceptions, the different indicators of month t are released by the corresponding agencies within month $t + 1$.[7] The time-series related to state economic activity are available at the website of the Federal Reserve Bank of Philadelphia.[8]

---

[6] Data on state citation shares is available upon request.

[7] For instance, the recent government shutdown in September 2013 forced the agencies to postpone their release announcements by two or three months.

Finally, I calculate the firm-specific contemporaneous regional economic activity proxy as the citation share-weighted average of economic activity growth rate over all relevant states:

$$CREA_{i,t} = \sum_{s=1}^{50} CitShare_{i,s,\tau-1} \cdot \frac{\Delta SCI_{s,t}}{SCI_{s,t-1}}, \tag{2}$$

where $\frac{\Delta SCI_{s,t}}{SCI_{s,t-1}}$ is the growth rate of the State Coincident Index of state s in month t and $CitShare_{i,s,\tau}$ is the citation share extracted from last year's annual report. This novel proxy measures firms' exposure to current macroeconomic conditions of relevant U.S. states. Specifically, $CREA$ can be interpreted as the average monthly growth rate of economic activity over all firm-relevant U.S. states. Similarly, I construct the predicted regional activity proxy using the forecast of the state activity indexes for the next six months:

$$PREA_{i,t} = \sum_{s=1}^{50} CitShare_{i,s,\tau-1} \cdot \frac{\Delta \widehat{SCI_{s,t+5}}}{SCI_{s,t-1}}. \tag{3}$$

Using the data extracted from the annual reports, I additionally construct two state-related variables that, as shown by García and Norli (2012), explain the cross-section of expected stock returns. First, I compute the state dispersion for each firm defined as the number of distinct state names mentioned in the 10-K report:

$$StateDisp_{i,t} = \sum_{s=1}^{50} \mathbb{1}\{n_{i,s,\tau-1} > 0\}, \tag{4}$$

where $n_{i,s,\tau-1}$ is the number of citations of state s in firm's i annual report available in month t. Figure A.1 in the Appendix shows a histogram of distinct state names cited in the annual reports across all firms and years. Most firms mention three different state names and, as expected, the distribution is right-skewed. Figure A.2 in the Appendix displays the average number of distinct state names over the sample period. Note that prior to May 1996 online filing at EDGAR was not mandatory and mostly only large and geographically dispersed firms reported their filing electronically. As expected, the states of Delaware, New York and California are the most cited states in the 10-K filings. An geographic overview of the citation counts of all U.S. states is illustrated in Figure A.3. Similar results on regional dispersion can be found in García and Norli (2012) and Bernile, Kumar, and Sulaeman (2012). An alternative measure to $StateDisp$ is the concentration measure of Herfindahl-Hirschman adapted to state counts. This variable incorporates important information on the economic relevance of each state. Specifically, the measure defines a company as local if one state receives nearly all the state counts despite the firm mentioning several other state names in its annual report. Formally, the proxy is defined as:

---

[8] The web site links are `http://www.philadelphiafed.org/research-and-data/regional-economy/indexes/coincident/` and `http://www.philadelphiafed.org/research-and-data/regional-economy/indexes/leading/`

$$HHI_{i,t} = \sum_{s=1}^{50} CitShare_{i,s,\tau-1}^{2}. \tag{5}$$

Both dispersion measures are employed as control variables throughout the empirical analyses.

2.2 Other Firm Characteristics

Besides the regional variables introduced above, I incorporate a list of other firm characteristics commonly used in the asset pricing literature. In particular, I compute the market capitalization and the book-to-market ratio accounting for the size and value effect (Banz, 1981; Fama and French, 1992).[9] Furthermore, I include market beta and idiosyncratic volatility, as motivated by Ang, Hodrick, Xing, and Zhang (2009). To obtain the two variables for each stock, I run rolling time-series regressions using the CAPM on six months worth of daily data.[10] The stock specific market beta is the loading on the market proxy and idiosyncratic volatility is the standard deviation of the error term. Additionally, I control for the short-term reversal effect (Jegadeesh, 1990) and the momentum effect (Jegadeesh and Titman, 1993), including the past month return and the cumulative return from month $t-12$ to $t-2$. To account for the influence of illiquidity on stock returns, I add the logarithmized bid-ask spread calculated as the average difference of the bid and ask price divided by the midquote using daily data of the previous six months as in Amihud and Mendelson (1986).[11] To calculate the aforementioned firm-specific characteristics, I obtain daily and monthly stock returns, stock prices, bid and ask quotes, trade volume and shares outstanding from the Center for Research on Security Prices (CRSP). Additionally, accounting variables, such as book value of equity, sales, income and headquarter information come from the CRSP-Compustat merged (CCM) file.

Following the standard finance literature, I merge monthly stock returns from July in year τ to June in year $\tau + 1$ with accounting data and the regional activity proxy of year $\tau - 1$. To match the state information extracted from the SEC filings with the returns and other firm characteristics, I use the Central Index Key (CIK), and the historical link tables of the CCM database and the WRDS SEC Analytics Suite. The final sample consists of all common stocks listed on the NYSE, AMEX and NASDAQ spanning the time period from January 1995 to

---

[9] The variable construction and data matching is similar to Fama and French (1992).

[10] Variation of the asset pricing model (e.g. Fama and French (1993)) or the estimation window does not change the main findings of the study.

[11] Alternatively, I use the Amihud (2002) illiquidity measure and find similar results (not reported). The findings are robust to changes in the calculation period.

December 2011. The average number of firms per month is around 4,100[12].

Additionally, I obtain time-series of the well-known Fama and French (1993) factors, market proxy($MKT - Rf$), size (SMB), value (HML), and the momentum factor (UMD) from Kenneth French's website.[13] Data on the Pástor and Stambaugh (2003) liquidity factor (LIQ) is downloaded from Ľuboš Pástor's website.[14]. These long-short portfolios are employed as risk factors throughout the portfolio asset pricing tests.

2.3 Summary Statistics

Table 1 displays summary statistics of state-related characteristics, other firm attributes and the five asset pricing risk factors. According to the figures of Panel A, the firms are exposed to an average monthly regional economic activity growth rate of 0.15 percent or 1.8 percent annually. As a comparison, the average annual GDP growth rate of the U.S. during the same time period amounts to 2.3 percent p.a. The two figures are not necessarily equal since GDP incorporates the market value of all final goods and services produced whereas, the calculation of $CREA$ strongly depends on the firm sample. Furthermore, by weighting all firm-month observations equally, the average of the regional activity proxy underweights observations in early periods with fewer firms traded on the NYSE, AMEX, or NASDAQ, but strong growth in economic activity. Given the construction of the predicted regional economic activity, the average value of $PREA$ is around six times higher than $CREA$'s mean. The median firm is operating in 8 different U.S. states and has a state concentration of around 0.31, according to the Herfindahl-Hirschman index.

[Insert Table 1]

To compute the Pearson correlation coefficients between the variables, I initially conduct important transformations of the variables. For instance, lagged monthly returns of around 8,000 percent could potentially inflate the estimation results. Therefore, for both lagged return and cumulative return, I set all outliers above the 99th percentile to the 99th percentile (winsorizing). Furthermore, I take the logarithm of the market capitalization, book-to-market ratio, idiosyncratic volatility, and bid-ask spread since the distribution of the aforementioned variables is considerably right-skewed. Then, I calculate the cross-correlation between the variables for each month and compute its time-series average for each pair of variables. Table 2 displays the average cross-correlation between $CREA$, $PREA$ and the other (transformed) variables.

---

[12] Note again that prior to May 1996 the companies were not obliged to report the 10-K filing electronically. As a consequence, the average number of firms per month from January 1995 to December 1996 is around 1,400.

[13] The web site link is `http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html`

[14] The web site link is `http://faculty.chicagobooth.edu/lubos.pastor/research/`

As expected, the two main variables of interest are highly correlated providing evidence that the forecasts of regional economic activity highly depend on the current regional economic conditions. Additionally, I observe that current regional activity is contemporaneously positively correlated with stock returns. The same holds for the predicted state activity proxy. In general, CREA and PREA are partially correlated with other explanatory variables suggesting that one should control for these additional variables before drawing any statistical inference on the relationship between returns and regional economic activity.

[Insert Table 2]

3 Regional Economic Activity and Stock Returns

This section investigates whether regional macroeconomic conditions have an impact on stock returns. Specifically, I employ two approaches commonly used in the finance literature: regression and portfolio tests.

3.1 Regression Tests

I conduct a regression analysis along the lines of Fama and MacBeth (1973) with monthly excess returns as the dependent variable. Namely, I run a cross-sectional regression for each month t:

$$Ret_{i,t} - Rf_t = \alpha_t + \beta_t \cdot CREA_{i,t-1} + \mathbf{x'_{i,t-1}b_t} + \varepsilon_{i,t}, \tag{6}$$

where $CREA_{i,t-1}$ denotes the regional economic activity proxy of stock i in the previous month and $\mathbf{x_i}$ represents a vector of control variables depending on the specification. Then, I calculate the time-series average of each estimated regression coefficient and its t-statistic. To account for autocorrelation and heteroskedasticity in the error terms, I use the Newey and West (1987) correction with six lags. This stock-specific approach allows to easily account for other firm characteristics and disentangles the state activity effect from other possible explanations. If macroeconomic conditions of economically relevant U.S. states have an influence on stock returns, I expect a significant positive estimate of β.[15] I report the estimation results for different specifications of Equation 6 in Table 3.

The first specification in Column 1 of Table 3 considers only $CREA_{t-1}$ as explanatory variable. Using this simple design, I find that the past economic activity of relevant states has a significant and positive impact on individual stock returns. The regression coefficient associated with the proxy is 2.167 with a corresponding t-value of 2.71. To economically interpret the regression coefficient, I sort the stocks in each month according to the lagged regional activity variable into deciles. The time-series average of the proxy in the lowest and highest decile is -0.000 and 0.289 percent, respectively. Computing the difference between these values and multiplying it with the regression coefficient of 2.167 shows that a change of the state activity from the bottom to the top decile increases the return by 0.626 percentage points.

---

[15] Note that this section analyzes the contemporaneous relationship between stock returns and regional economic activity rather than the predictive power of $CREA$. Despite the fact that I use the lagged $CREA$ proxy in this regression, recall that the macroeconomic indicators of the State Coincident Index are publicly available within the next month (exceptionally within the next two months). Section 4 considers the predictability of stock returns using publicly available information on regional economic activity.

However, since the summary statistics show that the regional economic activity proxy is correlated with other stock and firm characteristics that potentially explain the cross-section of returns, it is necessary to control for those variables to avoid spurious relationships. Therefore, the second regression specification includes the state activity proxy and a set of standard control variables: market beta, market capitalization, book-to-market ratio, idiosyncratic volatility, the bid-ask spread, and past month and cumulative past return of the stock. Column 2 of Table 3 shows the result of the second Fama and MacBeth (1973) regression specification. Including the standard controls as independent variables, I find that the coefficient on the lagged state activity proxy slightly decreases to 2.085 and is highly significant at the level of one percent with a t-statistic of 4.49. Thus, the effect of state activity is robust to common firm and stock characteristics and is of similar magnitude as the result reported in the univariate specification. Furthermore, the inclusion of the standard controls even purifies the regional activity effect and substantially increases the statistical significance of the estimated coefficient.

In addition to the pronounced estimate of $CREA$, the other regression coefficients only partially explain the cross-section of individual stock returns. Consistent with previous findings in the asset pricing literature, I do not find a significant effect of the lagged market beta estimate on current returns. Furthermore, I find no significant estimation coefficients for book-to-market, idiosyncratic volatility, cumulative past return, and the bid-ask spread. Moreover, firms' market capitalization, and especially the short-term reversal, exhibit on average a strong effect on returns. In particular, small (large) firms are associated with higher (lower) expected stock returns, and last month's return negatively predicts contemporaneous stock returns. The mixed results across the well-known standard controls could be partially attributed to the relatively short time period of the analysis.

[Insert Table 3]

Besides the standard controls in Column 2, it is important to account for the potentially confounding effect of other characteristics that are closely related to economic activity of relevant states, such as industry momentum, geographic and state dispersion, or the economic conditions of the headquarter state. These variables could potentially lead to a spurious correlation of $CREA$ and stock returns. To rule out this potential endogeneity issue, I apply alternative specifications in the remainder of this section.

3.1.1 Industry Momentum

In the third specification, in addition to the standard controls, I include the past month industry return and the cumulative past industry return from month $t-12$ to month $t-2$ to account for the industry momentum. Moskowitz and Grinblatt (1999) show that a strategy that buys winning industry stocks and sells losing industry stocks is highly profitable and partially explains the individual stock momentum. Moreover, Ellison and Glaeser (1997) find that industries tend to be clustered geographically. The existence of Silicon Valley in California or the dominance of the automotive industry in Michigan are just two examples of regional industry clustering. Therefore, given the correlation between firm location and industry, the effect of the state activity proxy on stock returns might be driven by the effect of industry momentum on individual stock returns.

The empirical results of the third and fourth regression specifications in Table 3 confirm the findings of Moskowitz and Grinblatt (1999). Both lagged monthly industry return and past cumulative industry return are significantly related to contemporaneous stock returns with regression coefficients of 0.150 and 0.020, respectively. Nevertheless, the lagged state activity coefficient remains significant at the one percent significance level and decreases only slightly to 1.982. This empirical finding eliminates the possibility that the state activity effect is purely driven by industry momentum and industry clustering.

3.1.2 State Dispersion

As described in Section 2.1, the construction of the state activity proxy requires two underlying variables: the economic relevance of the U.S. states for each firm and the growth rates of the relevant state coincident indexes. Thus, $CREA$ is indirectly related to the number of distinct state names mentioned in the SEC filings, *StateDisp*. According to the summary statistics in Table 2, these two variables are negatively correlated. Moreover, using *StateDisp*, García and Norli (2012) show that firms operating in fewer U.S. states outperform geographically dispersed firms. This finding is motivated by lower investor recognition of local companies and the compensation of investors for insufficiently diversified portfolios (Merton, 1987). To rule out the possibility that the influence of lagged state activity on returns is driven by the geographic dispersion of the relevant firm, I introduce the natural logarithm of *StateDisp* into the fifth regression specification. Additionally, I implement an established alternative measure for geographic dispersion into the estimation procedure, the Herfindahl-Hirschman index.

Column 5 and 6 of Table 3 present the regression estimates with the state activity proxy, the standard controls and the two geographic dispersion measures as independent variables. The lagged state activity coefficient remains highly significant with a value of 2.018 and 2.103, respectively. This finding implies that $CREA$ plays an important role in explaining returns and the effect is not driven by the state dispersion of the firm. Furthermore, I find weak evidence for the centralized effect reported by García and Norli (2012). The coefficient of $lnStateDisp$ is negative but not significant by standard confidence levels. Similar result holds when including HHI as explanatory variable. However, this result can be mainly attributed to the different sample period. Namely, in Column 7 and 8, I restrict the sample of this study to January 1995 to December 2008, as in García and Norli (2012), and find that state dispersion negatively predicts stock returns at the five percent significance level. As suggested by García and Norli (2012), the weaker effect of state dispersion in the recent sample indicates that the trading strategy related to state dispersion was spotted and extensively implemented by arbitrageurs after the publication of the effect. With this in mind, I observe in Column 7 and 8 that the impact of state activity on stock returns is robust to changes in the sample period.

