Dissanayake, Ruchith
(2021)
Geographic distribution of firms and expected stock returns.
Journal of Economic Dynamics and Control, *133*, Article number: 104267.

https://doi.org/10.1016/j.jedc.2021.104267

Geographic Distribution of Firms and Expected Stock Returns*

Ruchith Dissanayake†

*I would like to thank the Editor Xue-Zhong (Tony) He, the Associate Editor and the referee for helpful comments and suggestions. I thank Felipe Aguerrevere, Adam Clements, Mark Doolan, Adrian Fernandez-Perez, Neal Galpin, Lorenzo Garlappi, Patrick Grüning, Timothy Johnson, Hae Won Jung, Aditya Kaul, Egor Matveyev, Dimitris Papanikolaou, Marcello Pericoli, Pascalis Raimondos, Robert Ready, Stephen Thiele, Steven Riddiough, Alexander Vadilyev, Parianen Veeren, Akiko Watanabe, Masahiro Watanabe, Yanhui Wu, Hui Xu, Takeshi Yamada, Huizhong Zhang, Yu Zhang, Min Zhu, Qiaoqiao Zhu, and session participants at the American Finance Association Meetings, Auckland Finance Meetings, Australasian Finance and Banking Conference, and Midwest Finance Association Conference, and seminar participants at the Australian National University, Queensland University of Technology, and University of Melbourne for helpful comments and discussions.

†Queensland University of Technology, QUT Business School, Brisbane, Australia. E-mail: r.dissanayake@qut.edu.au, URL: http://www.rdissanayake.com/.

Geographic Distribution of Firms and Expected Stock Returns

ABSTRACT

I examine the effects of geographic distribution of firms on the expected stock returns. Information spillovers and coordinated actions by interacting managers increase the cyclicality of wages in agglomerated industries compared to dispersed industries. Consequently, geographic agglomeration provides firms a "natural hedge" against aggregate shocks. In contrast, geographically dispersed firms have higher exposure to aggregate shocks. A portfolio that goes long on geographically dispersed industries minus agglomerated industries – the GDMA portfolio – captures aggregate shocks. Stocks that co-vary closely with the GDMA portfolio returns earn higher expected returns. In the time-series, the premium is more pronounced during recessions when investors shrink from risk. In the cross-section, the premium is more pronounced among low profitable firms that are more vulnerable to adverse shocks.

Keywords: geographic distribution, expected stock returns, hedge factor, GDMA portfolio.

JEL classification: G12

When an industry has thus chosen a locality for itself, it is likely to stay there long: so great are the advantages which people following the same skilled trade get from near neighbourhood to one another. — Marshall, 1890

1 Introduction

Firm production is quite concentrated in space. The high concentration of advertising industry in Manhattan and auto industry in Detroit are well-known examples of geographic agglomeration of firms. Firms geographically concentrate near one another for many reasons. The primary advantage of agglomeration is to pool the demand for specialized labor, first emphasized by Marshall (1890). Other positive externalities of agglomeration include increasing returns to scale, information sharing, and intellectual spillovers.[1] Although knowledge spillovers and input sharing are important benefits of agglomeration, the evidence is strongest for labor market pooling (Rosenthal and Strange (2001); Dumais, Ellison, and Glaeser (2002)). The literature documents benefits of agglomeration on aggregate consumption growth (Davis, Fisher, and Whited (2014)), firm productivity (Henderson (2003)), and wages (Glaeser and Mare (2001); Amiti and Cameron (2007); Rosenthal and Strange (2008)). However, to the best of my knowledge, none of the prior studies have examined the time-varying risks associated with geographic distribution of firms and the implications on the cross-section of stock returns. The goal of this paper is to fill this gap in the literature.

The geographic agglomeration of firms allows managers the opportunity to establish informal local labor market networks to hire workers with better abilities and negotiate better wage contracts. Montgomery (1991) theorizes that firms learn about a potential worker's ability if the firm employs individuals from the potential worker's network. Furthermore, in equilibrium, individuals are more likely to receive and accept wage offers from businesses that employ others in their network. Underlying most network models is some form of information imperfection in which net-

[1] These sources are not mutually exclusive. Separating the sources of agglomeration of firms is beyond the scope of this study.

works serve, at least partially, to mitigate these imperfections. There is a large body of work that use employer-employee micro data at the establishment level to show the presence and importance of labor market networks based on proximity (e.g., Rees (1966); Bayer, Ross, and Topa (2008); Hellerstein, McInerney, and Neumark (2011); Hellerstein, Kutzbach, and Neumark (2014); Saygin, Weber, and Weynandt (2021); Hensvik and Skans (2016)). Managers use these social networks to attract workers with better qualities in hard-to-observe dimensions (Hensvik and Skans (2016)). Also, there is evidence that displaced workers are more likely to become re-employed at a firm in their geographical network that employs former co-workers of the displaced worker (Saygin, Weber, and Weynandt (2021)).

These local labor market networks provide firms in geographically agglomerated industries the opportunity to better negotiate wage contracts during market downturns. During low growth periods, workers are reluctant to move outside their labor market networks and have lower expectations about outside opportunities. This allows managers in agglomerated industries to better negotiate wage contracts with workers during recessions. Geographic proximity of firms increases the correlated actions among interacting managers and induces amplified responses to aggregate shocks (Guiso and Schivardi (2007)).[2] This amplification effect is supported by information spillover models; agents face a common problem in an uncertain environment and each agent holds private information, which can be inferred from other agents' actions (Banerjee (1992); Bikhchandani et al. (1992)). When one firm negotiates lower wage contracts, the information revealed triggers further actions, and start a self-reinforcing process that prompts other managers in the local labor market networks to adjust wages within a short time span. This wage adjustment provides firms in agglomerated industries a natural hedge (procyclicality of wages) against aggregate shocks. Since fewer people are making long-distance moves in the U.S. (Molloy, Smith, and Wozniak (2011)), the effects of geographical networks are likely to become even stronger in the future.

The procyclicality of wages in agglomerated industries, relative to dispersed industries, have

[2]See Glaeser and Scheinkman (2000) for a survey on the literature on the effects of social interaction on individual decision-making.

direct consequences on firm valuations. For firms in agglomerated industries, the lower average wage costs during market downturns reduces the *ex-ante* covariance between firm's cash flow and business conditions, which translates to lower cost of equity. Firms in dispersed industries, on the other hand, have relatively stickier wages. Consequently, firms in dispersed industries face a greater drop in cash flow following an adverse systematic shock, which generally leads to a recession. The lower valuations increase the expected returns for firms in dispersed industries relative to firms in agglomerated industries.[3]

Sophisticated investors can obtain information about firm geography through costly sources such as analyst reports. I use an indirect approach to identify industry agglomeration that is both unbiased and systematic; I use the Ellison and Glaeser (1997) (EG) index of geographic agglomeration in manufacturing industries. EG index is independent of the number of plants and of their distribution and controls for the industrial concentration. As shown by Dumais, Ellison, and Glaeser (2002), geographic concentrations for industries are strikingly stable over time. I map the EG index with the universe of stocks from the CRSP dataset. An appealing aspect of this approach is that the classification of firms is based on a highly researched and economically meaningful characteristic.[4]

An important caveat is that the entire system of informal labor market networks is not observable to the econometrician. However, I can infer the consequences of the networks, without observing the actual network connections, by examining changes in the aggregate variables of interest (i.e., wages, cash flow, and expected returns). It is important to emphasize that the geographically local informal network structure need not be interpreted literally. Instead, it can also describe connectedness among firms in terms of social attributes, research interests, compatibility

[3]Investors demand a premium to hold geographically dispersed firms, relative to agglomerated firms, since they carry significantly more cash flow risk following adverse systematic shocks (e.g., negative demand shocks). Geographically agglomerated firms, in contrast, can better adjust wages and reduce the impact of adverse shocks on cash flow.

[4]In fact, the theoretical literature examining geographic agglomeration dates to the seminal work by Marshall (1890). Other examples of notable theory include Krugman (1991), Holmes (1998), and Arzaghi and Henderson (2008).

of R&D programs, etc. I emphasize the labor market network since it is well established in the literature. Any other attributes related to economic geography that both co-vary over the business cycle and provide a natural hedge against aggregate shocks strengthen the arguments presented in this paper.

This paper makes three contributions. *First*, I show that cash flows are more procyclical in geographically dispersed firms, and hence more exposed to systematic risk, than that in agglomerated firms. Specifically, using manufacturing industry data, I test whether wages are more procyclical in agglomerated industries than that in dispersed industries. I control for unobserved heterogeneity by including industry fixed effects estimates, which remove any industry fixed characteristics. The regressions also include year fixed effects, which control for any changes in the aggregate investment opportunities over time. I find that the average wages are more procyclical (countercyclical) in geographically agglomerated (geographically dispersed) industries. This implies that firms in agglomerated industries with formal and informal labor market networks can better adjust wages during market downturns than firms in geographically dispersed industries.