3.1.3 Headquarter State Activity

Pirinsky and Wang (2006) show the importance of firm location for asset pricing and document a co-movement in stock returns of firms headquartered in the same U.S. state. Moreover, many studies dealing with local bias focus on the headquarter region as the variable of interest, neglecting other states or regions. Therefore, the correlation between the lagged state activity over all relevant states and stock returns could be driven fully by the past economic conditions of the headquarter state. To solve this possible endogeneity issue, I decompose the regional activity proxy into a part that captures only the macroeconomic conditions of the headquarter state:

$$CREA_{i,t}^{HQ} = \sum_{s=1}^{50} \mathbb{1}_{\{s=\mathbf{HQ}\}} \cdot \frac{\Delta SCI_{s,t}}{SCI_{s,t}} \tag{7}$$

and a proxy that captures the regional economic activity of all relevant states except the headquarter state:

$$CREA_{i,t}^{ExHQ} = \frac{1}{\sum_{s=1}^{50} \mathbb{1}_{\{s\neq\mathbf{HQ}\}} \cdot n_{i,s,\tau-1}} \sum_{s=1}^{50} \mathbb{1}_{\{s\neq\mathbf{HQ}\}} \cdot n_{i,s,\tau-1} \cdot \frac{\Delta SCI_{s,t}}{SCI_{s,t}} \tag{8}$$

where $\mathbb{1}_{\{s=\mathbf{HQ}\}}$ ($\mathbb{1}_{\{s\neq\mathbf{HQ}\}}$) is an indicator function yielding one if the state s is (is not) the headquarter state. I conjecture that the economic activity of all relevant states better explains the influence of regional macroeconomic conditions on returns than the economic activity of the mere headquarter state. In other words, I expect that the regression coefficients associated with both $CREA_{i,t}^{HQ}$ and $CREA_{i,t}^{ExHQ}$ are economically and statistically significant.

As expected, Column 9 of Table 3 shows that the effects of the two regional activity variables are highly significant. Moreover, the coefficient to $CREA_{i,t}^{ExHQ}$ is nearly twice as large as the corresponding coefficient of $CREA_{i,t}^{HQ}$.[16] In untabulated estimations, I observe that including the $CREA_{i,t-1}$ variable (instead of $CREA_{i,t}^{ExHQ}$) causes the effect of the headquarter state activity to vanish. These results have the following important implication for explaining the cross-section of individual stock returns: A proxy capturing the economic conditions of all relevant states matters over and above the proxy that incorporates only information on the economic conditions of the headquarter state.

3.2 Portfolio Tests

Besides the regression framework of Fama and MacBeth (1973), the finance literature provides an alternative approach to test asset pricing models using cross-sectional data: portfolio tests. Portfolio tests are relatively simple in nature. I first sort the stocks into portfolios according to the lagged economic activity of their relevant states in each month. Second, comparing the returns of the two extreme portfolios, I formally analyze whether state activity has a positive impact on stock returns across different asset pricing models.

As shown in the regression tests and in Section 2.3, state activity is correlated with a list of other firm and stock characteristics, which could in turn confound the effect of state activity on returns. To rule out this possibility, I orthogonalize the main variable of interest by regressing state activity on all control variables that could explain the cross-section of stock returns except those that are directly related to the risk factors (market capitalization, book-to-market ratio, market beta, past cumulative return)[17]. Formally, for each month, I run a cross-sectional regression with state activity as dependent variable and the aforementioned variables as the independent variables. Then, I define for each stock-month observation the orthogonalized regional activity, $CREA_{i,t}^{\perp}$, as the regression residual:

$$CREA_{i,t} = a_t + \mathbf{x_i'b} + \varepsilon_{i,t} \tag{9}$$

$$CREA_{i,t}^{\perp} := \varepsilon_{i,t} \tag{10}$$

---

[16] Note that the average cross-sectional variances of the proxies are essentially identical, while the two variables are, on average, not significantly cross-correlated.

[17] Omission of these variables is simply motivated by the interest in the sensitivity of the long-short portfolio to the common risk factors: $MKT - Rf$, SMB, HML, UMD, and LIQ

Next, at the beginning of each month t, I sort the stocks into deciles according to their $CREA^{\perp}$ proxy in month $t-1$. Column 1 and 2 of Table 4 report average excess returns over the risk-free rate of the equal- and value-weighted decile portfolios, respectively. The univariate sorts show that the relationship between the state activity proxy and the returns of the following month is non-linear. In the case of equal-weighted portfolios, one can observe that the effect of state activity is pronounced among the two extreme deciles, while there is a weak increasing relationship between state activity and returns for the remaining portfolios. The value-weighted portfolios yield a similar result, but exhibit high variation across deciles. To test the hypothesis of whether economic activity of U.S. states has an impact on the cross-section of returns, I form a zero-cost portfolio strategy by going long in the highest decile and going short in the lowest decile. The portfolio is rebalanced every month. If state activity positively affects stock returns, I expect that this strategy yields, on average, an economically and statistically significant return. The last row of Table 4 shows that the return difference between the tenth and the first decile is positive and statistically significant regardless of the weighting method. The equal-weighted long-short portfolio earns a monthly return of 0.684 percent ($t = 5.64$) whereas the value-weighted portfolio yields a lower return of 0.491 percent ($t = 2.01$). The decrease in economic and statistical significance using the value-weighted portfolio formation indicates that the state activity effect is stronger among small stocks.

[Insert Table 4]

To rule out the possibility that the portfolio returns are just compensation for the well-known risk factors, I run a list of time-series regressions to risk-adjust the abnormal returns. I account for the market risk, the Fama and French (1993) factors, the Carhart (1997) factor, and the Pástor and Stambaugh (2003) factor, respectively. Table 5 shows that the abnormal returns remain significant employing all four asset pricing models. For instance, in Panel A for the equal-weighted portfolio, the intercept (alpha) for the Pástor and Stambaugh (2003) Model is 0.631 percent ($t = 5.23$) while the same time-series regression for the value-weighted portfolio yields a risk-adjusted return of 0.493 percent ($t = 1.93$). Comparing these results to Table 4, the returns slightly decrease after the risk-adjustments, but remain statistically significant for the equal-weighted portfolio. The value-weighted long-short portfolio is statistically significant at the ten percent level.

[Insert Table 5]

A natural question that arises is how the portfolio strategy is exposed to other risk factors. Column 5 of Table 5 reports the factor loadings of the Pástor and Stambaugh (2003) Model for both the equal-weighted and value-weighted portfolio strategy. Panel A shows that the equal-weighted returns do not load significantly on the market ($MKT - Rf$), size (SMB), and momentum (UMD) factor indicating that the long-short portfolio is well-diversified with respect to the aforementioned risk factors. Interestingly, I find that the state activity portfolio has a significantly negative exposure to the value-minus-growth (HML) portfolio. The Pástor and Stambaugh (2003) factor is positively correlated with the long-short portfolio suggesting that the strategy might be affected by illiquid stocks. Nevertheless, as already mentioned, the risk-adjusted excess return remains economically and statistically significant. Interestingly, as shown in Panel B, the returns of the value-weighted portfolio exhibit no significant loadings on the risk factors. Taking all evidence into consideration, I find that the well-known risk factors only slightly correlate with the state activity strategy and the risk-adjusted abnormal returns of the strategies remain positive and significant with an average annual return of 7.6 and 5.9 percent, respectively, confirming the regression results in Section 3.1.

3.3 Robustness Checks

To examine the stability of the relationship between regional economic activity and stock returns, I conduct a battery of robustness tests. In particular, I assess whether the results are sensitive to return adjustments, sample selection and an alternative proxy. The estimation results for all robustness tests are available in Table A.1.

3.3.1 Return Adjustments

As shown in Section 3.1, the $CREA$ effect is not driven by the past performance of clustered industries. Nevertheless, I now address this issue from a different perspective by adjusting the contemporaneous returns directly by the relevant industry return. Specifically, I divide all stocks into 49 Fama-French industry portfolios and adjust the stock returns by their industry portfolio return. Then I repeat the regression with the full specification model of Equation 6 with industry-adjusted returns instead of raw excess returns as the dependent variable. I observe in Column 1 of Table A.1 a slight decrease of the $CREA$ effect relative to the results of Table 6, yet the results remain highly significant at the 1% level.

Furthermore, Daniel, Grinblatt, Titman, and Wermers (1997) and Wermers (2004) propose a different method to assess the portfolio/stock performance: adjusting the raw returns directly by the returns of benchmark portfolios based on market capitalization, book-to-market and past cumulative return[18]. Again, I find in Column 2 of the same Table that the adjustments leave the initial results unaffected.

---

[18] The DGTW benchmarks are available via `http://www.smith.umd.edu/faculty/rwermers/ftpsite/Dgtw/coverpage.htm`

3.3.2 Sample Selection

My original sample consists of all stocks with a share code of ten or eleven that are listed on at least one of the three major stock exchanges. However, to assure that the effect of regional economic activity is not solely driven by microcaps or penny stocks, I exclude in Column 3 and 4 stocks with a price lower than one and five dollars, respectively. Then, I run the same regressions with the limited sample and find that the $CREA$-associated regression coefficient slightly decreases compared to the main results. Having in mind the modest decrease in returns going from equal-weighted to value-weighted portfolio formation documented above, this finding in the robustness test is not surprising. All in all, after excluding penny stocks, the $CREA$ effect remains highly significant in both economic and statistical terms.

The exclusion of the financial industry stocks from the sample is well-known procedure in the asset pricing literature. Following this data restriction, I again find in Column 5 that the impact of regional economic activity on stock returns remains statistically and economically significant.

Another issue that could lead to confounding results is the definition of the State of Washington and the popularity of the State of Delaware among U.S. companies. Namely, conducting the parsing algorithm, the State of Washington could be mistaken by the capital of the U.S., Washington, D.C. To avoid this confusion, I simple exclude all counts of Washington. Furthermore, the exclusion of Delaware is motivated by the business-friendly corporation laws of Delaware and the fact that over 50 percent of U.S. companies are incorporated in this state. Therefore, I construct a new proxy $CREA^{EXDeWa}$ with the same formula as in Equation 2 but ignoring citations of Washington and Delaware. Running Regression 6 with the new proxy, I observe in Column 6 a slightly lower coefficient than in the main results. Nevertheless, the estimate remains highly significant.

3.3.3 An Alternative Proxy

To construct the firm-specific proxy of regional economic activity in the main analysis, I weight the State Coincident Indexes according to the citation share of the corresponding states. However, this assumes that the citation share is a reasonable proxy for economic relevance of a state. Alternatively, I construct a proxy by equal-weighting the SCIs of all states mentioned at least once in the annual report:

$$CREA^{EW}_{i,t} = \frac{1}{StateDisp_{i,t}} \sum_{s=1}^{50} \mathbb{1}_{\{n_{i,s,\tau-1}>0\}} \times \frac{\Delta SCI_{s,t}}{SCI_{s,t-1}} \tag{11}$$

From estimation results in Column 7, I find that the alternative regional activity measure yields similar findings as the original one from the previous section. The coefficient associated with $CREA^{EW}_{i,t-1}$ is positive and significant. This finding shows the robustness of the results with respect to the construction of the proxy.

[Insert Table A.1]

3.3.4 A Placebo Test

In my final robustness check, I conduct a placebo test to assure that the effect of $CREA$ is not driven mechanically or by omitted national-wide shocks. I assign state citation shares randomly across firm-year observations and construct the corresponding placebo regional activity proxies, $CREA^{Placebo}$. If the citation shares truly capture the link between the firm and the regions, I expect that the randomly generated regional activity proxy does not significantly influence stock returns. To test this conjecture, I run 1,000 Fama and MacBeth (1973) regressions with the placebo proxies and the standard control variables introduced above. Figure 2 displays the kernel density estimation of the estimated regression coefficients and clearly shows that the economic conditions of randomly assigned U.S. states do not drive stock returns. The average estimate associated with $CREA^{Placebo}$ is essentially zero, while only around one percent of the estimates are significantly positive. Recall from Section 3.1 that the estimated regression coefficient for $CREA$ is 2.085 and differs substantially, even from the 99th percentile placebo coefficient of 0.528. In short, the placebo test confirms the importance of the link between U.S. firms and the economic relevance of U.S. states also documented in Bernile, Korniotis, and Kumar (2013).

[Insert Figure 2]

4 Forecasts on Regional Economic Activity and Cross-Predictability of Returns

The findings of Section 3 indicate that returns are positively correlated to regional economic conditions of the previous month. Having this relationship in mind, I hypothesize that if investors incorporate (publicly available) information of regional activity with a delay, then forecasts of regional economic activity can predict the cross-section of stock returns. To answer this question, I use the predicted regional economic activity proxy ($PREA$) introduced in Section 2.1. This proxy contains the forecast of the state-specific growth rate of the economic activity for the next six months. To assure that the economic indicators are publicly available before measuring the stock price reaction, I lag the predicted state activity by two months. As in Section 3, I employ the firm-specific regression approach and the portfolio approach.

4.1 Regression Tests

To test whether stock returns are predictable using forecasts for state activity, I conduct Fama and MacBeth (1973) regression analyses with monthly excess returns as the dependent variable. For each month t, I run following cross-sectional regression:

$$Ret_{i,t} - Rf_t = \alpha_t + \beta_t \cdot PREA_{i,t-2} + \mathbf{x'_{i,t-1}b_t} + \varepsilon_{i,t}, \tag{12}$$

where $PREA_{i,t-2}$ denotes the expected or predicted state activity proxy of stock i in month $t-2$ and $\mathbf{x_{i,t-1}}$ represents a vector of control variables. Again, the time-series averages and the t-statistics of the regression coefficients are of main interest. Similar to Section 3.1, I account for a list of standard controls, and the three alternative explanatory variables: industry momentum, state dispersion and the expected economic activity of the headquarter state. If forecasts on economic activity of firm-relevant U.S. states can predict stock returns, I expect a significant positive estimate of β_t. I report the estimation results for different specifications of Equation 12 in Table 6.

Column 1 of Table 6 includes $PREA_{i,t-2}$ and the standard controls: market capitalization, book-to-market ratio, market beta, idiosyncratic volatility, bid-ask spread, lagged return and lagged cumulative return. The regression coefficient for the state activity forecast proxy is highly significant and equals 0.415 with a t-statistic of 4.42. Column 2 adds the two measures for industry momentum. Both coefficients associated with $(IndRet - Rf)_{t-1}$ and $(IndRet - Rf)_{t-2,t-12}$ are positive and statistically significant. Nevertheless, $PREA_{i,t-2}$ still significantly predicts individual stock returns with a slightly decreased regression coefficient of 0.391. In Column 3 and 4, I add the two regional dispersion proxies, the logarithmized number of relevant U.S. states and the Herfindahl Index applied on citation share data. While both measures have no significant explanatory power for stock returns, the β_t estimate remains almost unchanged. Furthermore, in specification I decompose $PREA$ into the headquarter-related proxy and the proxy related to the remaining U.S. states. Column 5 shows that both proxies are highly significant and predict the cross-section of individual stock returns. In other words, similar to the findings for the $CREA$ proxy, the economic activity forecast of all relevant U.S. states has greater relevance than the forecast for the headquarter state. Finally, Column 6 combines all control variables into the model and confirms the previous results. Taking all evidence together, forecasts on state activity in month $t-2$ predict individual stock returns in month t.

Note that publicly available data is employed to exploit the trading opportunity since the regional economic indicators are published with a lag of one month. To economically interpret the regression results, I sort the stocks according to the $PREA_{i,t-2}$ proxy in each month and find that the lowest decile has an average predicted regional growth rate of 0.161 percent, whereas the top decile is associated with an average growth rate of 1.672 percent. Taking the difference between these two figures and multiplying it with the coefficient estimate of 0.379, I find that stocks within the top ten percent portfolio outperform those in the lowest ten percent portfolio, on average, by 57 basis points per month.

[Insert Table 6]

4.2 Portfolio Tests

Similar to Section 3.2, I first construct an orthogonalized state activity forecast proxy by regressing $PREA_i$ on the standard controls that are not related to the common risk factors:

$$CREA_{i,t} = a_t + \mathbf{x'_{i,t}}\mathbf{b_t} + \varepsilon_{i,t} \tag{13}$$

$$CREA^{\perp}_{i,t} := \varepsilon_{i,t} \tag{14}$$

Second, I sort all stocks at the beginning of each month t into ten portfolios according to the orthogonalized economic activity forecast proxy of month $t-2$. If $PREA^{\perp}_{i,t-2}$ positively predicts stock returns, I expect that a zero-cost portfolio that is long in the top decile and short in the bottom decile generates a positive and statistically significant return. Again, I use the four factor models to account for possible alternative risk-based explanations.

Column 1 in Panel A of Table 7 confirms the results of the regression analysis. A portfolio that goes long in the highest and short in lowest decile generates a highly significant positive excess return over the sample period. The equal-weighted portfolio yields, on average, a return of 0.560 percent (t = 5.24). This finding confirms the idea that exploiting the information on state activity forecasts helps to predict the cross-section of stock returns. However, the abnormal returns generated by the strategy might be explained by common risk-factors. Therefore, Columns 2-5 of Panel A report the time-series alphas of the different asset pricing models. The abnormal return after employing the CAPM remains significant at the one percent level with a value of 0.555. Accounting for the Fama and French (1993) factors and the momentum factor does not change the results considerably. If I augment the Carhart (1997) Model by the Pástor and Stambaugh (2003) liquidity factor, I observe a risk-adjusted return of 0.527 percent with a t-statistic of 4.87.

Panel B of Table 7 shows the corresponding results for the value-weighted portfolios. The average excess return and the risk-adjusted return of the long-short strategy are positive and significant at the five percent level across all specification. Weighting the stocks by market capitalization within the portfolios decreases the statistical significance of the results, indicating that the firm size is an important determinant of the cross-predictability of returns. I provide further insights on the role of size and other firm characteristics in Section 5.3. All in all, the five risk factors cannot explain the excess return of the regional economic activity portfolio.

[Insert Table 7]

Figure 3a plots the average monthly Pástor and Stambaugh (2003) risk-adjusted returns over the sample years. The equal-weighted long-short strategy using the state activity forecast yields a positive return in all years of the sample period. Specifically, the return is positive in 130 out of the 204 months of the sample period. Particularly, in 73 months the return is over 1.0 percent while in only 25 months the strategy yields a return lower than -1.0 percent. Figure 3b shows that the value-weighted strategy is more volatile, but surprisingly yields on average higher returns after risk-adjustments. In the first two years, my sample consists mainly of large and dispersed firms. Therefore, the low heterogeneity in this time period could lead to a poor performance of the trading strategies.

In Figure 4, I compare the two $PREA$ long-short portfolios with the performance of the market and the momentum strategy (Jegadeesh and Titman, 1993). I find that the equal-weighted state activity portfolio performs quantitatively similar to the market and the momentum strategy, while the value-weighted portfolio outperforms the corresponding portfolios in the sample period by over 25 percent. The cumulative excess returns of the state activity strategies at the end of my sample period are 114.20 and 136.94 percent, respectively.

[Insert Figure 3]

[Insert Figure 4]

Besides the positive and significant abnormal return of the trading strategy, both retail and institutional investors might be interested in the sensitivity of the portfolio return to the market return and other risk factors. Table 7 also reports the factor loadings across the different asset pricing models. I focus here on the Pástor and Stambaugh (2003) Model in the last column. Similar to the findings in Section 3.2, where $CREA$ is used as the explanatory variable, I find that the equal-weighted portfolio does not load significantly on the market and momentum factors. On the other hand, the corresponding portfolio has a significant exposure to the SMB factor. Furthermore, the Pástor and Stambaugh (2003) factor is positively correlated with the state activity strategy suggesting that the long-short portfolio is associated with considerable illiquidity risk. This finding is related to Section 5.3, where I discuss the role of firm characteristics as determinants of the state activity portfolio returns in more detail. All in all, accounting for all five risk factors provide some interesting insight on the portfolio structure but does not fully explain the abnormal return of the portfolio.

Interestingly, as shown in Panel B, the value-weighted returns exhibit different loadings on the risk factors. Specifically, the momentum factor is positively correlated with the state activity portfolio, while SMB cannot explain the variation of the state activity portfolio at a significant level. Additionally, it is surprising that the coefficient associated with the liquidity factor is negative. These findings are mainly driven by underweighting small and illiquid stocks, which decreases or even reverses their influence on the state activity portfolio.

To summarize, I find that a trading strategy based on lagged forecasts of regional economic activity combined with information on firms' economic relevant regions is partially correlated with five important risk factors, except the market portfolio. However, the explanatory power of the factors is very low resulting in a risk-adjusted alpha of 6.32 percent p.a. (equal-weighted portfolio) and 8.26 percent p.a. (value-weighted portfolio), respectively.[19]

5 Understanding the Effect of Regional Economic Activity

The findings of the previous sections show that firms exposed to regions that are expected to do well (bad) in the future generate higher (lower) returns in the subsequent month. There are two possible explanations for the positive relationship between expected state activity and stock returns. The first one suggests that the economic activity of relevant regions positively influences the expectations on regional consumer demand and firms' cash flows. Consequently, positive cash flow news drives the stock prices and increases the returns. In other words, the increase of stock returns is based on the rise in expected future profitability of the firms. However, according to the second hypothesis, the stock price reaction is explained by combining the well-known local bias and time-region-varying risk aversion (see e.g. Korniotis and Kumar (2013)). The hypothesis states that higher regional activity decreases the risk aversion of regional investors and increases the demand for risky assets. Assuming that investors prefer trading stocks of regional firms over stocks of other firms, they buy local stocks and drive their prices up.

---

[19] Similar to Section 3, I conduct a battery of robustness tests as presented in Section 3.3. The results are robust across all specifications. Estimation results are available upon request.

To distinguish between these two hypothesis, I first test whether the forecasts of regional economic activity predict real operations of firms. If the price changes are purely based on changes in local risk-aversion, I should not find any effect on firm profitability. In the second step, I examine the long-run performance of the $PREA$ long-short portfolio. If the return predictability is solely driven by changes in local risk-aversion, I should observe a reversal of the trading strategy in the long-run.

5.1 State Activity and Firm Performance

Before testing whether the predicted state activity improves the forecast of firm profitability, one should be aware of the large strand of literature that has examined the determinants of firm profitability. Besides incorporating the expected state activity of the previous quarter $PREA^{\perp}_{q-1}$ into the model, I rely on previous findings in the real effects literature and control for lagged profitability, market capitalization, book-to-market ratio, change in net operating assets, dividend yield, and a dummy variable that takes value one if the firm reports a loss in the last quarter (see for instance Fama and French, 1995, 2000; Richardson, Sloan, Soliman, and Tuna, 2005). Additionally, I control for the cumulative past stock return and the cumulative industry return to disentangle the effect of state activity from information already incorporated into the stock market. To estimate the predictive power of the independent variables, I run regressions with firm and quarter fixed effects:

$$\begin{aligned} Profitability_{i,q} = \alpha + \beta_1 Profitability_{i,q-1} + \beta_2 PREA^{\perp}_{i,q-1} + \\ + \mathbf{y'_{i,q-1}} \mathbf{b} + \mu_i + \eta_q + \varepsilon_{i,q} \end{aligned} \tag{15}$$

where firm profitability is measured by sales scaled by assets (SOA) and operating income before depreciation scaled by book value of equity (ROE), both measured in percentage. $PREA^{\perp}_{i,q-1}$ is the value of the monthly orthogonalized $PREA$ at the end of the previous quarter. $\mathbf{y'_{i,q-1}}$ denotes the vector of control variables, and μ_i and η_q denote the firm fixed effect and quarter fixed effect, respectively. The main coefficient of interest interest of this analysis lies in β_2, which is expected to be positive.

Column 1-4 of Table 8 report the empirical results for SOA as the dependent variable. The lagged expected state activity in Columns 1-3 remain positive, stable and statistically significant across the different regression specifications. For instance, in Column 3, the $PREA^{\perp}_i$ beta estimate is equal to 14.633 with a t-statistic of 4.02 indicating that forecasts on regional economic activity predicts future firm fundamentals. The coefficients on the remaining controls are in line with previous studies (see, e.g., Li, Richardson, and Tuna (2012)). The fourth specification introduces a simple alternative measure for lagged expected state activity: a dummy variable that is equal to one if $PREA_i^{\perp}$ is lower than zero[20]. The alternative measure is statistically significant at the one percent level and, as expected, the sign of the coefficient is negative.

[Insert Table 8]

The results of Column 5-8 of Table 8 confirm the predictable power of the lagged regional activity forecast when using ROE as the dependent variable. In all four regression specifications, the coefficient associated with the proxy is statistically significant at least at the five percent level suggesting that $PREA_i^{\perp}$ positively predicts future profitability of firms. Finding an effect of regional economic activity on the future performance measures suggests that the stock market reaction documented in Sections 4.1 and 4.2 is based on the change in the fundamental value of the firms.

5.2 Long-run Effect

Investors' reactions to changes in regional conditions are significant and predictable over the cross-section of stock returns. Furthermore, I find that the positive state activity-return relationship is based on fundamental changes in firms' profitability. As a consequence, I expect that state activity has a permanent impact on prices and reflects the change in fundamental value with a lag. Alternatively, one might argue that local investors overreact to information on regional activity and/or temporarily change the risk aversion as suggested by (Korniotis and Kumar, 2013). According to this hypothesis the long-run return reaction shows a reversal. The previously conducted one-month horizon analysis does not allow for answering the question whether investors over- or underreact to information on regional economic activity.

To examine the return pattern in the long-run, I apply two methods. First, I run Fama and MacBeth (1973) regressions analogously to Section 3.1 with lagged firm-specific state activity forecasts up to month $t-24$ and test the significance and sign of the regression coefficient associated with $PREA_{i,t-p}$ where $p \in \{2, ..., 24\}$. The second method uses the long-short portfolio returns as the explanatory variable. Namely, I construct for each month a long-short portfolio using $t-2$ month's predicted state activity, obtain the monthly returns in the months $t+k-1$ where $k \in \{1, ..., 24\}$ and run a time-series regression with the Pástor and Stambaugh (2003) five factors for the corresponding months. The regression intercept is defined as the average risk-adjusted portfolio return for the long-short portfolio at month $t+k$:

---

[20] Note that $PREA_i^{\perp}$ is de-meaned and, therefore, approximately half of the observations are associated with a expected state activity lower than zero.

$$\begin{aligned} LS_{t+k-1} = \alpha_k + \beta_{MKT-Rf,k}(MKT-Rf)_{t+k-1} + \beta_{SMB,k}SMB_{t+k-1} + \\ + \beta_{HML,k}HML_{t+k-1} + \beta_{UMD,k}UMD_{t+k-1} + \beta_{LIQ,k}LIQ_{t+k-1} + \varepsilon_{t+k-1,k}. \end{aligned} \tag{16}$$

The risk-adjusted return of the portfolio in month k since formation is defined as:

$$AR_k := \alpha_k \tag{17}$$

Additionally, I compute the average holding period (cumulative) return for the next k months as the following:

$$ACR_k := \sum_{j=1}^{k} \alpha_j \tag{18}$$

Note that for $k=1$, the risk-adjusted returns AR_k and ACR_k are exactly the same magnitude as estimated in Section 4.

Figure 5 displays the value of the regression coefficients and the t-statistics associated with the regional activity forecast variable lagged by p months. In each regression, I account for $PREA_{i,t-p}$ and the standard controls from Section 3.1. One can observe that the regression coefficient of the first estimation is 0.415 and the t-statistic amounts to 4.42, as reported in Table 3. The coefficient decreases over the time periods and the statistical significance vanishes after 10 months. Furthermore, the estimate converges towards zero and does not become negative, indicating that the investors do not overreact to information about regional economic activity.

[Insert Figure 5]

In Figure 6, I plot the average risk-adjusted holding period returns, ACR_k, of the equal-weighted and value-weighted long-short portfolios for different holding periods k. Overall, the average cumulative abnormal return of the portfolios is increasing over the holding period, but with decreasing monthly returns. After 24 months, the average risk-adjusted holding period returns are around three percent. It is important to note that both portfolios do not show a significant reversal in their patterns and remain positive over the entire investment horizon.

[Insert Figure 6]

The results of the long horizon investment combined with the findings of the real operations analysis provide evidence that the change in stock prices is based on new information about firms' fundamental values that are gradually incorporated into the stock market. Note that there might be other macroeconomic indicators that proxy for changes in local risk-aversion (see for instance Korniotis and Kumar, 2013) that could cause a reversal of local portfolios in the long-run. Nevertheless, the results of this paper, and particularly this section, show that regional macroeconomic indicators also affect the real operations of firms and the corresponding stock prices in the same direction.

5.3 Difficult-to-arbitrage Firms

Finding a profitable trading strategy as presented in this study, brings up the natural question: What prevents the investors to invest and arbitrage away the strategy? In this section, I explore the role of limits to arbitrage (Shleifer and Vishny, 1997), which might sustain the cross-predictability of stock returns using $PREA$. Theoretically, arbitrage opportunities vanish immediately as a high number of investors participating in the market take positions against the mispricing, driving the stock price to its fundamental value. However, in reality the stock price might diverge in short-run even further from the fundamental value that requires more risky capital. This fact could prevent the investors from arbitraging away the abnormal return in the first place and set certain limits to arbitrage.

To test whether my findings support the theory of Shleifer and Vishny (1997), I define variables related to limits-to-arbitrage and implement them as an interaction with $PREA$ into a Fama and MacBeth (1973) regression framework:

$$Ret_{i,t} - Rf_t = \alpha_t + \beta_1 PREA_{i,t-2} + \beta_2 PREA_{i,t-2} \cdot M_{i,t-2} + \beta_3 M_{i,t-2} + \mathbf{x_i'b} + \epsilon_{i,t}, \qquad (19)$$

where M denotes the interaction variable and β_2 is the coefficient of interest. In the following, I consider three different variables that are closely related to the mechanism of limits to arbitrage and commonly used in the literature: Idiosyncratic volatility, bid-ask spread of the stock and market capitalization of the firm.