The cyclicality of wages leads to a reduction in cash flow risk for firms in geographically agglomerated industries. For example, a negative demand shock, which generally leads to an economic downturn, can reduce a firm's future cash flow. If a firm can reduce labor costs following the negative shock, then some of the adverse effects on cash flow are abated. Following a negative shock, firms in geographically agglomerated industries can better adjust wages and reduce the impact of the adverse effects on cash flow. This wage adjustment process in geographically agglomerated firms acts as a "natural hedge" against adverse systematic shocks.[5]

This natural hedge in geographically agglomerated firms should be reflected in a firm's cash flow over the business cycle. To test this conjecture, I construct a measure of firm-level geographic risk. I construct a "hedge factor," which is a mimicking portfolio that goes long on geographically dispersed industries minus agglomerated industries based on the EG index classification (*GDMA*

[5] I refer to this wage adjustment process that reduces cash flow risk following adverse shocks as a "natural hedge" to distinguish it from a financial hedge (e.g., financial derivatives that hedge cash flow risk following adverse shocks).

portfolio).[6] The *GDMA* portfolio captures aggregate shocks that are naturally hedged in geographically agglomerated firms but not in geographically dispersed firms. To measure the geographic risk at the firm level, at each point in time, I compute a stock's sensitivity to the *GDMA* portfolio returns (β_{GDMA}) using a 60-month window. Firms with low β_{GDMA} have low exposure to geographic risk since geographic agglomeration provides a natural hedge against aggregate shocks. In contrast, firms with high β_{GDMA} have high exposure to geographic risk.[7]

Using panel regressions on a sample of publicly listed firms in Compustat, I show that earnings - measured using ROA and cash flow to assets - are more cyclical in firms with high $\beta'_{GDMA}s$ controlling for both firm and year fixed effects. I also estimate the parameters for (1) the sample of manufacturing firms for which the geographic risk can be computed at the industry level *via* EG index and (2) the sample of firms for which the geographic risk can be computed using only the β_{GDMA}. The results are similar for both sub-samples; firms with high $\beta'_{GDMA}s$ have more cyclical earnings. Hence, during recessions when cash flow is needed the most, cash flows are significantly lower for firms with high geographic risk (high β_{GDMA}) than for firms with low geographic risk (low β_{GDMA}).

Second, I show that the expected stock returns are higher in geographically dispersed firms than in agglomerated firms. Since firms in geographically agglomerated industries are better naturally hedged against aggregate shocks, the expected returns are lower for agglomerated firms than for geographically dispersed firms; the investor is willing to pay a higher price, hence lower expected return, to hold stocks in geographically agglomerated industries than for stocks in geographically dispersed industries.

The benefit of the natural hedge in agglomerated firms is heterogenous across-time and across-firms. In the time-series, the marginal benefit of hedging is greater during recessions - periods of low demand and productivity - than during economic expansions. The natural hedge against aggregate shocks in agglomerated firms helps curb the cash flow risk during times of low growth.

[6] Note that the Ellison and Glaeser index can only be computed for manufacturing firms at the industry level.

[7] I refer to the exposure to the GDMA portfolio (β_{GDMA}) as geographic risk for brevity.

Firms in geographically dispersed industries, in contrast, are fully exposed to aggregate shocks, and hence more vulnerable to cash flow risk. These effects are significantly exacerbated during recessions - time periods of heightened economic uncertainty when investors avoid risk. During low growth periods, the exposure to adverse shocks plays even a greater role. Hence, the premium an investor demands to hold stocks with high geographic risk should be higher during recessions.

Following Fama and French (1993), I use value-weighted portfolio sorts to examine the geographic risk premium. I use the classification by NBER as the primary indicator of recessions.[8] I show that stocks that co-vary closely with the *GDMA* portfolio returns (high β_{GDMA} stocks) earn higher expected returns. Using univariate sorts on all stocks in CRSP, I show that stocks with high β_{GDMA} have significantly higher expected returns than firms with low β_{GDMA}, especially during recession months. During recessions, the annualized long-short portfolio spread over the CAPM and Carhart (1997) 4-factor model ($\alpha_{4-\text{factor}}$) is approximately 15.2 percent and 12.2 percent, respectively. The annualized long-short portfolio spread is statistically insignificant during times of economic growth. The findings are consistent with a time-varying geographic risk premium.

I also perform several tests to examine the robustness of the results. Hou, Xue, and Zhang (2020) show that majority of the anomalies reported in the asset pricing literature becomes statistically insignificant once microcap stocks are removed from the sample. To mitigate this concern, I follow Hou, Xue, and Zhang (2020) and examine the equal-weighted returns on β_{GDMA} sorted portfolios excluding all microcap stocks. I continue to find a significant geographic risk premium over the CAPM. The risk premium is significantly larger during recessions. I also conduct a sub-sample test using non-manufacturing stocks for which geographic risk can only be computed using $\beta'_{GDMA}s$. Again, I continue to find qualitatively similar results.

Third, I show that the geographic risk premium is more pronounced among firms that are more

[8] National Bureau of Economic Research (NBER) defines a recession as a period of significant decline in economic activity spread across the economy, lasting more than a few months, normally visible in real GDP, real income, employment, industrial production, and wholesale-retail sales. NBER recession months are not constructed using real time information. I address this issue by constructing model-generated low growth states that use only past and contemporaneous information.

vulnerable to adverse systematic shocks. If the time-varying premium is driven by changes in the exposure to systematic risk, then the effects should be more pronounced among firms that are more vulnerable to aggregate shocks. Firms that have higher production costs or lower revenues than competitors should have higher probability of default during market downturns. For such firms, the exposure to systematic risk plays a more consequential role.

To test the cross-sectional heterogeneity in the geographic risk premium, I use two-way portfolio sorts. I employ 5×3 double sorts on pre-ranked β_{GDMA} and previous year's profitability (ROA). I then examine the long-short portfolio spread on the β_{GDMA} sorted quintiles across ROA terciles. The β_{GDMA} sorted long-short spread is statistically significant for all profitability terciles during recessions when investors avoid risk. However, during economic expansions, the geographic risk premium is significantly larger for firms in the lowest ROA tercile. Firms with low earnings are more vulnerable to negative aggregate shocks even during high growth states. During high growth states, investors remain concerned about geographic risk for firms with low earnings and command a significant risk premium. I repeat the tests using double sorts on pre-ranked β_{GDMA} and previous year's cash flow-to-assets ratio and find consistent results.

The NBER definition of recessions is commonly used in the literature to proxy low growth states. However, this measure is unavailable in real time and are subject to subsequent revision as more macroeconomic information becomes available. To reduce the look-ahead bias, I compute an alternative measure of market downturns based on financial data. I estimate a Markov regime-switching dynamic model using only the excess market returns and its lags. I then predict low growth state probabilities based on the parameters obtained from the regime-switching model. A Kalman filter is used to predict the low growth states using only the past and contemporaneous information.[9] The results are consistent using the regime switching model generated economic growth states. In the time-series, the geographic risk premium is larger during recessions. In

[9] I also categorize recession dates are defined as two consecutive quarters of decline in real U.S. GDP. This measure avoids look-ahead bias. However, a clear disadvantage of this measure relative to NBER recessions is the lower frequency of low economic growth states. The small number of observations increases the Type II error in the tests. Again, I find consistent results, which are available upon request.

the cross-section, the geographic risk premium is larger firms with low earnings, especially during recessions.

The literature reporting the relationship between a firm's location and stock returns is limited. Pirinsky and Wang (2006) show a positive co-movement in stock returns among firms headquartered in the same geographical area. Garcia and Norli (2012) find that investors display a preference for investing in local firms. Korniotis and Kumar (2013) find that local economic conditions predict the stock returns of firms in the local geographical area. Tuzel and Zhang (2017) find that firm location affects firm risk through local factor prices. Smajlbegovic (2019) examines the diffusion of regional macroeconomic information into stock prices. I show that geographic agglomeration provides a natural hedge against aggregate shocks and the effects vary over the business cycle.

This paper also contributes to the labor-based asset pricing literature. Chen, Kacperczyk, and Ortiz-Molina (2011) show that the cost of capital is higher for industries with high unionization levels. Merz and Yashiv (2007) and Belo, Lin, and Bazdresch (2014) introduce labor frictions and Donangelo et al. (2019) introduce labor leverage in asset pricing models. Donangelo (2014) shows that differences in labor mobility leads to differences in risk premiums in the cross section. Zhang (2019) show asset pricing implications when firms replace routine task labor with technology. I emphasize the link between wage costs, geographic risk, and stock returns.

The paper is organized as follows. Section 2 introduces the testable hypotheses. Section 3 introduces data and the measures. Section 4 explores the geographic risk on industries and firms. Section 5 shows the asset pricing results. In section 6 , I conduct tests based on a regime switching model and section 7 concludes.

2 Testable Hypotheses

The literature has well established the importance of local labor market networks based on proximity (e.g., Rees (1966); Bayer, Ross, and Topa (2008); Hellerstein, McInerney, and Neumark (2011); Hellerstein, Kutzbach, and Neumark (2014); Saygin, Weber, and Weynandt (2021); Hensvik and Skans (2016)). Managers use these social networks to reduce information asymmetries and hire and retain workers with better capabilities. Employees are more likely to receive and accept wage offers from businesses that employ others in their network.

The local market networks have important dynamics for firm cash flow and asset prices. Firms in geographically agglomerated industries are better equipped to negotiate low wage contracts during market downturns. On the labor demand side, workers have low expectations about outside job opportunities, and they are more likely to accept wage contracts at a discount from firms in the local networks. On the labor supply side, geographic proximity of firms increases the correlated actions among interacting managers (Guiso and Schivardi (2007)), amplifying the labor cost cutting during market downturns.[10] The combination of low expectations on the demand side and coordinated wage reductions on the supply side lead to more cyclical wages in agglomerated firms than in dispersed firms.

Assumption 1. *[A1] The average wage costs are more cyclical in agglomerated industries compared to geographically dispersed industries.*

The cyclical wages in agglomerated industries provides a natural hedge against aggregate shocks. Consequently, the cyclicality of cash flow should be lower in firms with low geographic risk than in firms with high geographic risk. Since geographic dispersion increases firm exposure to the aggregate shocks (geographic risk), such firms should have more procyclical cash flows. Put differently, during recessions when investors shrink from risk, cash flows are lower for firms with high geographic risk than that for firms with low geographic risk. Agglomerated firms, through greater information spillover and coordinated actions, can lower their operational costs following adverse shocks that lead to recessions. Hence, the effects of aggregate shocks on firm cash flow are lower for firms with low geographic risk. Based on this intuition, I propose the first hypothesis:

[10] Literature has well established that geographic agglomeration induces locational spillovers through information sharing (Glaeser et. al. (1992); Jaffe, Trajtenberg, Henderson (1993); Audretsch and Feldman (1996); Lucas and Rossi–Hansberg (2002); Henderson (2003); Arzaghi and Henderson (2008); Ellison, Glaeser, and Kerr (2010)).