Column 1 of Table 9 shows the time-series average of the cross-sectional regression coefficients employing all previously introduced control variables and the interaction term between logarithmized idiosyncratic volatility and the predicted state activity proxy. In line with my prediction, the estimated coefficient on the interaction term amounts to 0.272 and is thereby statistically significant at the ten percent level. The intuition behind this finding is that stocks with higher volatility are less attractive to arbitrageurs and exhibit a larger predictable return than stocks with lower volatility. This finding is in line with previous studies exploring return predictability (e.g. Cohen and Lou (2012)). Furthermore, the higher the illiquidity of a stock, the slower and more costly it is traded on the market. Theses additional costs could prevent investors from fully exploiting arbitrage opportunities and taking advantage of the return predictability. Therefore, the hypothesis is that the predictability effect is stronger among illiquid stocks. I measure illiquidity by the logarithm of the average daily bid-ask spread of the previous six months. Column 2 confirms the hypothesis and reports a coefficient estimate of 0.151 ($t = 2.47$) on the interaction term between *lnBidAsk* and $PREA$. Finally, since illiquid and volatile stocks are in most cases stocks with low market capitalization, I expect the stocks of small firms to be more difficult to arbitrage. The immediate implication is that the return effect of lagged $PREA$ is stronger for smaller sized firms. Column 3 provides evidence for this prediction. The Fama and MacBeth (1973) regression coefficient of the interaction term between *lnSize* and $PREA$ is -0.058 ($t = -2.01$).[21]

[Insert Table 9]

Volatility, illiquidity and market capitalization are well-known proxies for limits to arbitrage. However, since all three variables are highly correlated, it is difficult to disentangle the effect of each proxy. Nevertheless, I argue that the variables mutually influence the effect of state activity on stock returns. Specifically, stocks that are small, volatile and illiquid are more difficult to arbitrage and display higher predictable returns.

6 Conclusion

This study investigates the link between stock returns and economic conditions of firm-relevant U.S. regions. Based on a textual analysis of annual financial reports in combination with regional economic indicators, I construct a novel proxy that measures the firm-specific exposure to the regional economy.

Using this proxy, I find that economic conditions of relevant regions have a positive impact on stock returns. This result is strengthened due to the fact that this effect cannot be explained by industry momentum, geographic dispersion, economic activity of the headquarter state and various other well-known cross-sectional effects. Furthermore, I provide evidence that the heterogeneity in local economic conditions generates predictable stock return patterns; higher predicted economic growth in the relevant regions increases future stock prices. In particular, using forecasts on state economic activity, I construct a trading strategy that yields a risk-adjusted return of over five percent p.a. Consistent with the theory of Shleifer and Vishny (1997), I find that the predictability is stronger among difficult-to-arbitrage firms. Finally, this study indicates that the economic activity of firms' locations has a strong impact on firms' profitability. This finding provides evidence that the stock market reaction is based on news about future cash flows rather than on changes in local risk-aversion, as suggested by Korniotis and Kumar (2013). Besides uncovering new links between regional economic conditions, firm profitability and stock returns, this study highlights the importance of all economically relevant regions as opposed to just the headquarter location (similar to Bernile, Kumar, and Sulaeman, 2012).

---

[21] Besides the three mechanism variables employed above, I additionally test other potential return predictability mechanisms such as investors' limited processing capacity or investors' limited attention. Using different proxies closely related to information complexity and investor attention, I find only weak evidence that supports these two hypotheses.

Besides the new empirical findings, this study additionally poses many interesting questions related to geographically segmented markets. For instance, a question that remains unanswered in this study is whether investors require compensation for holding stocks with higher sensitivity to regional economic conditions. Specifically, the comparison of asset pricing models with regional and aggregate macroeconomic factors is of particular interest. Moreover, this study reveals certain correlations between the regional economic activity portfolios and the common risk factors. Particularly, the strong loadings on the book-to-market and momentum portfolios remain unanswered. Similarly, the role of geographic regions in explaining the heterogeneity in other firm or stock characteristics such as liquidity (Bernile, Korniotis, Kumar, and Wang, 2013) provides an interesting area for future research.

References

Addoum, J. M., A. Kumar, and K. Law, 2013, “Geographic Diffusion of Information and Stock Returns,” SSRN Scholarly Paper ID 2343335, Social Science Research Network, Rochester, NY.

Amihud, Y., 2002, “Illiquidity and stock returns: cross-section and time-series effects,” Journal of Financial Markets, 5(1), 31–56.

Amihud, Y., and H. Mendelson, 1986, “Asset pricing and the bid-ask spread,” Journal of Financial Economics, 17(2), 223–249.

Ang, A., R. J. Hodrick, Y. Xing, and X. Zhang, 2009, “High idiosyncratic volatility and low returns: International and further U.S. evidence,” Journal of Financial Economics, 91(1), 1–23.

Banz, R. W., 1981, “The relationship between return and market value of common stocks,” Journal of Financial Economics, 9(1), 3–18.

Bartov, E., and G. M. Bodnar, 1994, “Firm Valuation, Earnings Expectations, and the Exchange-Rate Exposure Effect,” The Journal of Finance, 49(5), 1755–1785.

Becker, B., 2007, “Geographical segmentation of US capital markets,” Journal of Financial Economics, 85(1), 151–178.

Bernile, G., G. M. Korniotis, and A. Kumar, 2013, “Geography of Firms and Propagation of Local Economic Shocks,” SSRN Scholarly Paper ID 2064141, Social Science Research Network, Rochester, NY.

Bernile, G., G. M. Korniotis, A. Kumar, and Q. Wang, 2013, “Local Business Cycles and Local Liquidity,” SSRN Scholarly Paper ID 2052159, Social Science Research Network, Rochester, NY.

Bernile, G., A. Kumar, and J. Sulaeman, 2012, “Home Away From Home: Economic Relevance and Local Investors,” SSRN Scholarly Paper ID 1711653, Social Science Research Network, Rochester, NY.

Breeden, D. T., M. R. Gibbons, and R. H. Litzenberger, 1989, “Empirical Test of the Consumption-Oriented CAPM,” The Journal of Finance, 44(2), 231–262.

Brooks, H. E., and C. A. Doswell, 2002, “Deaths in the 3 May 1999 Oklahoma City Tornado from a Historical Perspective,” Weather and Forecasting, 17(3), 354–361.

Campbell, J. Y., 1996, "Understanding Risk and Return," Journal of Political Economy, 104(2), 298–345.

Carhart, M. M., 1997, "On Persistence in Mutual Fund Performance," The Journal of Finance, 52(1), 57–82.

Chan, K., N.-f. Chen, and D. A. Hsieh, 1985, "An exploratory investigation of the firm size effect," Journal of Financial Economics, 14(3), 451–471.

Chen, N.-F., R. Roll, and S. A. Ross, 1986, "Economic Forces and the Stock Market," The Journal of Business, 59(3), 383–403.

Cohen, L., K. Diether, and C. Malloy, 2013, "Misvaluing Innovation," Review of Financial Studies, 26(3), 635–666.

Cohen, L., and A. Frazzini, 2008, "Economic Links and Predictable Returns," The Journal of Finance, 63(4), 1977–2011.

Cohen, L., and D. Lou, 2012, "Complicated firms," Journal of Financial Economics, 104(2), 383–400.

Crone, T. M., and A. Clayton-Matthews, 2005, "Consistent Economic Indexes for the 50 States," The Review of Economics and Statistics, 87(4), 593–603.

Daniel, K., M. Grinblatt, S. Titman, and R. Wermers, 1997, "Measuring Mutual Fund Performance with Characteristic-Based Benchmarks," The Journal of Finance, 52(3), 1035–1058.

Darrat, A. F., B. Li, and J. C. Park, 2011, "Consumption-based CAPM models: International evidence," Journal of Banking & Finance, 35(8), 2148–2157.

Eiling, E., 2013, "Industry-Specific Human Capital, Idiosyncratic Risk, and the Cross-Section of Expected Stock Returns," The Journal of Finance, 68(1), 43–84.

Ellison, G., and E. L. Glaeser, 1997, "Geographic Concentration in U.S. Manufacturing Industries: A Dartboard Approach," Journal of Political Economy, 105(5), 889–927.

Fama, E. F., and K. R. French, 1992, "The Cross-Section of Expected Stock Returns," The Journal of Finance, 47(2), 427–465.

——— , 1993, "Common Risk Factors in the Returns On Stocks And Bonds," Journal of Financial Economics, 33, 3–56.

——— , 1995, "Size and Book-to-Market Factors in Earnings and Returns," The Journal of Finance, 50(1), 131–155.

——— , 2000, "Forecasting Profitability and Earnings," The Journal of Business, 73(2), 161–175.

Fama, E. F., and J. D. MacBeth, 1973, "Risk, Return, and Equilibrium: Empirical Tests," Journal of Political Economy, 81(3), 607–636.

Ferson, W. E., and C. R. Harvey, 1991, "The Variation of Economic Risk Premiums," Journal of Political Economy, 99(2), 385–415.

——— , 1993, "The Risk and Predictability of International Equity Returns," Review of Financial Studies, 6(3), 527–566.

García, D., and y. Norli, 2012, "Geographic dispersion and stock returns," Journal of Financial Economics, 106(3), 547–565.

Harvey, C. R., Y. Liu, and H. Zhu, 2013, "…and the Cross-Section of Expected Returns," SSRN Scholarly Paper ID 2249314, Social Science Research Network, Rochester, NY.

Hirshleifer, D., P.-H. Hsu, and D. Li, 2013, "Innovative efficiency and stock returns," Journal of Financial Economics, 107(3), 632–654.

Hong, H., J. D. Kubik, and J. C. Stein, 2008, "The only game in town: Stock-price consequences of local bias," Journal of Financial Economics, 90(1), 20–37.

Hong, H., and J. C. Stein, 1999, "A Unified Theory of Underreaction, Momentum Trading, and Overreaction in Asset Markets," The Journal of Finance, 54(6), 2143–2184.

Hong, H., and J. C. Stein, 2007, "Disagreement and the Stock Market," Journal of Economic Perspectives, 21(2), 109–128.

Huang, X., 2012, "Gradual Information Diffusion in the Stock Market: Evidence from U.S. Multinational Firms," SSRN Scholarly Paper ID 1963648, Social Science Research Network, Rochester, NY.

Jagannathan, R., and Z. Wang, 1996, "The Conditional CAPM and the Cross-Section of Expected Returns," The Journal of Finance, 51(1), 3–53.

Jegadeesh, N., 1990, "Evidence of Predictable Behavior of Security Returns," The Journal of Finance, 45(3), 881–898.

Jegadeesh, N., and S. Titman, 1993, "Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency," The Journal of Finance, 48(1), 65–91.

Korniotis, G. M., 2008, "Habit Formation, Incomplete Markets, and the Significance of Regional Risk for Expected Returns," Review of Financial Studies, 21(5), 2139–2172.

Korniotis, G. M., and A. Kumar, 2013, "State-Level Business Cycles and Local Return Predictability," The Journal of Finance, p. n/a–n/a.

Lettau, M., and S. Ludvigson, 2001, "Resurrecting the (C)CAPM: A Cross-Sectional Test When Risk Premia Are Time-Varying," Journal of Political Economy, 109(6), 1238–1287.

Li, N., S. Richardson, and A. Tuna, 2012, "Macro to Micro: Country Exposures, Firm Fundamentals and Stock Returns," SSRN Scholarly Paper ID 2017091, Social Science Research Network, Rochester, NY.

Menzly, L., and O. Ozbas, 2010, "Market Segmentation and Cross-predictability of Returns," The Journal of Finance, 65(4), 1555–1580.

Merton, R. C., 1987, "A Simple Model of Capital Market Equilibrium with Incomplete Information," The Journal of Finance, 42(3), 483–510.

Moskowitz, T. J., and M. Grinblatt, 1999, "Do Industries Explain Momentum?," The Journal of Finance, 54(4), 1249–1290.

Newey, W. K., and K. D. West, 1987, "A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix," Econometrica, 55(3), 703–708.

Nguyen, Q., 2012, "Does Investor Attention Stop at the Water's Edge? The Predictability of Multinational Firms' Returns," SSRN Scholarly Paper ID 1921537, Social Science Research Network, Rochester, NY.

Parker, J. A., and C. Julliard, 2005, "Consumption Risk and the Cross Section of Expected Returns," Journal of Political Economy, 113(1), 185–222.

Pirinsky, C., and Q. Wang, 2006, "Does Corporate Headquarters Location Matter for Stock Returns?," The Journal of Finance, 61(4), 1991–2015.

Pástor, u., and R. F. Stambaugh, 2003, "Liquidity Risk and Expected Stock Returns," Journal of Political Economy, 111(3), 642–685.

Richardson, S. A., R. G. Sloan, M. T. Soliman, and r. Tuna, 2005, "Accrual reliability, earnings persistence and stock prices," Journal of Accounting and Economics, 39(3), 437–485.

Shanken, J., and M. I. Weinstein, 2006, "Economic forces and the stock market revisited," Journal of Empirical Finance, 13(2), 129–144.

Shleifer, A., and R. W. Vishny, 1997, "The Limits of Arbitrage," The Journal of Finance, 52(1), 35–55.

Stock, J. H., and M. W. Watson, 1989, "New Indexes of Coincident and Leading Economic Indicators," NBER chapters, National Bureau of Economic Research, Inc.

Wermers, R., 2004, "Is Money Really 'Smart'? New Evidence on the Relation Between Mutual Fund Flows, Manager Behavior, and Performance Persistence," SSRN Scholarly Paper ID 414420, Social Science Research Network, Rochester, NY.

Yogo, M., 2006, "A Consumption-Based Explanation of Expected Stock Returns," The Journal of Finance, 61(2), 539–580.

7 Figures and Tables

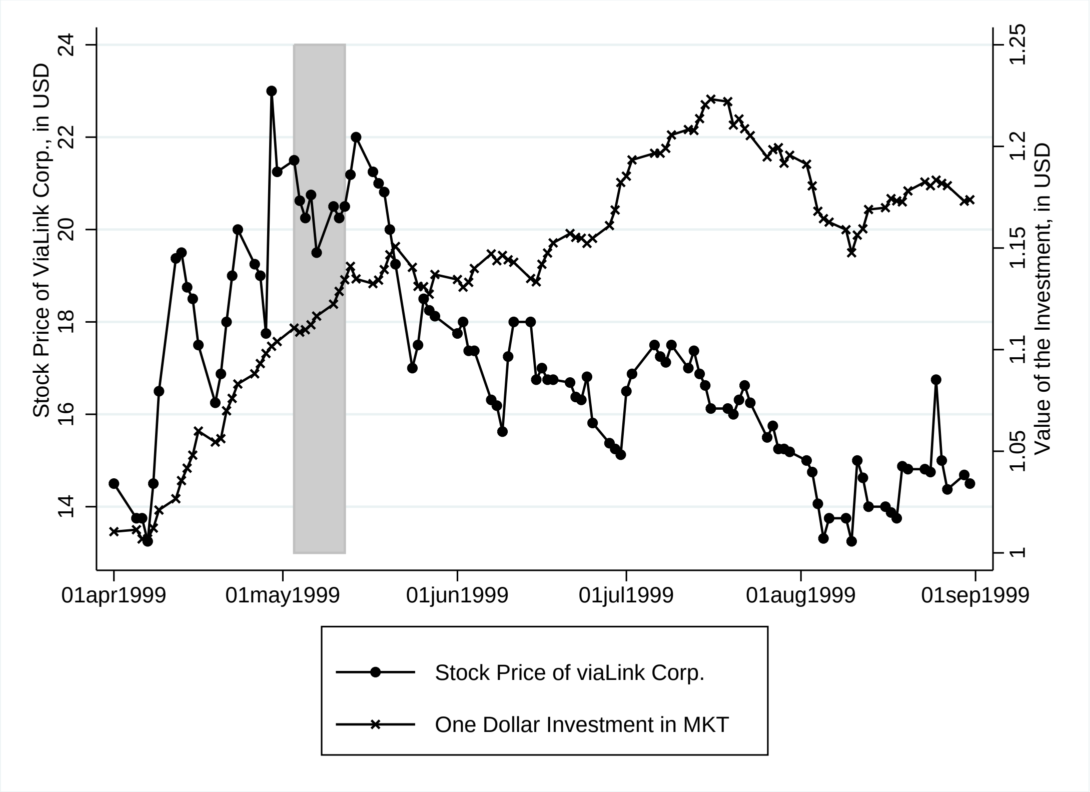


Figure 1: Example of the viaLink Corp. Stock price During the 1999 Tornado outbreak

This figure displays the stock price of viaLink Corp. and the overall financial market (proxied by a one dollar investment in the equal-weighted market portfolio) around the tornado breakout in Central United States. The grey shaded region highlights the seven days period of the storm, while viaLink's stock price and the market portfolio investment value are displayed by the circle- and cross-connected lines, respectively.