Hypothesis 1. *[H1] Cash flows are more cyclical for firms with high geographic risk than that for firms with low geographic risk.*

The cyclicality of cash flow for firms in geographically dispersed industries should be reflected in their expected stock returns. Investors command additional compensation, in the form of higher expected returns, for holding firms in dispersed industries that have higher exposure to aggregate shocks. Put differently, despite the low expected returns, investors are willing to hold stocks in agglomerated firms because such firms are naturally hedged against negative shocks and thus have low cash flow risk. The marginal benefit of this natural hedge in agglomerated firms is greater during recessions when firms and investors shrink from risk. Hence, the premium to compensate for high geographic risk should be higher during recessions than during economic expansions. Based on this intuition, I hypothesize the following:

Hypothesis 2. *[H2] The expected returns are higher for stocks with high geographic risk than for stocks with low geographic risk, especially during recessions.*

In the cross-section, the geographic risk is a greater concern for firms that are more vulnerable to adverse aggregate shocks. I consider two characteristics, although not mutually exclusive, to approximate a firm's vulnerability to adverse shocks: firm profitability (ROA) and cash flow-to-assets ratio. Even during economic expansions, firms with low profitability and cash flow are more vulnerable to adverse shocks. Based on this intuition, I hypothesize the following:

Hypothesis 3. *[H3] The geographic risk premium is more pronounced for firms that are more vulnerable to aggregate shocks.*

In the remainder of the paper, I test the validity of the assumption *A1* and the hypotheses *H1-H3*.

3 Data and Measures

3.1 Geographic Agglomeration of Industries

Industries are geographically relatively concentrated. For example, the auto parts manufacturing is concentrated in the midwestern states and in California.[11] To measure geographic agglomeration of manufacturing industries, I use the index developed by Ellison and Glaeser (1997). This measure controls for both the differences in the size distribution of plants and the differences in the size of geographic areas. The index of geographic concentration is computed as follows:

$$\gamma \equiv \frac{\sum_{i=1}^{M}(s_i - x_i)^2 - \left(1 - \sum_{i=1}^{M} x_i^2\right)^2 \sum_{j=1}^{N} z_j^2}{\left(1 - \sum_{i=1}^{M} x_i^2\right)\left(1 - \sum_{j=1}^{N} z_i^2\right)} = \frac{G - \left(1 - \sum_{i=1}^{M} x_i^2\right) H}{\left(1 - \sum_{i=1}^{M} x_i^2\right)(1 - H)}, \tag{1}$$

where s_i is the share of the industry's employment in each of M geographic areas and x_i is the share of aggregate manufacturing employment in each of the M areas. G measures the geographic concentration of an industry. The employment data are from the U.S. Census Bureau. For accuracy, the measure controls for the Herfindahl index of the industry plant size distribution, H. Higher values of γ imply higher geographic agglomeration, whereas a value of $\gamma = 0$ implies no agglomeration, where location choices are random without any labor pooling and information spillover advantages.

The index has several desirable properties. An investor can compute the index given the availability of employment data at the state level. The coefficient γ is independent of the number of plants and of their distribution. Finally, the index controls for industrial concentration. EG index is calculated at the 4-digit SIC level within the manufacturing sector. Dumais, Ellison, and Glaeser (2002) find that the geographic concentration of industries is highly stable over time, especially post-1970. To be consistent with Ellison and Glaeser (1997), I use γ computed using data from

[11] Figure A1 in the Appendix shows the auto parts manufacturing establishments in the U.S. by county in 2010. There is a clear agglomeration of firms in the auto parts sector.

1987. I then assign the 4-digit SIC industry's γ to each firm within the industry.

EG index is high for industries that are known for being highly geographically concentrated. For example, γ's for the automobile industry, automobile parts industry, photographic equipment industry, carpet industry, computer industry, computer storage devices industry, and semiconductors and related devices industry are 0.127, 0.089, 0.174, 0.378, 0.059, 0.142, and 0.064, respectively.

3.2 Financial and Accounting Data

The financial data consists of stocks in the New York Stock Exchange (NYSE), American Stock Exchange (AMEX), and Nasdaq from the Center for Research in Security Prices (CRSP) of the University of Chicago. I exclude financial firms (SIC $6000-6799$) and utilities (SIC $4900-4949$) from the sample. To exclude foreign incorporated firms and ADRs, I restrict the sample to common stocks as identified by the CRSP share code (SHRCD) of 10 or 11.

Accounting data are from the Compustat database. To avoid survivorship bias in the data, a firm must have a December fiscal-year end and at least two years of data to be included in the sample. The market value of equity (ME), the stock price times the number of shares outstanding, is computed at the end of June each year using CRSP data. Following Fama and French (1993), the book value of equity (BE) is computed as the Compustat book value of stockholders' equity (data item 60), plus balance sheet deferred taxes (data item 74) and investment tax credits (data item 208) minus the book value of preferred stock. Depending on the availability of data, I use the redemption (data item 56), liquidation (data item 10), or par value (data item 130) of preferred stock. The book-to-market equity (BE/ME) is the book equity for the fiscal year ending in calendar year $t-1$ divided by market equity at the end of December of $t-1$. Negative or zero book values are treated as missing. I use further screening to satisfy the standard requirements in finance literature.

The debt to assets ratio, cash flow-to-assets ratio (CF/Assets), return on assets (ROA), and investment-to-assets ratio are calculated using Compustat data. The CF/Assets ratio is the income before extraordinary items (data item 18) plus depreciation and amortization (data item 133) divided by total assets (data item 6). The investment-to-assets ratio is the change in total assets divided by lag total assets.

3.3 *GDMA* Portfolio

I classify firms in the top tercile of γ industries as agglomerated firms. Similarly, I classify firms in the bottom tercile of γ industries as geographically dispersed firms. I map the γ's with stock returns from NYSE, AMEX, and Nasdaq to construct the two portfolios. I exclude all microcap firms when forming the portfolios.[12] For the sample from July 1947 to December 2018, the average number of firms equals 197.5 and 143.0 for the agglomerated portfolio and geographically dispersed portfolio, respectively.

At each point in time, I independently sort stocks based on NYSE size cutoffs; small (less than 50^{th}percentile), medium (50^{th} to 80^{th}percentile), and large (greater than 80^{th}percentile). I construct six portfolios from the intersection of geographic agglomeration and size (GD/S, GD/M, GD/L, A/S, A/M, A/L). Although the splits are arbitrary, results are not sensitive to these choices.

In each month, I compute the spread between the simple average of the value weighted returns on the three geographically dispersed (GD) portfolios (GD/S, GD/M, GD/L) and the simple average of the value weighted returns on the three geographically agglomerated (A) portfolios (A/S, A/M, A/L). There are enough firms in each of the 6 portfolios to diversify away most of the idiosyncratic effects. Hence, *GDMA* is the return spread between firms with geographically dispersed minus agglomerated portfolios with approximately the same weighted average market equity.[13]

For a visual inspection, Figure 1 plots the *GDMA* portfolio returns and recessions as defined by the NBER.[14] The returns on the *GDMA* portfolio returns capture aggregate shocks that are naturally hedged in geographically agglomerated firms but not in dispersed firms.

Table 1 shows the correlations between the *GDMA* portfolio returns and market, size, value,

[12] Microcaps are stocks that are smaller than the 20^{th} percentile of market equity for NYSE stocks.

[13] Table A1 in the Appendix shows the time-series average and the alphas for the GDMA portfolio returns.

[14] I plot the time-series at a quarterly frequency for a clear visual. The quarterly returns are computed using compounded monthly returns.

and momentum factors. Panel A, B, and C show the correlations for the entire sample, recessions, and expansions. For all samples, the correlations are relatively low between the *GDMA* portfolio and other factors.

4 Geographic Risk

4.1 Industry-level Analysis

The literature documents direct benefits of firm agglomeration on wages (Glaeser and Mare (2001); Rosenthal and Strange (2008)). However, the time-varying benefits of agglomeration on wages across business cycles are less understood. To test whether labor hedging assumption (*A1*) is valid, I examine the industry level wage dynamics across the business cycle. I use the industry sample from NBER and U.S. Census Bureau's Center for Economic Studies (CES) dataset. The sample includes annual industry-level data from 1958 to 2011.

To examine the effects of geographic agglomeration on industry wages, I estimate the following panel regression:

$$Y_{i,t+1} = Industry_i + Time_t + \delta_1\ Rec_t + \delta_2\ high\_\gamma + \delta_3\ Rec_t \cdot high\_\gamma + \delta_4\ Controls_{i,t} + \varepsilon_{i,t+1}, \tag{2}$$

where $Industry_i$ and $Time_t$ are industry fixed effects and time fixed effects, respectively. $Y_{i,t+1}$ is the dependent variable and $Controls_t$ is a set of controls. The variable Rec_t is an indicator that equals one if the year is an NBER recession year, and zero otherwise. The variable $high\_\gamma$ is an indicator that equals the value unity if the industry's geographic agglomeration is in the top quartile, and zero otherwise. Industry controls include the change in total employment, change in real equipment, change in real capital structures. Since the level of the dependent variable is likely correlated overtime and across industries, I cluster the standard errors by industry and year (Petersen (2009)).

Equation 2 includes *year fixed effects*, which effectively demean each observation by its yearly average. I do not include Rec_t in the regression by itself since yearly time-series variables have no explanatory power in a regression that includes time fixed effects. An alternative approach is to remove the year fixed effects and include the Rec_t term in the regression. For robustness, I also examine the results excluding year fixed effects and including the Rec_t indicator, inflation, log of real GDP growth, T-bill rate, BAA minus t-bond spread.

The dependent variables are the change in wages (total payroll divided by total value added) in columns 1 to 3 and change in employment (total employment in 1000s) in columns 4 to 6. To test whether wages are more procyclical in agglomerated industries, I introduce an interaction term between the high agglomerated industry indicator and the recession indicator, $Rec_t \cdot high\_\gamma$. If agglomeration leads to lower wages in recession, then the interaction term should be negative.

Table 2 shows the results. I show three different regression models. The first specification includes industry fixed effects. The second specification includes industry fixed effects, industry controls, and aggregate macroeconomic controls. The third specification adds year fixed effects. In the third specification, macroeconomic controls cannot be included with year fixed effects, since doing so would mechanically absorb all the explanatory power of the aggregate controls.

Columns 1 to 3 show the results for regressions with the next period's change in wage as the dependent variable. The interaction term, $Rec_t \cdot high\_\gamma$, is negative and significant in all three specifications. This validates the assumption that wages are more procyclical in agglomerated industries than in geographically dispersed industries. The results support the assumption *A1* that the average wage costs are more cyclical in agglomerated industries than in geographically dispersed industries. Columns 4 to 6 show the results for regressions with the next period's change in employment as the dependent variable. The interaction term is statistically insignificant. Hence, firms in agglomerated industries are able to reduce wage costs during market downturns without losing a significant part of the workforce.

4.2 Geographic Risk: Firm-level Measure

To measure geographic risk at the firm level, I compute a firm's exposure to the aggregate hedge factor, GDMA portfolio. I calculate a firm's stock return beta with respect to the *GDMA* portfolio returns. Specifically, for each firm i, I estimate the following univariate time-series regression:

$$r_{i,t} - r_{f,t} = \alpha_{i,t} + \beta^{i}_{GDMA,t}\ r_{GDMA,t} + \epsilon_{i,t}, \tag{3}$$

where $r_{i,t}$ is the return on asset i, $r_{f,t}$ is the risk-free rate, and $r_{GDMA,t}$ is the returns on the *GDMA* portfolio. Following Fama and French (1992), I use the standard 60-month rolling window to estimate the parameters and require at least 24 observations to be included in the sample. The parameter β^{i}_{GDMA} is firm i's exposure to the *GDMA* portfolio returns.

I use $\beta'_{GDMA}s$ as a measure of geographic risk at the firm-level. There are several advantages for using $\beta'_{GDMA}s$ to measure geographic risk. This measure can be computed for all publicly listed firms. Hence, I can expand the analysis to include firms outside of the manufacturing sector. Another advantage is that this geographic risk measure can be computed at a higher frequency, which better suites asset pricing studies.

4.3 Firm-level Analysis

Using the firm level measure of geographic risk, I test the first hypothesis *H1*. Specifically, I test whether geographic risk increases the cyclicality of firm cash flow. I create an indicator variable, $High\_\beta_{GDMA,t}$, which equals one if a firm's β_{GDMA} is in the top tercile in a given year, and zero otherwise. Firms with high β_{GDMA} are treated as firms with high geographic risk, and the rest of the firms act as the control group. To identify the treatment period, I create an indicator variable, Rec_t, which equals one for NBER recessions, and zero otherwise. The variable of interest is the interaction term between high β_{GDMA} indicator and the recession indicator, $High\_\beta_{GDMAi,t} \cdot Rec_t$.

The dependent variables are firm level ROA and CF/Assets ratio. If geographic dispersion increases the cyclicality of cash flow, then the interaction term should be negative. Put differently, following an adverse shock, the coordinated actions taken to lower wages by managers in close proximity abate the adverse impact on cash flow in agglomerated firms. To test this hypothesis, I estimate the following panel regression:

$$Y_{i,t+1} = Firm_i + Time_t + \delta_2\ High\_\beta_{GDMA i,t} + \delta_3\ High\_\beta_{GDMA i,t} \cdot Rec_t + \delta_4\ Controls_{i,t} + \varepsilon_{i,t+1}, \quad (4)$$

where $Firm_i$ and $Time_t$ are firm fixed effects and year fixed effects, respectively. $Y_{i,t+1}$ is the dependent variable (next period's ROA and CF/Assets ratio), and X_t is a set of firm level controls. Firm level controls include CF/Assets ratio, ROA, Debt/Assets ratio, BE/ME ratio, and the log of firm size. Since the level of the dependent variable is likely correlated overtime and across firms, I cluster standard errors by firm and year to correct for inflated t-statistics (Petersen (2009)). Because the regressions use accounting data from Compustat, the sample starts from 1963.

Table 3 reports the results estimating 4 . Columns (1) and (2) report the regression results using the entire sample of publicly listed firms. The coefficient δ_2 is positive in both specifications implying that high β_{GDMA} firms, on average, have higher ROA and CF/Assets ratio. The coefficient δ_3 on the interaction term, $High\_\beta_{GDMA i,t} \cdot Rec_t$ is negative and statistically significant for both ROA and CF/Assets ratio. The results imply that earnings are more cyclical for firms with high geographic risk than that for firms with low geographic risk, consistent with the hypothesis *H1*.

The results also counter the argument that cyclicality is driven by higher abnormal returns for high β_{GDMA} firms during recessions. The results show that high β_{GDMA} firms experience a lower cash flow during recessions. The negative interaction term is more consistent with a risk-based argument.

For robustness, I perform a sub-sample analysis. One of the advantages of using β_{GDMA} to measure geographic risk is that the betas can be computed for all publicly listed firms and not just for the firms in the manufacturing sector. I split the sample into two: 1) the sample of manufacturing firms for which the geographic risk can be computed at the industry level using the EG index in 1 and 2) the sample of non-manufacturing firms for which the geographic risk can only be computed using $\beta'_{GDMA}s$.

Column (3) and (4) report the regression results using the manufacturing sector. Column (5) and (6) report the regression results using the sample of firms for which the geographic risk is computed using $\beta'_{GDMA}s$. Both samples produce similar results. In both sub-samples, firms with high $\beta'_{GDMA}s$ have more cyclical earnings than for firms with low $\beta'_{GDMA}s$.

5 Geographic Risk and Equity Returns

In this section, I formally examine the pricing of geographic risk. Following Fama and French (1993), I perform asset pricing tests using both univariate and multivariate portfolio sorts.

5.1 Portfolio Sorts: Value-weighted Returns

To expand the analysis to all stocks, I use β'_{GDMA}s from equation 3 to proxy geographic risk. I include all stocks from NYSE, AMEX, and Nasdaq exchanges for the period July 1947 to December 2018. At the end of June of each year *t,* I sort stocks into quintiles based on the pre-ranked β_{GDMA} measured at the fiscal year ending in calendar year *t-1* and calculate the value-weighted returns from July of year t to June of t *+1.*

Table 4 reports the summary statistics at a monthly frequency. Panel A reports the firm characteristics across the portfolios. Portfolio A comprises of stocks with the lowest pre-ranked β'_{GDMA}s (lowest geographic risk portfolio) and portfolio GD comprises of stocks with the highest pre-ranked β'_{GDMA}s (highest geographic risk portfolio). The average pre-ranked β'_{GDMA}s range from -2.8 to 1.7, which is sizable.[15] In contrast, there is little variation in the pre-ranked β'_{Mkt}s, which captures a stock's sensitivity to the market portfolio. There is limited variation in firm size

[15]Small variation in betas often lead to erroneously measured factor premiums. I also find that ex-post β'_{GDMA}s align well with the pre-ranked β'_{GDMA}s.

and BE/ME ratio.

Panel B reports the average value-weighted portfolio excess returns (returns in excess of the risk-free rate) and the Sharpe ratios. Stocks with the lowest β'_{GDMA}s have the lowest average returns, while stocks with the highest β'_{GDMA}s have the highest average returns. Going from quintile A, the portfolio of firms with the lowest geographic risk, to quintile GD, the portfolio of firms with the highest geographic risk, the average value-weighted returns increase almost monotonically. The Sharpe ratio, which is the excess return divided by the standard deviation, increases monotonically as geographic risk increases. The Sharpe ratio for the highest geographic risk portfolio is almost twice in magnitude compared to the Sharpe ratio for the lowest geographic risk portfolio.

Table 5 reports the excess returns adjusted for the CAPM (α_{CAPM}) and the Carhart (1997) 4-factor ($\alpha_{4-\text{factor}}$) model for the pre-ranked β_{GDMA} sorted quintiles. I also report the $\alpha\prime$s for the return spread between the highest (quintile GD) and the lowest pre-ranked β_{GDMA} sorted portfolio (quintile A), dubbed the *GD-A* spread. To examine the time-series properties of the *GD-A* spread, I report the results for the full sample, NBER recessions, and economic expansions.

Panel A reports the results for the full sample. α_{CAPM} increases monotonically as geographic risk increases. The unexplained portion of the annualized return spread on the *GD-A* spread controlling for the CAPM is approximately 4.1 percent. Hence, an investor will only buy stocks with geographic risk at a discounted price. In other words, an investor requires compensation in the form of high expected returns to hold stocks with high geographic risk. However, for the full sample, the *GD-A* spread becomes weaker controlling for the 4-factor model.