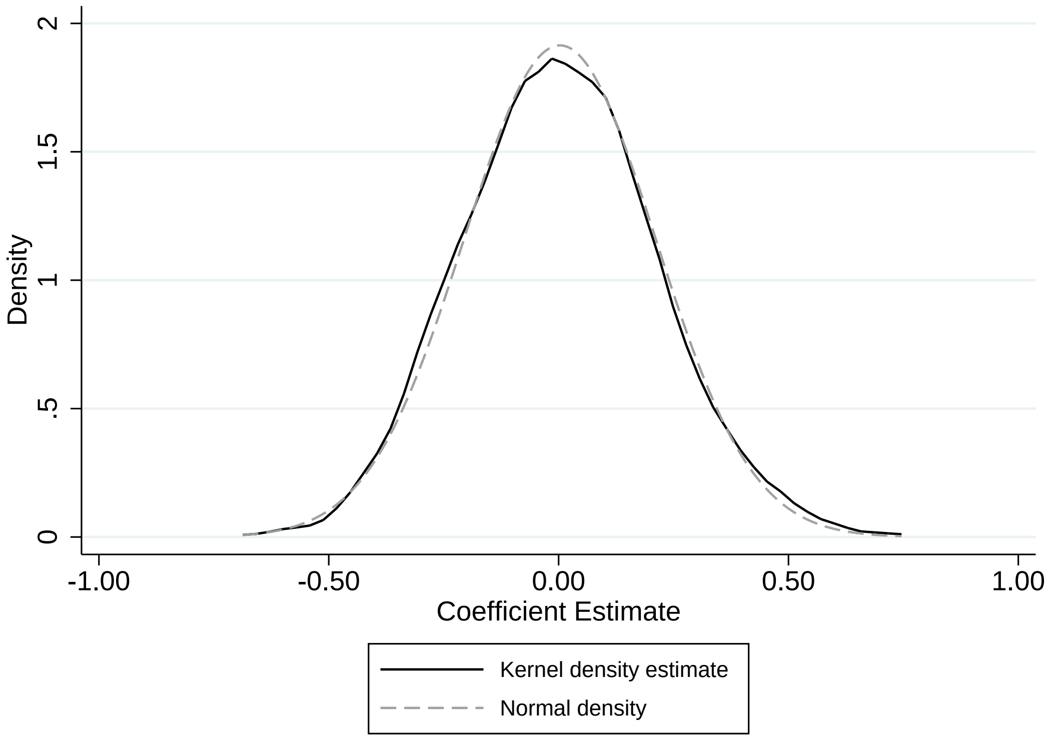
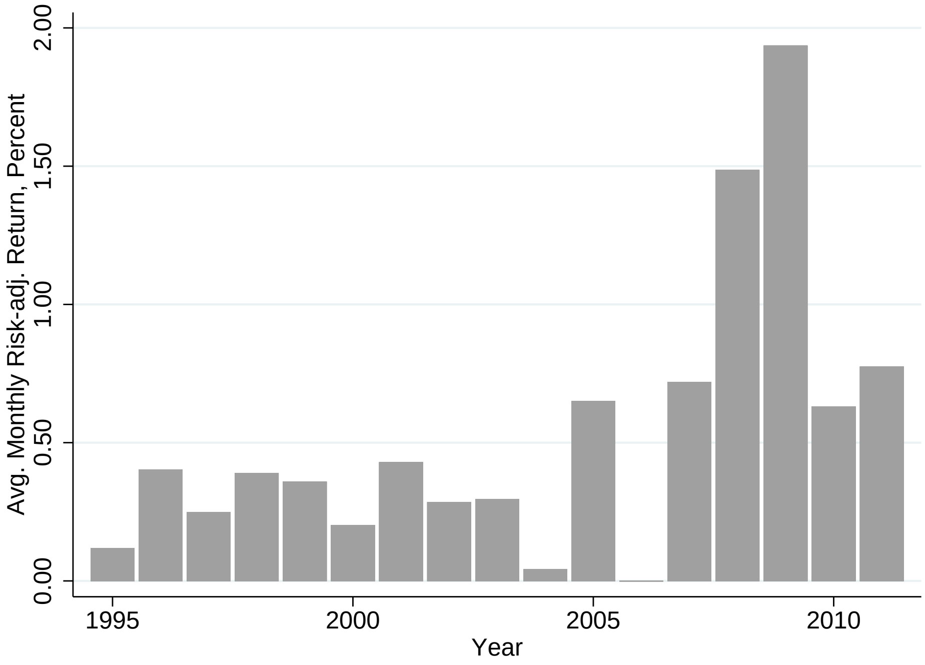
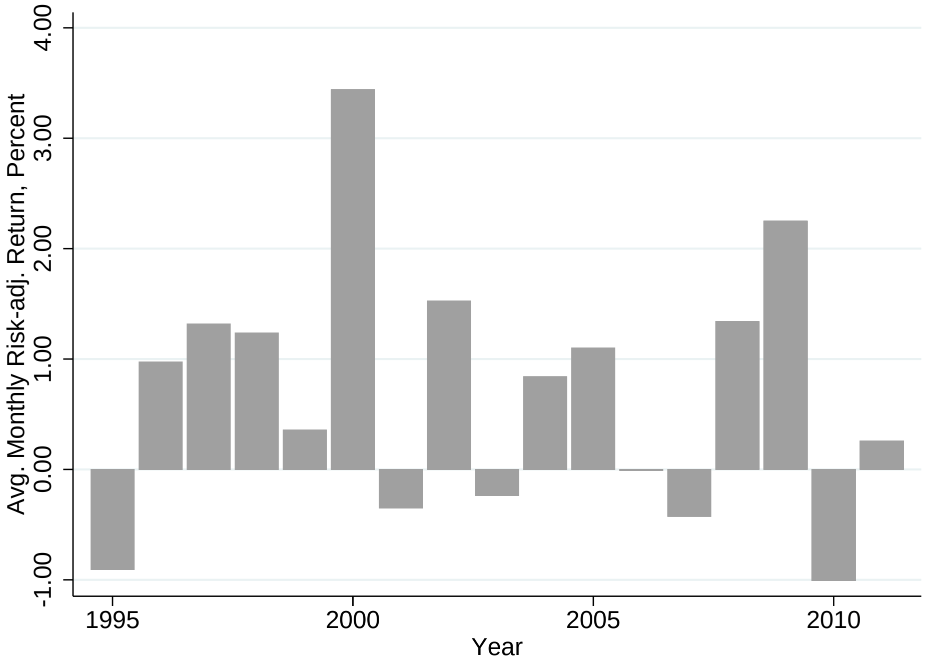


Figure 2: Kernel Density Estimation of the Placebo Regression Coefficients

This figure shows the kernel density estimation (black solid line) of the regression coefficients to the 1,000 $CREA_{t-1}$ placebo proxies controlling for the standard control variables. The grey dashed line displays the corresponding normal density. The sample period is January 1995 through December 2011.



(a) Equal-weighted Portfolio



(b) Value-weighted Portfolio

Figure 3: Performance of the Predicted Regional Activity-based Trading Strategy Over the Sample Period

This figure shows for each year of the sample period the average monthly Pástor and Stambaugh (2003) risk-adjusted return of the zero-investment LS portfolio. The LS portfolio is constructed each month by going long in the highest decile and short in the lowest decile according to $PREA_{t-2}^{\perp}$. The Figure in (a) plots the results of the equal-weighted LS portfolio and Figure (b) plots the value-weighted LS portfolio results. The sample period is January 1995 through December 2011.

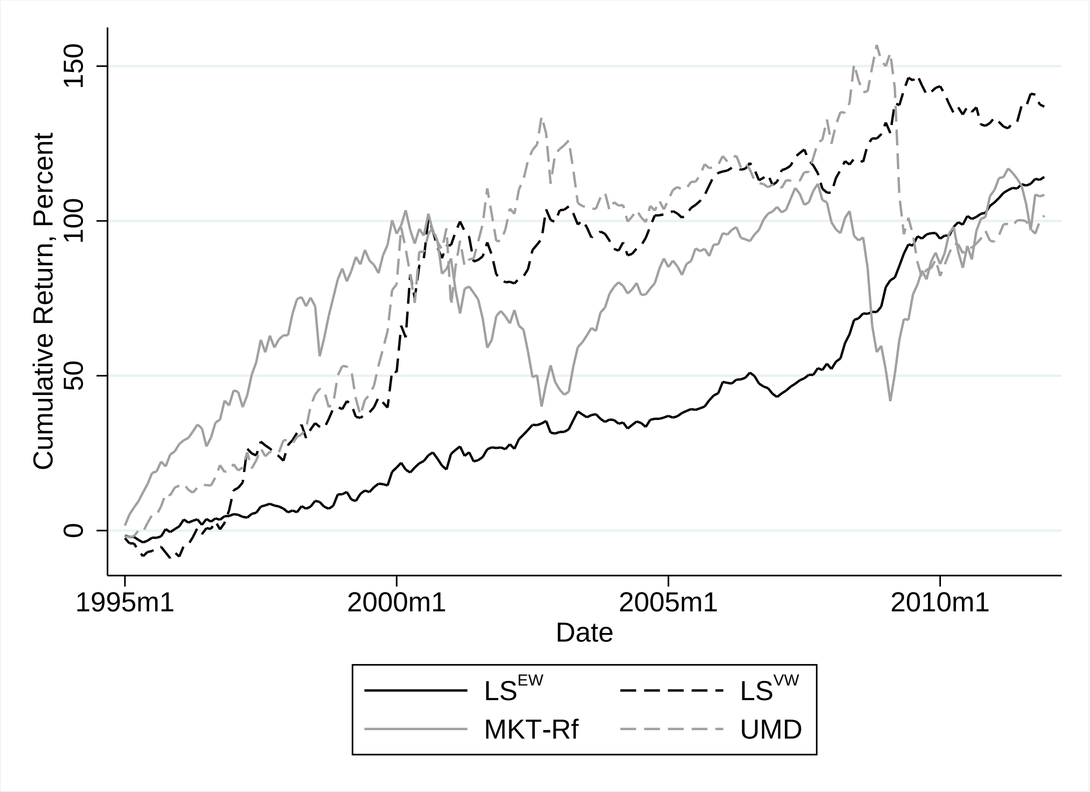


Figure 4: Cumulative Return of the Regional Activity-based Trading Strategies, the Market Proxy, and the Momentum Strategy

This figure shows the cumulative return performance of the equal- and value-weighted long-short $PREA^{\perp}_{t-2}$ portfolio (black solid and black dashed line, respectively), the value-weighted market portfolio over the risk-free rate (grey solid line), and the momentum portfolio (grey dashed line). The sample period is January 1995 through December 2011.

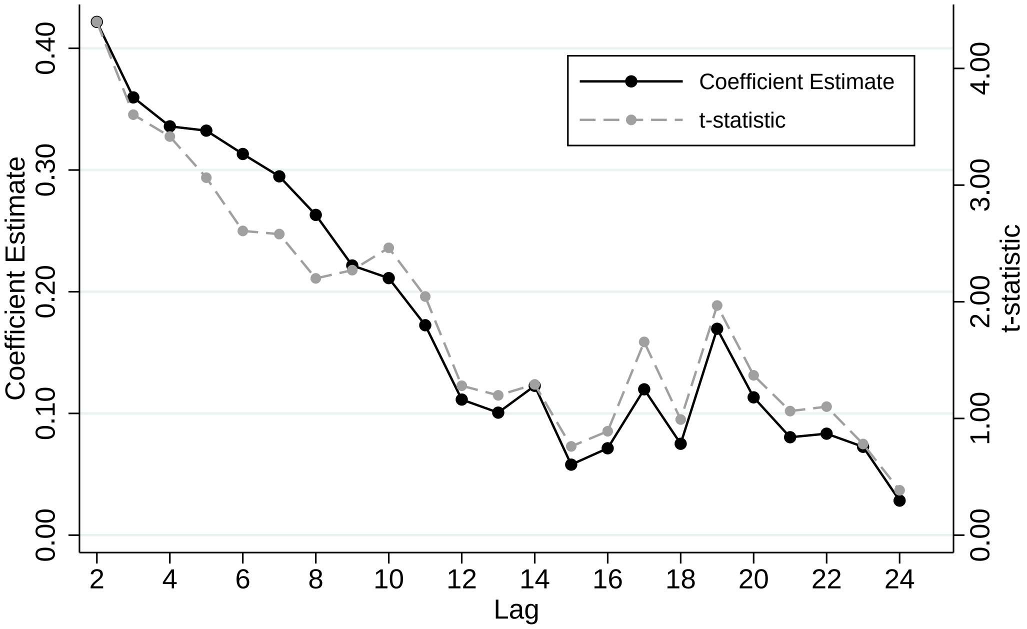


Figure 5: $PREA$ Regression Coefficients in Long-run
I run separate Fama and MacBeth (1973) regressions with standard control variables and lagged firm-specific state activity forecasts up to month $t-24$. This figure plots the regression coefficients (black solid line) and the corresponding t-statistics (grey dashed line) associated with $PREA_{i,t-p}$ where $p \in \{2, ..., 24\}$. The sample period is January 1995 through December 2011.