Panel B reports the risk adjusted returns during recessions. The intercept increases from the low geographic risk to high geographic risk portfolios. After controlling for the CAPM and the 4-factor model, the high geographic risk portfolios earn higher expected returns than low geographic risk portfolios. The unexplained portion of the annualized *GD-A* return spread controlling for the CAPM and 4-factor model are approximately 15.2 percent and 13.2 percent, respectively, during recessions. The α's are both statistically and economically significant. The results are consistent with the hypothesis *H2:* the expected returns are higher for stocks with high geographic risk than for stocks with low geographic risk and the premium is larger during recessions when the marginal benefit of hedging risk is high.

For robustness, I also conduct a sub-sample test using non-manufacturing stocks for which geographic risk can only be computed using $\beta'_{GDMA}s$. Importantly, these non-manufacturing stocks are not used in the construction of the *GDMA* portfolio. Hence, this sub-sample of stocks provides a clean out-of-sample test. I find a strong geographic risk premium for the sub-sample. Both the α_{CAPM} and $\alpha_{4-\text{factor}}$ are significant for the full sample. The premium is significantly larger during the recession months, which is consistent with the hypothesis *H2*.[16]

5.2 Equal-weighted Returns Excluding Microcaps

Hou, Xue, and Zhang (2020) find that most anomalies reported in the literature disappear when microcap firms are excluded from samples. In fact, 65 percent of the more than 450 anomalies examined in their paper failed to clear a t-statistic of 1.96. To examine whether the geographic risk premium is robust to such critique, I run the portfolio sorts using NYSE breakpoints and equal-weighted returns excluding all microcap firms. Microcaps are stocks that are smaller than the 20^{th} percentile of the market equity for NYSE stocks. Specifically, I exclude all microcaps, and then split stocks at the end of June of each year t into quintiles based on β_{GDMA} measured at the fiscal year ending in calendar year *t-1* and calculate equal-weighted returns from July of year t to June of t *+1*.

Table 6 reports the equal-weighted returns adjusted for the CAPM (α_{CAPM}) and the Carhart (1997) 4-factor ($\alpha_{4-\text{factor}}$) model for the β_{GDMA} sorted quintiles. The results continue to hold. For the full sample, α_{CAPM} increases, albeit non-monotonic, as geographic risk increases. The geographic premium is stronger during recessions when the benefit of hedging risk is high.

[16] For brevity, I report the sub-sample results in the Appendix Table A2 .

5.3 Heterogeneity in Geographic Premium

The geographic risk is exacerbated for firms that are more vulnerable to adverse systematic shocks. Such firms have an *ex-ante* higher probability of default. Profitable firms, in contrast, are less vulnerable to adverse shocks since they have high cash flow to reduce the impact of the shocks. To test this hypothesis, I independently double sort stocks into 5 $\times$ 3 portfolios based on the pre-ranked β_{GDMA} and the previous fiscal year's return on asset (ROA). For adequate diversification, I include all stocks from July 1963 to December 2018.[17] To form β_{GDMA} quintiles, at the end of June of each year *t,* I sort stocks into quintiles based on β_{GDMA} measured at the fiscal year ending in calendar year *t-1.* To form the ROA portfolios, in June of each year *t*, I sort stocks into terciles using NYSE breakpoints based on ROA measured at the fiscal year ending in calendar year *t-1.* I then calculate the value-weighted returns from July of year *t* to June of *t +1* for the 15 portfolios sorted on pre-ranked β_{GDMA} and ROA.

Table 7 shows the time-series average of the return spread between the highest and the lowest pre-ranked β_{GDMA} sorted quintiles (*GD-A* spread) for each ROA tercile. I also show the results for NBER recessions and economic expansions. The *GD-A* spread is significantly larger during recessions than during expansions for all ROA terciles. This is expected since the marginal benefit of hedging against aggregate shocks is greater during recessions.

For the full sample, the *GD-A* spread is statistically significant only for the lowest ROA tercile. For high ROA firms, the impact of aggregate shocks is low during times of economic growth. For low ROA firms, however, geographic risk remains a concern even during expansions. Hence, even in times of economic growth, geographic risk matters for firms that are more vulnerable to adverse aggregate shocks, consistent with the hypothesis *H3.* This pattern also holds when low cash flow to assets ratio is used to identify firms that are more vulnerable to adverse systematic shocks.

[17] Accounting data on ROA and CF/Assets ratio are available only from 1963 through Compustat database.

6 Discussion

In this section, I examine the robustness of the results. First, I examine the sensitivity of the results to the definition of recessions. Second, I examine whether other characteristics can potentially explain the geographic risk premium.

6.1 Markov Regime Switching Model based Low Growth States

For the baseline, I use recessions defined by NBER to measure of economic downturns. NBER uses data on GDP, real income, employment, industrial production, and wholesale-retail sales to determine market downturns. A potential concern is that macroeconomic data are not available in real time and are subject to subsequent revision as more information becomes available. In fact, revisions to early estimates of macro variables such as GDP can be large. To mitigate this look ahead bias, I propose a measure of low growth states using a Markov regime-switching dynamic model.

Following Hamilton (1989) and Hamilton (1990), I estimate the probabilities of low growth states using a Markov regime-switching dynamic model. I estimate the model using market returns, which proxy economic conditions. Specifically, I estimate the following model:

$$r_{mkt,t} - r_{f,t} = \alpha_{s_t} + z_t \beta_{st} + \epsilon_s, \tag{5}$$

where $r_{mkt,t}$ is the return on the market portfolio, μ_{s_t} is the state-dependent intercept, z_t is a vector of independent variables with state-dependent coefficients β_s, and ϵ_s is an independent and identically distributed ($i.i.d.$) normal error with zero mean and state-dependent variance σ_s^2. The vector of independent variables includes 3 lags of the dependent variable. I estimate the parameters of the model using an expectation–maximization (EM) algorithm to serve as initial value for the quasi-Newton optimizer. The model is estimated in monthly frequency using data from January 1947 to December 2018.

Table 8 shows the estimated parameters of the regime-switching model. State 1 is the low growth state with a monthly intercept of -1.7 percent. State 2 is the high growth state with an intercept of 2.2 percent. Wald test rejects the null of equal intercepts ($\alpha_1 = \alpha_2$). Wald tests also confirm that the coefficients on the independent variables are state-dependent. Low growth states have higher volatility than high growth states ($\sigma_1^2 > \sigma_2^2$); the test rejects the null of equal volatility ($\sigma_1^2 = \sigma_2^2$). This illustrates the well-known conditional heteroskedasticity in stock returns. The high growth state is more persistent than the low growth state. The ratio of the average duration of the high growth state to low growth state is 1.5.

I estimate the state probabilities using only the past and contemporaneous data employing a Kalman filter. Figure 2 , top panel, shows the 6-month moving average of the model fit and the 6-month moving average of the excess market returns. The model fit and the realized returns co-vary closely and have a correlation of 0.76. The lower panel shows the estimated probability of being in the low growth state. The probability of a low growth state increases as market returns decrease. I also show the NBER recessions using dashed red lines for reference. The probability of being in a low growth state tends to increase in NBER recessions. However, there are many occurrences in which low growth state probabilities are high outside of the NBER recession dates.

I construct an indicator variable to identify low growth periods, I_{S1}, which takes a value of 1 if the estimated probability of being in a low growth state is higher than 0.7 in month t and 0 otherwise.[18] Of 858 months from July 1947 through December 2018, 126 months belong to low growth states and 732 months belong to high growth states. Hence, 17 percent of the months in the sample are part of low growth states. In comparison, 122 months belong to NBER recessions and 39 months belong to recessions defined as two consecutive quarters of decline in GDP.

Using the regime switching model generated probabilities, I re-examine the two-way portfolio sorts. Again, I independently sort stocks into 5 $\times$ 3 portfolios based on the pre-ranked β_{GDMA} and the previous fiscal year's ROA. Instead of NBER recessions, I use the model based I_{S1} indicator to

[18]While the choice of the probability threshold is somewhat arbitrary, I find that the use of other reasonable levels between 0.5 and 1 leads to similar asset pricing results.

capture recessions.

Table 9 shows the averages of the return spread between the highest and the lowest β_{GDMA} sorted portfolios (*GD-A* spread) for the full sample, low growth states, and high growth states. *GD-A* spread is significant across all ROA terciles during low growth states. In contrast, *GD-A* spread is statistically trivial or negative during high growth periods. The high geographic premium during low growth states is consistent with the hypothesis *H2*. The geographic premium is also stronger for low profitable and low cash flow firms that are more vulnerable to aggregate shocks, consistent with the hypothesis *H3*.

6.2 Alternative Explanations

Finally, I examine whether the other characteristics can explain the *GDMA* hedge factor. I use a sample from January 1967 to December 2018 for the tests because of data restrictions. As before, I start with the unexplained portion of the returns controlling for CAPM (α_{CAPM}), Fama and French (1993) 3-factor ($\alpha_{3-\text{factor}}$), and Carhart (1997) 4-factor ($\alpha_{4-\text{factor}}$) models. Table 10 shows the α's controlling for the empirical factor models. The positive and significant α's imply that the information embedded in the *GDMA* hedge factor is not subsumed by the classical empirical models.

There is a clear association between geographic distribution of firms and international tradability. Firms that internationally trade goods tend to be, on average, more geographically concentrated than other firms (Jensen et al. (2005); Hlatshwayo and Spence (2014)). Agglomeration raises the probability of export market entry because of spillover externalities (Clerides, Lach, and Tybout (1998); Greenaway and Kneller (2008)). Since industries with high tradability have more cyclical stock returns (Tian (2018)), it is possible that tradability alone could explain the geographic risk premium. Hence, I examine the robustness of the geographic risk premium controlling for international tradability.

In the spirit of Tian (2018), I create an index of firms based on the tradability of their output.