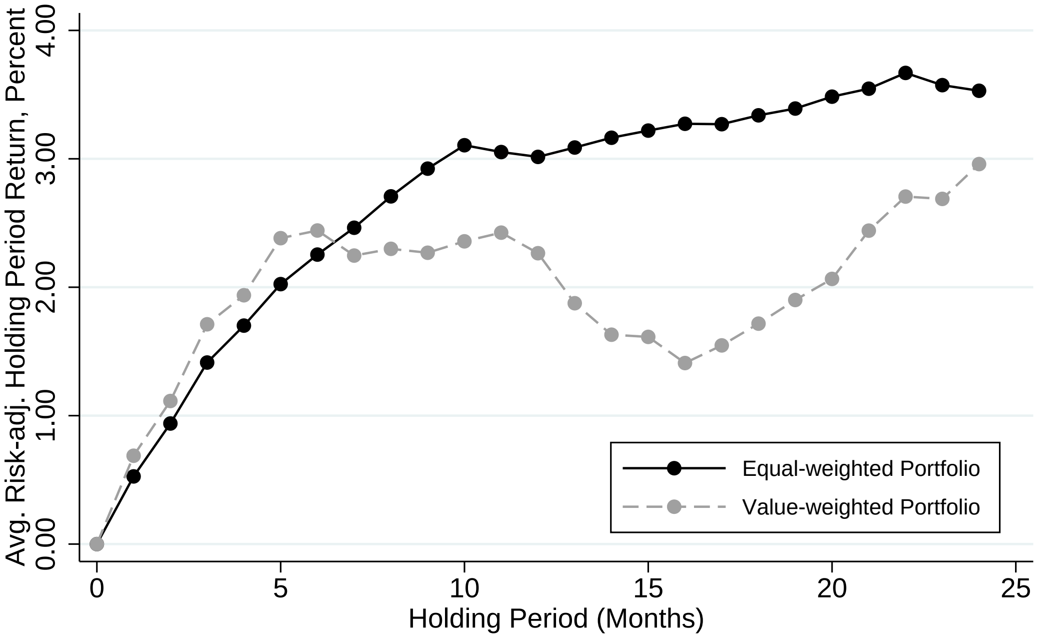


Figure 6: Long-horizon Performance of the Predicted Regional Activity Portfolio

This figure plots the average cumulative risk-adjusted return of the $PREA^{\perp}_{t-2}$ long-short portfolio. To calculate the long-horizon performance, I first construct each month the long-short portfolio according to $PREA^{\perp}_{t-2}$ and obtain the monthly returns in the months $t + k - 1$ where $k \in \{1, ..., 24\}$. Second, I run for each horizon a time-series regression with the Pástor and Stambaugh (2003) five factors for the corresponding months. The regression intercept is defined as the average risk-adjusted portfolio return for the long-short portfolio at month $t + k$:

$$\begin{aligned} LS_{t+k-1} = \alpha_k + \beta_{MKT-Rf,k}(MKT - Rf)_{t+k-1} + \beta_{SMB,k}SMB_{t+k-1} + \\ + \beta_{HML,k}HML_{t+k-1} + \beta_{UMD,k}UMD_{t+k-1} + \beta_{LIQ,k}LIQ_{t+k-1} + \varepsilon_{t+k-1,k}. \end{aligned}$$

For the third and final step, the average holding period (cumulative) risk-adjusted return for the next k months since formation is defined as:

$$ACR_k := \sum_{j=1}^{k} \alpha_j$$

The sample period is January 1995 through December 2011.

Table 1: Summary Statistics of Explanatory Variables and Risk Factors

This table reports univariate statistics (i.e. mean, standard deviation, and the 1st, 25th, 50th, 75th and 99th percentile) for a set of variables. Panel A displays the state-related variables: the contemporaneous regional economic activity proxy ($CREA$), the predicted regional economic activity proxy ($PREA$), the number of distinct state names cited in firms' annual reports ($StateDisp$), and the Herfindahl-Hirschman concentration measure based on state citations (HHI). $CREA$ ($PREA$) is constructed from a linear combination of current (predicted) state economic activity growth rates weighted by the citation share of economically relevant states. The additional firm characteristics in Panel B include standard control variables employed throughout the analyses: $Beta$ is the stock-specific market beta calculated using rolling regressions with daily returns of the past 252 days and the market return as the explanatory variable (CAPM). $ISVolatility$ is the standard deviation of the corresponding error term. $Size$ (the market capitalization) and BM (the book-to-market ratio) are computed as in Fama and French (1992). The $Bid-AskSpread$ is calculated as the average difference of the bid and ask price divided by the midquote using daily data of the previous six months as in Amihud and Mendelson (1986). Furthermore, $(Ret-Rf)_{t-1}$ is the lagged excess return (Jegadeesh, 1990), while $(Ret-Rf)_{t-12,t-2}$ denotes the cumulative excess return from month $t-12$ to $t-2$ capturing the momentum effect (Jegadeesh and Titman, 1993). Panel C displays the descriptive statistics of the five tradable common risk factors $MKT-Rf$, SMB, HML, UMD, and LIQ (Fama and French, 1993; Carhart, 1997; Pástor and Stambaugh, 2003).

			Percentile				
Variable	Mean	SD	1st	25th	Median	75th	99th
Panel A: State-related Variables							
$CREA$	0.148	0.269	-0.790	0.030	0.217	0.331	0.524
$PREA$	0.917	1.332	-3.778	0.339	1.258	1.822	2.820
$StateDisp$	10.892	8.818	1	5	8	14	46
HHI	0.369	0.223	0.061	0.200	0.314	0.485	1
Panel B: Other Firm Characteristics							
$Beta$	0.799	0.635	-0.364	0.325	0.726	1.184	2.584
$Size$	2656.798	13900	3.371	52.552	217.037	982.211	47600
BM	0.806	1.728	0.039	0.319	0.562	0.918	4.725
$ISVolatility$	0.037	0.027	0.008	0.019	0.029	0.046	0.131
$Bid-AskSpread$	0.025	0.047	0	0.003	0.011	0.029	0.196
$(Ret_i-Rf)_{t-1}$	0.009	0.198	-0.434	-0.075	-0.001	0.073	0.632
$(Ret_i-Rf)_{t-12,t-2}$	0.111	0.824	-0.881	-0.264	0.010	0.300	2.882
Panel C: Portfolios							
$MKT-Rf$	0.531	4.799	-10.76	-2.315	1.325	3.595	9.24
SMB	0.223	3.680	-6.750	-2.170	-0.110	2.475	7.730
HML	0.249	3.477	-9.780	-1.650	0.220	1.910	9.120
UMD	0.499	5.652	-16.29	-1.325	0.770	3.135	13.200
LIQ	0.750	4.149	-9.257	-1.432	0.598	2.981	11.001

Table 2: Correlation Between the Explanatory Variables And the Risk Factors

This table reports the time-series average of the cross-sectional correlation coefficients between a set of variables (Panel A) and the correlation between the long-short portfolios $MKT - Rf$, SMB, HML, UMD, and LIQ (Panel B). Particularly, the correlation matrix in Panel A displays average correlations between $CREA$, $PREA$, HHI, $Beta$, logarithmized $StateDisp$, $Size$, BM, $ISVola$, and $Bid - AskSpread$ and the current and past cumulative stock return. I take the logarithm of some of the aforementioned variables due to a skewed empirical distribution of the variables. Panel B shows the simple Pearson correlation coefficient of the common risk factors over the sample period. The sample period is January 1995 through December 2011.

Panel A: Firm Characteristics

	$CREA$	$PREA$	$lnStateDisp$	HHI	$Beta$	$lnSize$	$lnBeme$	$lnISVola$	$lnBidAsk$	$(Ret_i - Rf)_{t-1}$	$(Ret_i - Rf)_{t-12,t-2}$
$CREA$	1										
$PREA$	0.818	1									
$lnStateDisp$	-0.021	-0.022	1								
HHI	-0.004	-0.003	-0.762	1							
$Beta$	0.046	0.051	0.130	-0.168	1						
$lnSize$	-0.006	-0.009	0.381	-0.268	0.377	1					
$lnBeme$	-0.037	-0.036	0.020	0.042	-0.251	-0.271	1				
$lnISVola$	0.057	0.065	-0.191	0.043	0.078	-0.630	-0.030	1			
$lnBidAsk$	-0.006	-0.009	-0.288	0.219	-0.459	-0.852	0.291	0.585	1		
$(Ret_i - Rf)$	0.007	0.008	-0.004	0.003	-0.019	-0.009	0.025	-0.011	0.009	1	
$(Ret_i - Rf)_{t-12,t-2}$	0.032	0.036	-0.015	0.008	0.048	0.187	0.019	-0.073	-0.184	0.012	1

Panel B: Portfolios

	MKT-Rf	SMB	HML	UMD	LIQ
MKT-Rf	1				
SMB	0.257	1			
HML	-0.236	-0.361	1		
UMD	-0.278	0.089	-0.151	1	
LIQ	0.128	0.051	-0.109	0.047	1

Table 3: Regional Economic Activity and Stock Returns

This table reports the average cross-sectional regression coefficients using the Fama and MacBeth (1973) framework:

$$Ret_{i,t} - Rf_t = \alpha_t + \beta_t \cdot CREA_{i,t-1} + \mathbf{x'_{i,t-1}b_t} + \varepsilon_{i,t},$$

where $Ret_{i,t}$ is the return of stock i in month t and Rf_t is the monthly yield on 30-day Treasury bills, and $\mathbf{x'_i}$ is a vector of other firm characteristics. $CREA_{i,t-1}$ is the sum of growth rates of regional economic activity across U.S. states (Crone and Clayton-Matthews, 2005) weighted by the corresponding citation shares of the states extracted from firms' 10-K reports. Other control variables include: $Beta$ is the stock-specific market beta calculated using rolling regressions with daily returns of the past 252 days and the market return as the explanatory variable (CAPM). $lnISVola$ is the logarithmized standard deviation of the corresponding error term. $lnSize$ (the logarithmized market capitalization)and $lnBeme$ (the logarithmized book-to-market ratio) are computed as in Fama and French (1992). The $lnBidAsk$ is calculated as the logarithm of the average difference of the bid and ask price divided by the midquote using daily data of the previous six months as in Amihud and Mendelson (1986). Furthermore, $(Ret - Rf)_{t-1}$ is the lagged excess return (Jegadeesh, 1990), while $(Ret - Rf)_{t-1}$ denotes the cumulative excess return from month $t-12$ to $t-2$ capturing the momentum effect (Jegadeesh and Titman, 1993). $(IndRet - Rf)_{t-1}$ is the lagged excess industry return (Jegadeesh, 1990) and $(IndRet - Rf)_{t-12,t-2}$ denotes the cumulative industry-specific excess return from month $t-12$ to $t-2$ capturing the industry momentum effect. $lnStateDisp$ is the logarithmized number of distinct state names cited in firms' annual report, and the Herfindahl-Hirschman concentration measure is based on state citations (HHI). $CREA^{HQ}_{i,t-1}$ is the growth rate of economic activity of the headquarter state. The t-statistics computed with Newey and West (1987) standard errors are reported in the parentheses. The sample period is January 1995 through December 2011.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
$CREA_{t-1}$	2.167	2.085	2.003	1.982	2.018	2.103	1.892	1.969		1.887
	(2.71)	(4.49)	(5.10)	(5.03)	(4.66)	(4.55)	(3.67)	(3.58)		(5.42)
$CREA^{ExHQ}_{t-1}$									1.436	
									(3.22)	
$CREA^{HQ}_{t-1}$									0.830	
									(2.97)	
$Beta$		-0.002	-0.002	-0.002	-0.002	-0.002	-0.002	-0.002	-0.002	-0.002
		(-0.86)	(-0.89)	(-0.95)	(-0.88)	(-0.86)	(-0.86)	(-0.82)	(-0.86)	(-0.96)
$lnSize_{t-2}$		-0.002	-0.002	-0.002	-0.002	-0.002	-0.001	-0.001	-0.002	-0.001
		(-3.20)	(-3.14)	(-3.08)	(-2.64)	(-2.90)	(-1.45)	(-1.74)	(-3.11)	(-2.58)
$lnBeme$		0.001	0.002	0.002	0.001	0.001	0.002	0.002	0.001	0.002
		(1.43)	(1.75)	(1.72)	(1.54)	(1.47)	(2.12)	(1.88)	(1.50)	(1.98)
$lnISVola$		0.000	-0.000	-0.000	0.000	0.000	-0.000	-0.000	0.000	-0.001
		(0.08)	(-0.05)	(-0.10)	(0.07)	(0.07)	(-0.02)	(-0.00)	(0.09)	(-0.16)
$(Ret - Rf)_{t-12,t-2}$		0.001	0.001	0.000	0.001	0.001	0.005	0.005	0.001	0.000
		(0.35)	(0.25)	(0.10)	(0.33)	(0.35)	(2.99)	(3.03)	(0.38)	(0.07)
$(Ret - Rf)_{t-1}$		-0.047	-0.053	-0.048	-0.047	-0.047	-0.046	-0.046	-0.047	-0.054
		(-6.16)	(-7.15)	(-6.36)	(-6.23)	(-6.24)	(-5.43)	(-5.43)	(-6.19)	(-7.21)
$lnBidAsk$		-0.002	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001	-0.002	-0.001
		(-1.50)	(-1.34)	(-1.25)	(-1.34)	(-1.31)	(-0.66)	(-0.71)	(-1.43)	(-0.99)
$(IndRet - Rf)_{t-1}$			0.150							0.146
			(7.18)							(8.37)
$(IndRet - Rf)_{t-12,t-2}$				0.020						0.015
				(3.13)						(3.07)
$lnStateDisp$					-0.001		-0.002			-0.001
					(-0.73)		(-2.11)			(-0.91)
HHI						0.000		0.005		-0.001
						(0.14)		(1.89)		(-0.41)
Constant	0.005	0.020	0.016	0.015	0.021	0.019	0.015	0.011	0.020	0.014
	(1.02)	(1.09)	(0.94)	(0.84)	(1.11)	(1.01)	(0.75)	(0.54)	(1.07)	(0.79)
R2	0.002	0.062	0.065	0.065	0.063	0.063	0.063	0.063	0.062	0.069
Nobs	897709	799101	799101	799101	799101	799101	671526	671526	784276	799101

Table 4: Portfolio Sorts Based on $CREA^{\perp}$

$CREA^{\perp}$ is the sum of growth rates of regional economic activity across U.S. states (Crone and Clayton-Matthews, 2005) weighted by the corresponding citation shares of the states extracted from firms' 10-K reports orthogonalized by monthly cross-sectional regressions using the standard control variables (except common risk factor-related variables) as independent variables. The portfolios are sorted according to $CREA^{\perp}$ of the previous month in deciles. This table reports the average monthly portfolio return (in percent) and their corresponding t-statistics for both, equal-weighted and value-weighted portfolios. The last row reports the average monthly portfolio return (in percent) and the corresponding t-statistic of a portfolio that by goes long in the highest and short in the lowest decile. The sample period is January 1995 through December 2011.

Decile	$LS^{EW}_{CREA^{\perp}}$	t-statistic	$LS^{VW}_{CREA^{\perp}}$	t-statistic
1 (low)	0.677	1.72	0.390	1.04
2	0.845	1.94	0.468	1.33
3	0.954	2.07	0.698	2.03
4	0.896	1.93	0.632	1.87
5	0.829	1.73	0.430	1.25
6	0.887	1.95	0.434	1.28
7	1.047	2.29	0.672	1.88
8	1.050	2.27	0.622	1.66
9	1.107	2.47	0.649	1.70
10 (high)	1.361	3.30	0.881	2.37
$10-1$	0.684	5.64	0.491	2.01

Table 5: $CREA^{\perp}$ Portfolio Time-Series Regression

$CREA^{\perp}$ is the sum of growth rates of regional economic activity across U.S. states (Crone and Clayton-Matthews, 2005) weighted by the corresponding citation shares of the states extracted from firms' 10-K reports orthogonalized by monthly cross-sectional regressions using the standard control variables (except common risk factor-related variables) as independent variables. Firms are sorted according to $CREA^{\perp}$ of the previous month in deciles and LS is the return of the zero-cost strategy formed by going long in the top decile firms and going short in the bottom decile firms. The portfolio is updated monthly. The Table reports the coefficient estimates (Jensen's alpha and regression coefficients) of the following time-series regression model:

$$LS_{CREA^{\perp},t} = \alpha + \mathbf{X_t'}\beta + \varepsilon$$

where LS is the long-short portfolio formed according to $CREA^{\perp}_{t-1}$, and $\mathbf{X_t'}$ is a set of the five tradable common risk factors $MKT - Rf$, SMB, HML, UMD, and LIQ (Fama and French, 1993; Carhart, 1997; Pástor and Stambaugh, 2003). Panel A shows the results for the equal-weighted long-short portfolio, $LS^{EW}_{CREA^{\perp},t}$, and Panel B shows the results for the value-weighted counterpart, $LS^{EW}_{CREA^{\perp},t}$. The t-statistics are reported in the parentheses. The sample period is January 1995 through December 2011.