I use the 2007 BEA NIPA Input-Output Tables to compute the tradability ratio for over 400 industries. This ratio is the value of exports divided by the total industry output. I construct a high minus low tradability portfolio, dubbed *TMNT*, defined as the difference in value-weighted excess returns of high tradable firms minus low tradable firms.[19] Next, I examine the α controlling for the excess market returns and the *TMNT* factor. The results show that geographic risk premium is not subsumed by international tradability premium. Although agglomeration increases the probability of export market entry, it is unlikely a primary driver for the geographic premium.

Another characteristic related to geographic agglomeration of firms is the durability of output. Gomes, Kogan, and Yogo (2009) show that the durability of a firm's output is a characteristic related to systematic risk. They show that the returns of durable-goods producers are higher and more volatile than those of service producers and nondurable-goods producers. The aggregate expenditure on durable goods is more cyclical than that on nondurable goods and services. It is possible that durability of output is driving the results in this study. To mitigate such concerns, I examine whether the geographic premium is robust controlling for the time-varying durability premium.

Following Gomes, Kogan, and Yogo (2009), I construct portfolios based on a firm's output durability. NIPA tables classify personal consumption expenditure into durable goods, nondurable goods, and services consumption. Using the 1987 benchmark input-output accounts from the BEA, I assign each industry a category based on the final demand to which it contributes the highest value: personal consumption expenditure on durable goods, personal consumption expenditure on nondurable goods, personal consumption expenditure on services, investment, government expenditures, and net exports. I focus on the three categories of consumption based on durability. I map the industry classifications with the universe of stocks from CRSP. I construct a durability factor, durable minus non-durable portfolio. I then examine the unexplained portion of the *GDMA* returns controlling for the time-varying durability premium. The positive α implies that the effects

[19] Following Tian (2018), I use portfolio quintiles to construct *TMNT*. Portfolio T is the quintile five and portfolio NT is the quintile one.

of durability and economic geography are largely independent.

Another factor that could potentially explain the time-varying geographic risk premium is variation in market liquidity, which is the ability to trade large quantities swiftly at low cost without moving the price. During market downturns, liquidity is significantly lower than during economic expansions. It is possible that the geographic risk premium is capturing cross-sectional variations in an equity's sensitivity to market liquidity, which raises its equilibrium expected returns. To capture variations in market wide liquidity, I employ the aggregate liquidity measure introduced by Pastor and Stambaugh (2003).[20] The results in Table 10 shows that the unexplained part of the returns controlling for liquidity remains significant. In fact, the unexplained part of the returns increases controlling for market wide liquidity.

7 Conclusion

This paper examines the dynamic effects of geographic distribution of firms on asset prices. I show that agglomeration benefits go beyond the static labor market pooling externalities. Agglomeration of firms increases information spillover and coordinated actions by interacting managers. As a consequence, wage costs are more cyclical in agglomerated industries than in geographically dispersed industries. The procyclicality of wages provides firms in agglomerated industries a natural hedge against aggregate shocks.

I show that the cash flows are more cyclical for firms with high geographic risk than for firms with low geographic risk. Investors command additional compensation for holding firms with high geographic risk, especially during low growth states. I show that the expected returns are higher for stocks with high geographic risk than for stocks with low geographic risk, especially during recessions when investors shrink from risk. In the cross-section, the geographic risk is a greater concern for firms that are more vulnerable to aggregate shocks.

[20] Pastor and Stambaugh (2003) aggregate liquidity measure is a cross-sectional average of individual-stock liquidity measures.

References

Amiti, M., Cameron, L., 2007. Economic geography and wages. The Review of Economics and Statistics 89, 15-29.

Arzaghi, M., Henderson, J.V., 2008. Networking off madison avenue. The Review of Economic Studies 75, 1011-1038.

Audretsch, D.B., Feldman, M.P., 1996. R&D Spillovers and the Geography of Innovation and Production. American Economic Review 86, 630-640.

Banerjee, A.V., 1992. A Simple Model of Herd Behavior. The Quarterly Journal of Economics, 797-817.

Bayer, P., Ross, S.L., Topa, G., 2008. Place of work and place of residence: Informal hiring networks and labor market outcomes. Journal of Political Economy 116, 1150-1196.

Belo, F., Lin, X., Bazdresch, S., 2014. Labor hiring, investment, and stock return predictability in the cross section. Journal of Political Economy 122, 129-177.

Bikhchandani, S., Hirshleifer, D., Welch, I., 1992. A theory of fads, fashion, custom, and cultural change as informational cascades. Journal of Political Economy 100, 992-1026.

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

Chen, H.J., Kacperczyk, M., Ortiz-Molina, H., 2011. Labor unions, operating flexibility, and the cost of equity. Journal of Financial and Quantitative Analysis 46, 25-58.

Clerides, S.K., Lach, S., Tybout, J.R., 1998. Is learning by exporting important? Micro-dynamic evidence from Colombia, Mexico, and Morocco. The Quarterly Journal of Economics 113, 903-947.

Davis, M.A., Fisher, J.D., Whited, T.M., 2014. Macroeconomic implications of agglomeration. Econometrica 82, 731-764.

Donangelo, A., 2014. Labor mobility: Implications for asset pricing. The Journal of Finance 69, 1321-1346.

Donangelo, A., Gourio, F., Kehrig, M., Palacios, M., 2019. The cross-section of labor leverage and equity returns. Journal of Financial Economics 132, 497-518.

Dumais, G., Ellison, G., Glaeser, E.L., 2002. Geographic concentration as a dynamic process. Review of Economics and Statistics 84, 193-204.

Ellison, G., Glaeser, E.L., 1997. Geographic concentration in US manufacturing industries: a dart-board approach. Journal of Political Economy 105, 889-927.

Ellison, G., Glaeser, E.L., Kerr, W.R., 2010. What causes industry agglomeration? Evidence from coagglomeration patterns. American Economic Review 100, 1195-1213.

Fama, E., 1991. Efficient Capital Markets: II. Journal of Finance 46, 1575-1617.

Fama, E.F., French, K.R., 1992. The Cross-Section of Expected Stock Returns. Journal of Finance, 427-465.

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

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

Glaeser, E., Maré, D., 2001. Cities and Skills. Journal of Labor Economics 19, 316-342.

Glaeser, E.L., Kallal, H.D., Scheinkman, J.A., Shleifer, A., 1992. Growth in Cities. Journal of Political Economy 100.

Glaeser, E.L., Scheinkman, J., 2000. Non-Market Interactions. NBER Working Paper.

Gomes, J.F., Kogan, L., Yogo, M., 2009. Durability of output and expected stock returns. Journal of Political Economy 117, 941-986.

Greenaway, D., Kneller, R., 2008. Exporting, productivity and agglomeration. European Economic Review 52, 919-939.

Guiso, L., Schivardi, F., 2007. Spillovers in industrial districts. The Economic Journal 117, 68-93.

Hamilton, J.D., 1989. A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle. Econometrica 57, 357-384.

Hamilton, J.D., 1990. Analysis of time series subject to changes in regime. Journal of Econometrics 45, 39-70.

Hellerstein, J.K., Kutzbach, M.J., Neumark, D., 2014. Do labor market networks have an important spatial dimension? Journal of Urban Economics 79, 39-58.

Hellerstein, J.K., McInerney, M., Neumark, D., 2011. Neighbors and coworkers: The importance of residential labor market networks. Journal of Labor Economics 29, 659-695.

Henderson, J.V., 2003. Marshall's scale economies. Journal of Urban Economics 53, 1-28.

Hensvik, L., Skans, O.N., 2016. Social networks, employee selection, and labor market outcomes. Journal of Labor Economics 34, 825-867.

Hlatshwayo, S., Spence, M., 2014. Demand and defective growth patterns: The role of the tradable and non-tradable sectors in an open economy. American Economic Review 104, 272-277.

Holmes, T.J., 1998. The Effect of State Policies on the Location of Manufacturing: Evidence from State Borders. Journal of Political Economy 106, 667-705.

Hou, K., Xue, C., Zhang, L., 2020. Replicating anomalies. The Review of Financial Studies 33, 2019-2133.

Jaffe, A.B., Trajtenberg, M., Henderson, R., 1993. Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations. The Quarterly Journal of Economics, 577-598.

Jensen, J.B., Kletzer, L.G., 2005. Tradable Services: Understanding the Scope and Impact of Services Offshoring, Brookings Trade Forum. Brookings Institution Press, pp. 75-116.

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

Krugman, P., 1991. Increasing Returns and Economic Geography. Journal of Political Economy 99, 483-499.

Lucas, R.E., Rossi–Hansberg, E., 2002. On the internal structure of cities. Econometrica 70, 1445-1476.

Marshall, A., 1890. 1920. Principles of economics. London: Mac-Millan, 1-627.

Merz, M., Yashiv, E., 2007. Labor and the Market Value of the Firm. American Economic Review 97, 1419-1431.

Molloy, R., Smith, C.L., Wozniak, A., 2011. Internal migration in the United States. Journal of Economic Perspectives 25, 173-196.

Montgomery, J.D., 1991. Social Networks and Labor-Market Outcomes: Toward an Economic Analysis. American Economic Review 81, 1407-1418.

Pástor, Ľ., Stambaugh, R.F., 2003. Liquidity Risk and Expected Stock Returns. Journal of Political Economy 111, 642-685.

Petersen, M.A., 2009. Estimating Standard Errors in Finance Panel Data Sets: Comparing Approaches. The Review of Financial Studies 22.

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

Rees, A., 1966. Information networks in labor markets. The American Economic Review 56, 559-566.

Rosenthal, S.S., Strange, W.C., 2001. The determinants of agglomeration. Journal of Urban Economics 50, 191-229.

Rosenthal, S.S., Strange, W.C., 2008. The attenuation of human capital spillovers. Journal of Urban Economics 64, 373-389.

Saygin, P.O., Weber, A., Weynandt, M., 2021. Coworkers, Networks, and Job-Search Outcomes among Displaced Workers. ILR Review 74, 95-130.

Smajlbegovic, E., 2019. Regional economic activity and stock returns. Journal of Financial and Quantitative Analysis 54, 1051-1082.

Tian, M., 2018. Tradability of output, business cycles and asset prices. Journal of Financial Economics 128, 86-102.

Tuzel, S., Zhang, M.B., 2017. Local risk, local factors, and asset prices. The Journal of Finance 72, 325-370.

Zhang, M.B., 2019. Labor-technology substitution: Implications for asset pricing. The Journal of Finance 74, 1793-1839.

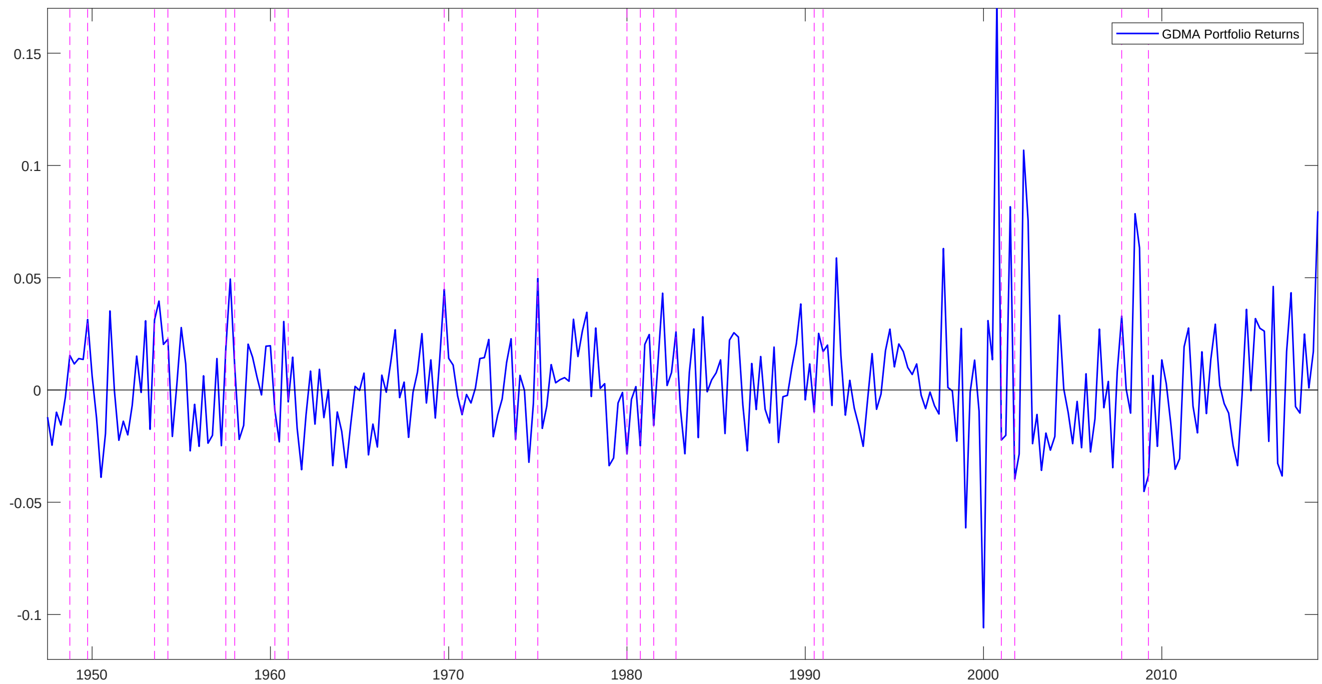

Figure 1: GDMA return spread over the business cycles

The figure plots the GDMA portfolio returns and NBER recessions. GDMA is the long-short value-weighted portfolio of geographically dispersed (D) minus agglomerated (A) industries in the manufacturing sector. A recession is a significant decline in economic activity which spread across the economy, lasting more than a few months, normally visible in real GDP, real income, employment, industrial production, and wholesale-retail sales as defined by NBER. The sample covers the post war time period from 1947 Q3 to 2018 Q4.

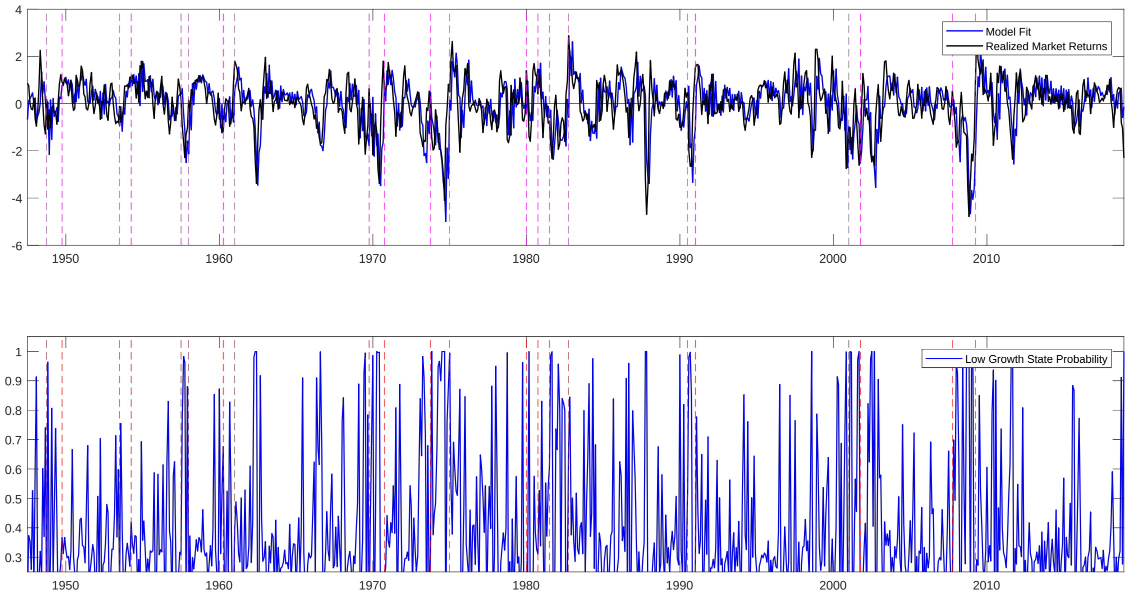

Figure 2: Regime-Switching Model Fit and Low Growth State Probabilities

I run the regime switching model, $r_t^e = \alpha_{s_t} + z_t\beta_{st} + \epsilon_s$, where r_t^e is the excess return on the market portfolio, μ_{s_t} is the state-dependent intercept, z_t is a vector of independent variables with state-dependent coefficients β_s, and ϵ_s is an independent and identically distributed (i.i.d.) normal error with zero mean and state-dependent variance σ_s^2. I use three lags of the dependent variable as controls. The top panel shows the 6-month moving average of the model fit and the 6-month moving average of the excess market returns. The bottom panel shows the estimated probability of being in the low growth state. The sample consists of data from July 1947 through December 2018.

Table 1: Factor Correlations

	GDMA	Mkt-rf	SMB	HML	MOM
Panel A: Full Sample					
Mkt-rf	-0.186	1.000			
	(0.00)				
SMB	-0.075	0.256	1.000		
	(0.03)	(0.00)			
HML	0.016	-0.192	-0.165	1.000	
	(0.65)	(0.00)	(0.00)		
MOM	0.090	-0.107	-0.032	-0.183	1.000
	(0.01)	(0.00)	(0.34)	(0.00)	
Panel B: Recessions					
Mkt-rf	-0.201	1.000			
	(0.03)				
SMB	0.066	0.363	1.000		
	(0.47)	(0.00)			
HML	0.112	-0.149	-0.054	1.000	
	(0.22)	(0.10)	(0.55)		
MOM	0.323	-0.437	-0.397	-0.171	1.000
	(0.00)	(0.00)	(0.00)	(0.06)	
Panel C: Expansions					
Mkt-rf	-0.180	1.000			
	(0.00)				
SMB	-0.102	0.230	1.000		
	(0.01)	(0.00)			
HML	-0.005	-0.208	-0.192	1.000	
	(0.90)	(0.00)	(0.00)		
MOM	0.029	0.024	0.078	-0.190	1.000
	(0.43)	(0.52)	(0.03)	(0.00)	

The table reports the correlation and standard errors (in paranthesis) between risk factors used in the baseline study. GDMA is the long-short value-weighted portfolio of geographically dispersed (D) minus agglomerated (A) industries in the manufacturing sector. MKT is the excess returns on the value-weighted market portfolio, SMB is the portfolios of small stocks minus big stocks, HML is the difference between the portfolios high and low book-to-market stocks. MOM is the momentum factor. The sample includes monthly data from July 1947 to December 2018. .