Panel A: Equal-weighted Portfolio

	(1)	(2)	(3)	(4)	(5)
α	0.684	0.669	0.704	0.681	0.631
	(5.64)	(5.49)	(5.92)	(5.68)	(5.26)
$MKT - Rf$		0.029	0.001	0.014	0.006
		(1.13)	(0.06)	(0.52)	(0.21)
SMB			0.035	0.030	0.032
			(1.00)	(0.85)	(0.91)
HML			-0.117	-0.108	-0.102
			(-3.19)	(-2.88)	(-2.75)
UMD				0.031	0.027
				(1.36)	(1.20)
LIQ					0.073
					(2.56)
R2	0.000	0.006	0.075	0.083	0.113
Nobs	204	204	204	204	204

Panel B: Value-weighted Portfolio

	(1)	(2)	(3)	(4)	(5)
α	0.491	0.483	0.488	0.448	0.493
	(2.01)	(1.96)	(1.96)	(1.78)	(1.93)
$MKT - Rf$		0.016	0.009	0.032	0.040
		(0.30)	(0.17)	(0.55)	(0.68)
SMB			0.014	0.005	0.003
			(0.19)	(0.06)	(0.04)
HML			-0.020	-0.004	-0.009
			(-0.26)	(-0.04)	(-0.12)
UMD				0.054	0.057
				(1.14)	(1.21)
LIQ					-0.066
					(-1.10)
R2	0.000	0.000	0.001	0.008	0.014
Nobs	204	204	204	204	204

Table 6: Regional Economic Activity and Cross-predictability of Stock Returns

The Table reports the average cross-sectional regression coefficients using the Fama and MacBeth (1973) framework:

$$Ret_{i,t} - Rf_t = \alpha_t + \beta_t \cdot PREA_{i,t-2} + \mathbf{x'_{i,t-1}b_t} + \varepsilon_{i,t},$$

where $Ret_{i,t}$ is return of stock i in month t and Rf_t is the monthly yield on 30-day Treasury bills, and $\mathbf{x'_i}$ is a vector of other firm characteristics. $PREA_{i,t-2}$ is the sum of predicted growth rates of regional economic activity across U.S. states (Crone and Clayton-Matthews, 2005) weighted by the corresponding citation share of the states extracted from firms' 10-K reports. Other control variables are: $Beta$ is the stock-specific market beta calculated using rolling regressions with daily returns of the past 252 days and the market return as the explanatory variable (CAPM). $lnISVola$ is the logarithmized standard deviation of the corresponding error term. $lnSize$ (the logarithmized market capitalization)and $lnBeme$ (the logarithmized book-to-market ratio) are computed as in Fama and French (1992). The $lnBidAsk$ is calculated as the logarithm of the average difference of the bid and ask price divided by the midquote using daily data of the previous six months as in Amihud and Mendelson (1986). Furthermore, $(Ret - Rf)_{t-1}$ is the lagged excess return (Jegadeesh, 1990), while $(Ret - Rf)_{t-1}$ denotes the cumulative excess return from month $t-12$ to $t-2$ capturing the momentum effect (Jegadeesh and Titman, 1993). $(IndRet - Rf)_{t-1}$ is the lagged excess industry return (Jegadeesh, 1990) and $(IndRet - Rf)_{t-12,t-2}$ denotes the cumulative industry-specific excess return from month $t-12$ to $t-2$ capturing the industry momentum effect. $lnStateDisp$ is the logarithmized number of distinct state names cited in firms' annual report, and the Herfindahl-Hirschman concentration measure is based on state citations (HHI). $PREA^{HQ}_{i,t-2}$ is the predicted growth rate of economic activity of the headquarter state. The t-statistics computed with Newey and West (1987) standard errors are reported in the parenthesis. The sample period is January 1995 through December 2011.

	(1)	(2)	(3)	(4)	(5)	(6)
$PREA_{t-2}$	0.415	0.391	0.405	0.417		0.379
	(4.42)	(5.28)	(4.60)	(4.52)		(5.40)
$PREA^{ExHQ}_{t-2}$					0.268	
					(3.34)	
$PREA^{HQ}_{t-2}$					0.189	
					(4.20)	
$Beta$	-0.002	-0.002	-0.002	-0.002	-0.002	-0.002
	(-0.88)	(-0.93)	(-0.89)	(-0.88)	(-0.87)	(-0.97)
$lnSize_{t-2}$	-0.002	-0.002	-0.002	-0.002	-0.002	-0.001
	(-3.21)	(-3.10)	(-2.64)	(-2.91)	(-3.12)	(-2.58)
$lnBeme$	0.001	0.002	0.001	0.001	0.001	0.002
	(1.41)	(1.89)	(1.53)	(1.46)	(1.48)	(1.97)
$lnISVola$	0.000	-0.001	0.000	0.000	0.000	-0.001
	(0.06)	(-0.15)	(0.06)	(0.06)	(0.08)	(-0.17)
$(Ret - Rf)_{t-12,t-2}$	0.001	0.000	0.001	0.001	0.001	0.000
	(0.34)	(0.07)	(0.33)	(0.34)	(0.38)	(0.06)
$(Ret - Rf)_{t-1}$	-0.047	-0.053	-0.047	-0.047	-0.047	-0.054
	(-6.17)	(-7.18)	(-6.24)	(-6.25)	(-6.21)	(-7.21)
$lnBidAsk$	-0.002	-0.001	-0.001	-0.001	-0.002	-0.001
	(-1.49)	(-1.18)	(-1.33)	(-1.31)	(-1.42)	(-0.98)
$(IndRet - Rf)_{t-1}$		0.152				0.146
		(8.51)				(8.35)
$(IndRet - Rf)_{t-12,t-2}$		0.015				0.015
		(3.06)				(3.05)
$lnStateDisp$			-0.001			-0.001
			(-0.75)			(-0.91)
HHI				0.001		-0.001
				(0.18)		(-0.38)
Constant	0.019	0.012	0.020	0.019	0.019	0.013
	(1.05)	(0.73)	(1.08)	(0.97)	(1.01)	(0.75)
R2	0.062	0.068	0.063	0.063	0.062	0.069
Nobs	799101	799101	799101	799101	784276	799101

Table 7: $PREA^{\perp}$ Portfolio Time-Series Regression

$PREA^{\perp}$ is the sum of growth rates of predicted regional economic activity across U.S. states (Crone and Clayton-Matthews, 2005) weighted by the corresponding citation shares of the states extracted from firms' 10-K reports orthogonalized by monthly cross-sectional regressions using the standard control variables (except common risk factor-related variables) as independent variables. Firms are sorted according to $PREA^{\perp}$ lagged by two months in deciles and LS is the return of the zero-cost strategy formed by going long in the top decile firms and going short in the bottom decile firms. The portfolio is updated monthly. The Table reports the coefficient estimates (Jensen's alpha and regression coefficients) of the following regression model:

$$LS_{PREA^{\perp},t} = \alpha + \mathbf{X_t'}\beta + \varepsilon$$

where LS is the long-short portfolio formed according to $PREA^{\perp}_{t-2}$, and $\mathbf{X_t'}$ is a set of the five tradable common risk factors $MKT - Rf$, SMB, HML, UMD, and LIQ (Fama and French, 1993; Carhart, 1997; Pástor and Stambaugh, 2003). Panel A shows the results for the equal-weighted long-short portfolio, $LS^{EW}_{PREA^{\perp},t}$, and Panel B shows the results for the value-weighted counterpart, $LS^{EW}_{PREA^{\perp},t}$. The t-statistics are reported in the parentheses. The sample period is January 1995 through December 2011.

Panel A: Equal-weighted Portfolio

	(1)	(2)	(3)	(4)	(5)
α	0.560	0.555	0.570	0.560	0.527
	(5.24)	(5.15)	(5.37)	(5.21)	(4.87)
$MKT - Rf$		0.009	-0.015	-0.009	-0.015
		(0.38)	(-0.64)	(-0.36)	(-0.60)
SMB			0.060	0.058	0.059
			(1.93)	(1.85)	(1.89)
HML			-0.065	-0.060	-0.056
			(-1.97)	(-1.81)	(-1.69)
UMD				0.014	0.011
				(0.68)	(0.55)
LIQ					0.048
					(1.88)
R2	0.000	0.001	0.054	0.056	0.072
Nobs	204	204	204	204	204

Panel B: Value-weighted Portfolio

	(1)	(2)	(3)	(4)	(5)
α	0.671	0.673	0.717	0.601	0.688
	(2.69)	(2.67)	(2.85)	(2.43)	(2.76)
$MKT - Rf$		-0.004	-0.040	0.023	0.038
		(-0.07)	(-0.74)	(0.41)	(0.68)
SMB			0.055	0.029	0.026
			(0.74)	(0.40)	(0.36)
HML			-0.148	-0.100	-0.111
			(-1.91)	(-1.30)	(-1.45)
UMD				0.153	0.159
				(3.30)	(3.47)
LIQ					-0.126
					(-2.15)
R2	0.000	0.000	0.028	0.078	0.099
Nobs	204	204	204	204	204

Table 8: Profitability and Regional Economic Activity

The Table reports the regression coefficients using two-way fixed effect estimation with the following specification:

$$Profitability_{i,q} = \alpha + \beta_1 Profitability_{i,q-1} + \beta_2 PREA^{\perp}_{i,q-1} + \mathbf{y'_{i,q-1}b} + \mu_i + \eta_q + \varepsilon_{i,q}$$

where firm profitability is measured by sales scaled by assets (SOA) and operating income before depreciation scaled by book value of equity (ROE), both measured in percentage. $PREA_{i,q-1}$ is the sum of growth rates of predicted regional economic activity across U.S. states (Crone and Clayton-Matthews, 2005) weighted by the citation shares of the corresponding states at the end of the last quarter. $PREA^{\perp}_{i,q-1}$ is $PREA_{i,q-1}$ orthogonalized by standard control variables, except the variables related to common risk factors. $\mathbf{y'_{i,q-1}}$ denotes the vector of control variables, while μ_i and η_q denote the firm fixed effect and quarter fixed effect, respectively. The standard control variables include lagged profitability, logarithm of market capitalization ($lnSize_{i,q-1}$), logarithm of book-to-market ratio ($lnBeme_{i,q-1}$), change in net operating assets ($DNOA_{i,q-1}$), dividend yield ($Div_{i,q-1}$), and a dummy variable that takes a value one if the firm reported a loss in the last quarter ($LOSS_{i,q-1}$) (see for instance Fama and French, 1995, 2000; Richardson, Sloan, Soliman, and Tuna, 2005). Additionally, I include firms' excess stock return ($Ret_{i,q-1} - Rf_{q-1}$) and excess industry return ($IndRet_{i,q-1} - Rf_{q-1}$) of the past quarter. $PREA^{\perp}_{i,q-1} < 0$ is a simple and alternative proxy that takes a value of 1 if the de-meaned $PREA^{\perp}_{i,q-1}$ proxy is less than 0. The t-statistics are calculated from two-way clustered standard errors and reported in the parentheses. The sample period is January 1995 through December 2011.

Dependent Variable	SOA_q				ROE_q			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
SOA_{iq-1}	0.643	0.614	0.614	0.614				
	(45.27)	(38.76)	(38.77)	(38.76)				
ROE_{iq-1}					0.479	0.416	0.416	0.416
					(37.94)	(30.58)	(30.65)	(30.65)
$PREA^{\perp}_{iq-1}$	19.130	14.636	14.662		7.177	3.722	3.774	
	(5.72)	(4.01)	(4.03)		(4.80)	(2.52)	(2.57)	
$PREA^{\perp}_{iq-1} < 0$				-0.166				-0.035
				(-4.16)				(-2.27)
$lnSize_{iq-1}$		-1.380	-1.382	-1.382		-0.167	-0.168	-0.168
		(-16.39)	(-16.51)	(-16.51)		(-4.57)	(-4.60)	(-4.61)
$lnBeme_{iq-1}$		-1.734	-1.724	-1.725		-0.651	-0.647	-0.647
		(-17.55)	(-17.24)	(-17.24)		(-11.71)	(-11.37)	(-11.37)
$DNOA_{iq-1}$		0.115	0.117	0.116		-0.035	-0.033	-0.033
		(0.72)	(0.73)	(0.73)		(-0.22)	(-0.21)	(-0.21)
$LOSS_{iq-1}$		1.151	1.154	1.154		-0.161	-0.161	-0.161
		(10.48)	(10.52)	(10.53)		(-3.34)	(-3.32)	(-3.33)
$Div_{i,q-1}$		0.017	0.017	0.017		0.070	0.070	0.070
		(0.18)	(0.18)	(0.18)		(3.00)	(3.01)	(3.01)
$Ret_{i,q-1} - Rf_{q-1}$			0.063	0.064			0.004	0.004
			(0.87)	(0.87)			(0.14)	(0.15)
$IndRet_{i,q-1} - Rf_{q-1}$			0.106	0.108			0.216	0.216
			(0.25)	(0.25)			(1.25)	(1.25)
R2	0.921	0.925	0.925	0.925	0.527	0.609	0.609	0.609

Table 9: Predictability of Returns and Difficult-to-arbitrage Firms

This table reports the relevant average cross-sectional regression coefficients using the Fama and MacBeth (1973) framework:

$$Ret_{i,t} - Rf_t = \alpha_t + \beta_1 PREA_{i,t-2} + \beta_2 PREA_{i,t-2} \cdot M_{i,t-2} + \beta_3 M_{i,t-2} + \mathbf{x_i'b} + \varepsilon_{i,t},$$

where $Ret_{i,t}$ is return of stock i in month t, and Rf_t is the monthly yield on 30-day Treasury bills, and $\mathbf{x_i'}$ is a vector of other firm characteristics. $PREA_{i,t-2}$ is the sum of predicted growth rates of regional economic activity across U.S. states (Crone and Clayton-Matthews, 2005) weighted by the corresponding citation shares of the states extracted from firms' 10-K reports. $M_{i,t-2}$ is a firm characteristic defined according to the specification below. Other control variables include: *Beta* is the stock-specific market beta calculated using rolling regressions with daily returns of the past 252 days and the market return as the explanatory variable (CAPM). $lnISVola$ is the logarithmized standard deviation of the corresponding error term. *lnSize* (the logarithmized market capitalization)and *lnBeme* (the logarithmized book-to-market ratio) are computed as in Fama and French (1992). The *lnBidAsk* is calculated as the logarithm of the average difference of the bid and ask price divided by the midquote using daily data of the previous six months as in Amihud and Mendelson (1986). Furthermore, $(Ret - Rf)_{t-1}$ is the lagged excess return (Jegadeesh, 1990), while $(Ret - Rf)_{t-1}$ denotes the cumulative excess return from month $t-12$ to $t-2$ capturing the momentum effect (Jegadeesh and Titman, 1993). $(IndRet-Rf)_{t-1}$ is the lagged excess industry return (Jegadeesh, 1990) and $(IndRet-Rf)_{t-12,t-2}$ denotes the cumulative industry-specific excess return from month $t-12$ to $t-2$ capturing the industry momentum effect. *lnStateDisp* is the logarithmized number of distinct state names cited in firms' annual report, and the Herfindahl-Hirschman concentration measure is based on state citations (HHI). $CREA^{HQ}_{i,t-1}$ is the growth rate of economic activity of the headquarter state. This Table reports only the regression coefficients β_1 and β_2. The t-statistics computed with Newey and West (1987) standard errors are reported in the parentheses. The sample period is January 1995 through December 2011.

	(1)		(2)		(3)	
Independent Variable	Coeff.	t-statistic	Coeff.	t-statistic	Coeff.	t-statistic
$PREA_{t-2}$	1.264	(2.19)	0.960	(3.49)	1.039	(2.66)
$PREA_{t-2} \times lnISVola_{t-2}$	0.272	(1.82)				
$PREA_{t-2} \times lnBidAsk_{t-2}$			0.151	(2.47)		
$PREA_{t-2} \times lnSize_{t-2}$					-0.058	(-2.01)
Stand. Controls	Yes		Yes		Yes	
Avg. R^2	0.07		0.07		0.07	
Nobs	799105		799105		799105	

A Appendix: Additional Figures and Tables

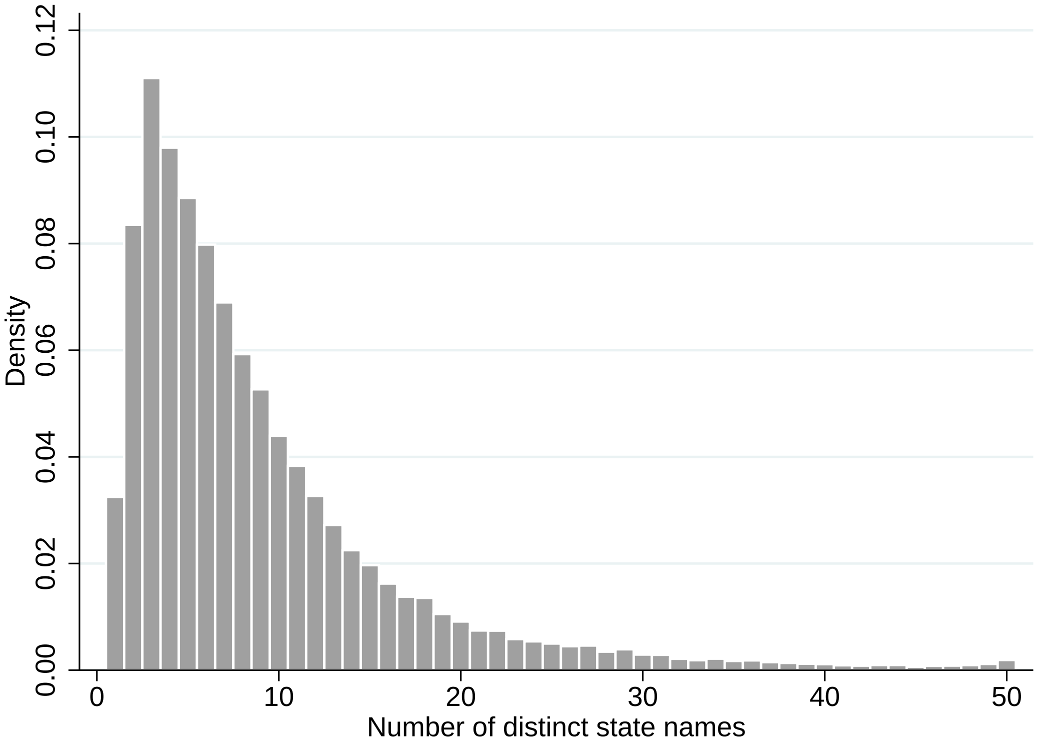


Figure A.1: Histogram of Distinct Number of State Names

This figure plots the histogram of state dispersion (number of distinct U.S. states mentioned in the 10-K filings) across all firm-year observations. 10-K filings that do not mention any state are excluded. The sample period is January 1995 through December 2011.

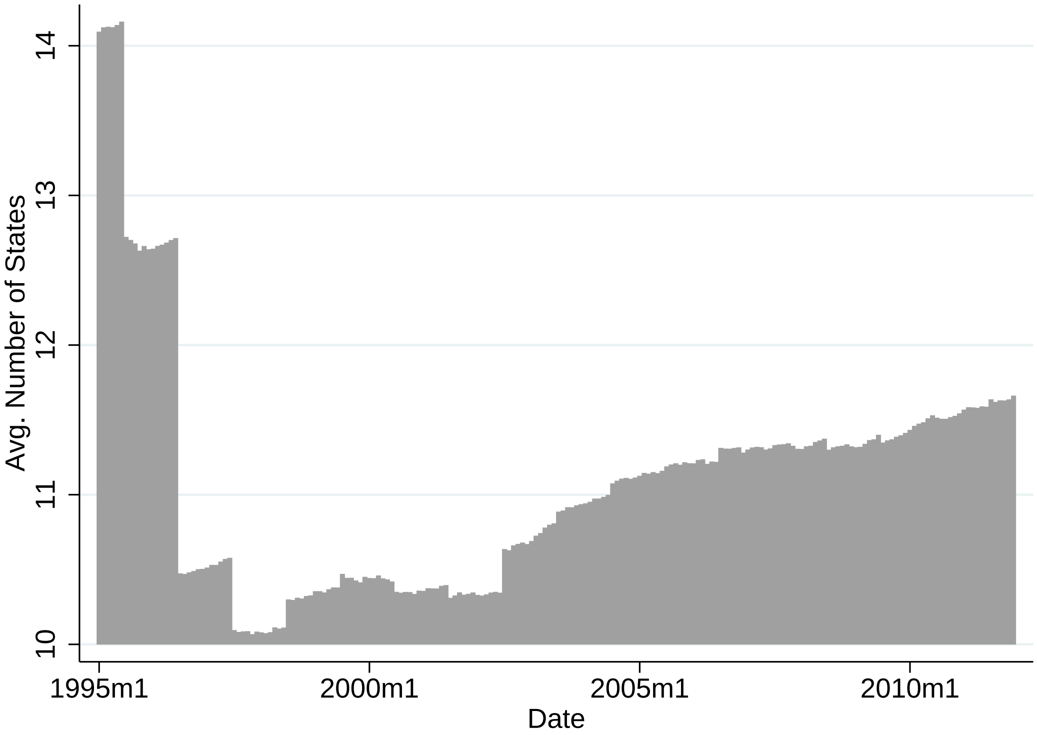


Figure A.2: Average Number of States over the Sample Period

This figure plots the cross-sectional average of state dispersion (number of distinct U.S. states mentioned in the 10-K filings) on monthly frequency. Observations from July of year τ to June of year $\tau + 1$ are assigned to the annual reports (state dispersion) of year $\tau - 1$. The sample period is January 1995 through December 2011.

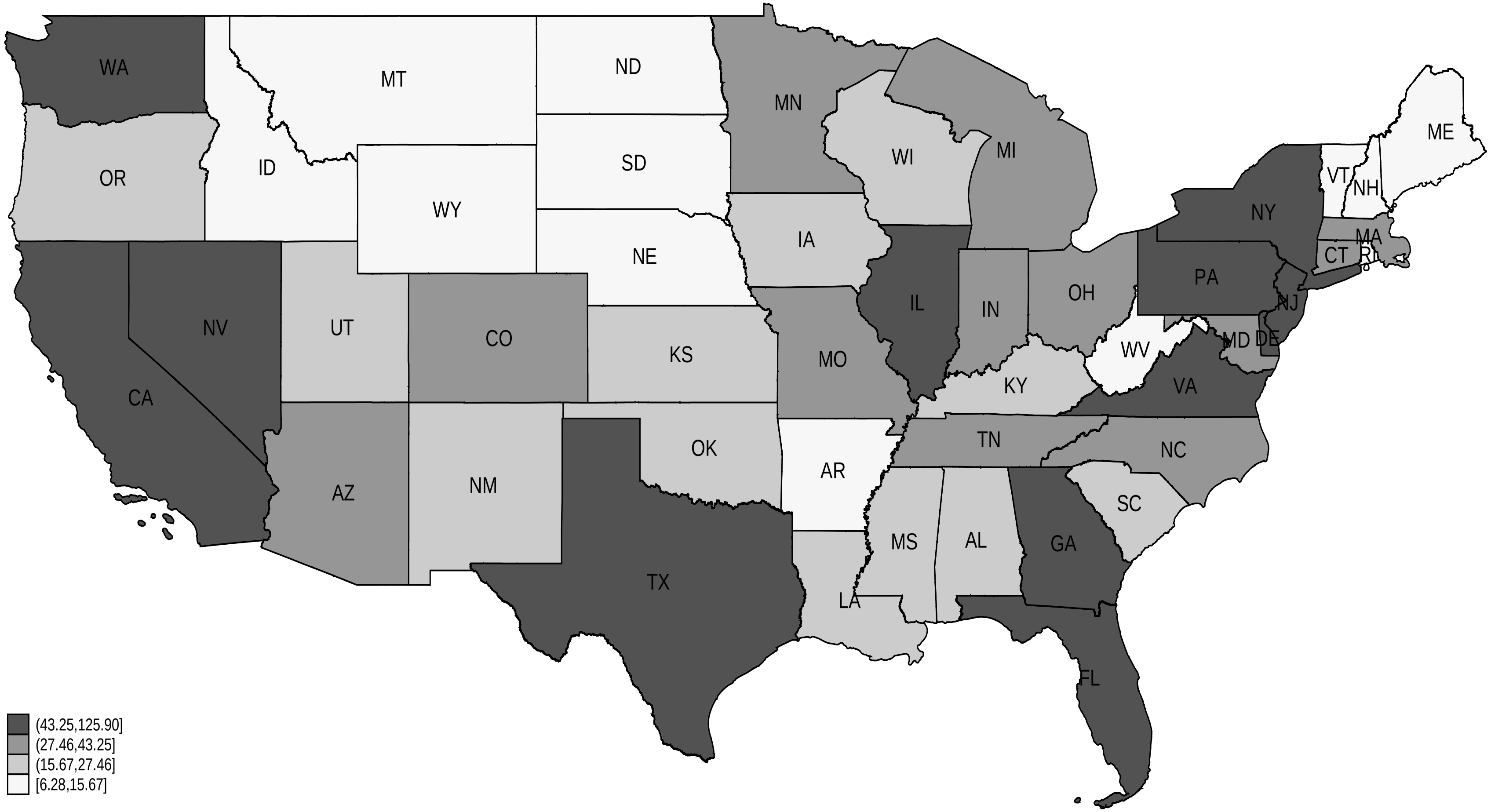

WA
OR
CA
NV
ID
MT
WY
UT
AZ
CO
NM
ND
SD
NE
KS
OK
TX
MN
IA
MO
AR
LA
WI
IL
MI
IN
OH
KY
TN
MS
AL
GA
FL
SC
NC
VA
WV
PA
NY
MD
DE
NJ
CT
RI
MA
VT
NH
ME
(43.25,125.90]
(27.46,43.25]
(15.67,27.46]
[6.28,15.67]

Figure A.3: The U.S. State Citation Counts

This figure graphically illustrates the economic relevance of U.S. states proxied by the the number of state citations (in thousands) across all 10-K filings reported by publicly traded firms. Darker states have a higher number of citations across all firm-year observations than lighter states. The top three regions according to citations are Delaware, New York, and California while North Dakota, South Dakota, and Maine are the least cited states. The sample period is January 1995 through December 2011.

Table A.1: Robustness Tests

This table reports the average cross-sectional regression coefficients using the Fama and MacBeth (1973) framework analyzing alternative specifications in the robustness checks:

$$Ret_{i,t} - Rf_t = \alpha_t + \beta_t \cdot CREA_{i,t-1} + \mathbf{x'_{i,t-1}b_t} + \varepsilon_{i,t},$$

where $CREA^{EX}_{i,t-1}$ is the sum of growth rates of regional economic activity across U.S. states (Crone and Clayton-Matthews, 2005) weighted by the corresponding citation shares of the states extracted from firms' 10-K reports excluding Delaware and Washington. $CREA^{EW}_{i,t-1}$ is the sum of growth rates of regional economic activity across U.S. states (Crone and Clayton-Matthews, 2005) weighted equally. In specification 1 and 2, the dependent variable is adjusted by the 49 Fama-French industry returns and the DGTW (1997) portfolio returns, respectively. In specification 3, 4, and 5 I exclude stocks priced less than \$1, \$5, and financial firms, respectively. Specification 6 and 7 use the alternative $CREA$ proxies mentioned above. All other variables are defined in Table 3. The t-statistics computed with Newey and West (1987) standard errors are reported in the parentheses. The sample period is January 1995 through December 2011.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$CREA_{t-1}$	1.499	1.866	1.729	1.640	1.954		
	(5.31)	(5.34)	(5.30)	(5.73)	(4.82)		
$CREA^{EXDeWa}_{t-1}$						1.522	
						(4.21)	
$CREA^{EW}_{t-1}$							2.382
							(4.36)
$Beta$	-0.001	-0.001	-0.001	-0.002	-0.002	-0.002	-0.002
	(-0.78)	(-0.82)	(-0.51)	(-0.79)	(-1.03)	(-0.97)	(-0.95)
$lnSize_{t-2}$	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001
	(-2.87)	(-1.89)	(-2.15)	(-1.24)	(-2.55)	(-2.60)	(-2.58)
$lnBeme$	0.002	0.000	0.002	0.001	0.002	0.002	0.002
	(2.99)	(0.02)	(2.19)	(1.41)	(2.38)	(1.98)	(1.98)
$lnISVola$	-0.001	-0.000	-0.003	-0.003	0.000	-0.001	-0.001
	(-0.40)	(-0.14)	(-0.86)	(-1.31)	(0.02)	(-0.17)	(-0.15)
$(Ret - Rf)_{t-12,t-2}$	0.000	-0.001	0.003	0.004	-0.000	0.000	0.000
	(0.02)	(-0.41)	(0.82)	(1.21)	(-0.07)	(0.08)	(0.08)
$(Ret - Rf)_{t-1}$	-0.053	-0.054	-0.036	-0.029	-0.051	-0.054	-0.053
	(-7.19)	(-7.40)	(-6.57)	(-5.44)	(-6.83)	(-7.22)	(-7.20)
$lnBidAsk$	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001
	(-0.71)	(-0.75)	(-1.20)	(-1.58)	(-0.74)	(-0.97)	(-0.99)
$(IndRet - Rf)_{t-1}$	-0.006	0.143	0.130	0.099	0.126	0.146	0.146
	(-0.48)	(8.31)	(7.48)	(5.84)	(7.57)	(8.37)	(8.40)
$(IndRet - Rf)_{t-12,t-2}$	-0.002	0.013	0.013	0.008	0.011	0.015	0.015
	(-0.58)	(2.88)	(2.93)	(1.86)	(2.55)	(3.07)	(3.10)
$lnStateDisp$	-0.000	-0.000	-0.001	-0.001	-0.000	-0.001	-0.001
	(-0.30)	(-0.61)	(-1.33)	(-0.98)	(-0.43)	(-0.82)	(-0.83)
HHI	0.001	-0.000	-0.002	-0.002	0.002	-0.001	-0.001
	(0.46)	(-0.12)	(-0.76)	(-0.78)	(0.94)	(-0.37)	(-0.38)
Constant	0.006	0.000	0.000	-0.007	0.015	0.014	0.014
	(0.49)	(0.03)	(0.02)	(-0.58)	(0.84)	(0.81)	(0.77)
R2	0.043	0.047	0.072	0.086	0.065	0.069	0.069
Nobs	799101	793655	768451	608434	712483	798168	799101