Table 2: Industry dynamics across the business cycle

	(1) Δ Payroll$_{t+1}$	(2) Δ Payroll$_{t+1}$	(3) Δ Payroll$_{t+1}$	(4) Δ Emp$_{t+1}$	(5) Δ Emp$_{t+1}$	(6) Δ Emp$_{t+1}$
High γ x Rec	-0.238***	-0.236***	-0.237***	0.101	0.055	0.061
	(-3.527)	(-3.431)	(-3.374)	(0.239)	(0.131)	(0.144)
High γ	0.318***	0.296***	0.219**	0.195	0.272*	0.319*
	(3.049)	(3.083)	(2.633)	(1.423)	(1.779)	(1.771)
Rec	-0.026	-0.518**		-1.720***	-0.726	
	(-0.123)	(-2.114)		(-3.213)	(-1.154)	
Industry Controls	NO	YES	YES	NO	YES	YES
Macro Controls	NO	YES	NO	NO	YES	NO
Industry FE	YES	YES	YES	YES	YES	YES
Year FE	NO	NO	YES	NO	NO	YES
Observations	22,203	22,203	22,203	22,203	22,203	22,203
Adjusted R^2	0.021	0.070	0.188	0.020	0.051	0.085

The table presents the results estimating 4 on the sample of industries in the National Bureau of Economic Research (NBER) and U.S. Census Bureau's Center for Economic Studies (CES) annual industry-level dataset from 1958 – 2011. Models 1 to 3 present the results using the change in payroll (total payroll divided by total value added) as the dependent variable. Model 4 to 6 present the results using the change in employment (Total employment in 1000s) as the dependent variable. All regressions include industry fixed effects. Models 3 and 6 include year fixed effects. Industry controls include the change in total employment, change in real equipment, change in real capital structures. Macro controls include the change in inflation, log of real GDP growth, t-bill rate, BAA minus t-bond spread. T-statistics in parentheses are based on standard errors adjusted for two-way clustering within industry and year. Observations are the total number of industry-year observations. ***, **, *, indicates significance at the 1%, 5%, and 10% level, respectively.

Table 3: Geographic risk across the business cycle

	(1)	(2)	(3)	(4)	(5)	(6)
	CF/Assets$_{t+1}$	ROA$_{t+1}$	CF/Assets$_{t+1}$	ROA$_{t+1}$	CF/Assets$_{t+1}$	ROA$_{t+1}$
High_β_{GDMA}x Rec	-0.007***	-0.008***	-0.007***	-0.005*	-0.008**	-0.011***
	(-3.071)	(-2.868)	(-2.962)	(-1.704)	(-2.419)	(-2.915)
High_β_{GDMA}	0.002**	0.003**	0.004***	0.004***	0.001	0.003*
	(2.178)	(2.591)	(2.758)	(3.028)	(0.884)	(1.827)
ROA	-0.058***	-0.078**	-0.082***	-0.011	-0.044**	-0.078***
	(-2.962)	(-2.544)	(-2.802)	(-0.293)	(-2.229)	(-3.554)
CF/Assets	0.283***	0.311***	0.317***	0.242***	0.235***	0.291***
	(9.934)	(6.471)	(7.833)	(5.167)	(5.868)	(4.711)
Log Size	0.018***	0.020***	0.018***	0.020***	0.018***	0.022***
	(9.900)	(9.293)	(8.128)	(9.721)	(7.428)	(6.885)
BE/ME	-0.016***	-0.023***	-0.014**	-0.020***	-0.017***	-0.024***
	(-4.316)	(-4.096)	(-2.030)	(-3.663)	(-4.343)	(-3.626)
Debt/Assets	-0.004	0.001	-0.026**	-0.026**	0.012	0.020*
	(-0.496)	(0.134)	(-2.475)	(-2.592)	(0.933)	(1.755)
Investment	-0.021***	-0.015***	-0.012*	-0.009	-0.026***	-0.018***
	(-4.404)	(-2.898)	(-1.676)	(-1.243)	(-4.583)	(-2.789)
Constant	-0.150***	-0.233***	-0.158***	-0.216***	-0.158***	-0.262***
	(-6.035)	(-7.514)	(-4.686)	(-6.823)	(-4.859)	(-6.076)
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	27,897	27,913	13,237	13,241	14,600	14,612
Adjusted R^2	0.490	0.293	0.498	0.450	0.497	0.233

The table presents the results estimating 4 on the sample of Compustat firms (excluding microcap firms) from 1963 – 2018. Model (1) and (2) use the full sample. Model (3) and (4) use the sample for which agglomeration can be calculated directly using 1 . Model (5) and (6) use the sample for which agglomeration can only be computed using $\beta'_{GDMA}s$. The dependent variables are cashflow to assets (CF/Assets$_{t+1}$) and return on assets (ROA$_{t+1}$). All regressions include firm fixed effects and year fixed effects. T-statistics in parentheses are based on standard errors adjusted for two-way clustering within firms and years (Petersen (2009)). Observations are the total number of firm-year observations. ***, **, *, indicates significance at the 1%, 5%, and 10% level, respectively.

Table 4: Firm Fundamentals Over the Business Cycle

	L	2	3	4	H
		Panel A: Characteristics			
# of firms	397.7	401.0	402.5	402.5	400.4
Size	13.825	14.302	14.484	14.608	14.575
BE/ME	1.190	1.020	0.975	0.907	0.868
β_{GDMA}	-2.801	-1.060	-0.283	0.393	1.681
β_{Mkt}	1.474	1.200	1.086	1.033	1.077
		Panel B: Returns			
Excess Returns	0.502	0.607	0.639	0.621	0.661
(t-stat)	(2.24)	(3.36)	(3.92)	(4.03)	(4.07)
Std. Dev.	6.341	5.096	4.600	4.357	4.581
Sharpe Ratio	7.913	11.911	13.882	14.256	14.420

The table reports the time-series averages of the value-weighted excess returns for portfolios sorted on β_{GDMA}. GDMA is the long-short value-weighted portfolio of geographically dispersed (GD) minus agglomerated (A) industries in the manufacturing sector. β_{GDMA} is the exposure to the returns on the GDMA portolio and β_{Mkt} is the exposure to the excess returns on the market portfolio. Betas are computed by estimating univariate regressions using the prior 60 months of data. I also report the average book equity to market equity (BE/ME) ratio and the market equity (Size). The sample includes data from July 1947 to December 2018.

Table 5: Portfolios sorted on GDMA beta: Value-weighted Excess Returns

	A	2	3	4	GD	GD-A
			Panel A: Full Sample			
α_{CAPM}	-0.235	-0.037	0.036	0.065	0.103	0.338
(t-stat)	(-1.93)	(-0.51)	(0.71)	(1.10)	(1.34)	(1.91)
$\alpha_{4-\text{factor}}$	-0.079	0.129	0.095	0.034	0.047	0.127
(t-stat)	(-0.64)	(1.78)	(1.80)	(0.56)	(0.60)	(0.70)
			Panel B: Recessions			
α_{CAPM}	-0.762	0.028	0.232	0.229	0.503	1.265
(t-stat)	(-1.91)	(0.13)	(1.33)	(1.39)	(2.38)	(2.25)
$\alpha_{4-\text{factor}}$	-0.636	0.106	0.261	0.249	0.463	1.099
(t-stat)	(-1.64)	(0.53)	(1.48)	(1.48)	(2.24)	(2.00)
			Panel C: Expansions			
α_{CAPM}	-0.163	-0.060	0.018	0.045	0.025	0.188
(t-stat)	(-1.28)	(-0.78)	(0.35)	(0.72)	(0.31)	(1.02)
$\alpha_{4-\text{factor}}$	-0.011	0.146	0.070	-0.016	-0.025	-0.014
(t-stat)	(-0.08)	(1.89)	(1.30)	(-0.24)	(-0.29)	(-0.07)

The table reports the summary statistics of the unexplained part of the value-weighted portfolios sorted on β_{GDMA}. GDMA is the long-short value-weighted portfolio of geographically dispersed (GD) minus agglomerated (A) industries in the manufacturing sector. The t-statistics for the return spreads are reported in parentheses. "GD-A" is the return difference between the highest and lowest β_{GDMA}sorted portfolios. I report the unexplained part of the returns over the CAPM (α_{CAPM}) and Carhart (1997) 4-factor model ($\alpha_{4-Factor}$). The sample includes data from July 1947 to December 2018. A recession is a significant decline in economic activity which spread across the economy, lasting more than a few months, normally visible in real GDP, real income, employment, industrial production, and wholesale-retail sales as defined by NBER. Expansion periods are all the non-recession months.

Table 6: GDMA beta Sorts Excluding Microcaps: Equal-Weighted Excess Returns

	A	2	3	4	GD	GD-A
			Panel A: Full Sample			
α_{CAPM}	0.126	0.481	0.528	0.563	0.491	0.365
(t-stat)	(1.03)	(6.24)	(7.30)	(7.79)	(5.90)	(2.66)
$\alpha_{4-\text{factor}}$	0.311	0.542	0.518	0.512	0.437	0.126
(t-stat)	(3.25)	(10.44)	(10.33)	(9.13)	(6.43)	(0.92)
			Panel B: Recessions			
α_{CAPM}	0.154	0.840	1.067	1.034	1.121	0.967
(t-stat)	(0.38)	(3.01)	(4.41)	(4.40)	(4.25)	(2.18)
$\alpha_{4-\text{factor}}$	0.154	0.790	0.940	0.892	0.948	0.794
(t-stat)	(0.51)	(4.48)	(7.06)	(6.63)	(5.58)	(1.89)
			Panel C: Expansions			
α_{CAPM}	0.127	0.444	0.468	0.504	0.394	0.267
(t-stat)	(1.00)	(5.75)	(6.37)	(6.75)	(4.55)	(1.87)
$\alpha_{4-\text{factor}}$	0.318	0.482	0.444	0.438	0.348	0.030
(t-stat)	(3.18)	(9.07)	(8.26)	(7.15)	(4.74)	(0.21)

The table reports the summary statistics of the unexplained part of the equal-weighted portfolios sorted on β_{GDMA} excluding microcap firms. GDMA is the long-short value-weighted portfolio of geographically dispersed (GD) minus agglomerated (A) industries in the manufacturing sector. The t-statistics for the return spreads are reported in parentheses. "GD-A" is the return difference between the highest and lowest β_{GDMA}sorted portfolios. I report the unexplained part of the returns over the CAPM (α_{CAPM}) and Carhart (1997) 4-factor model ($\alpha_{4-Factor}$). The sample includes data from July 1947 to December 2018. A recession is a significant decline in economic activity which spread across the economy, lasting more than a few months, normally visible in real GDP, real income, employment, industrial production, and wholesale-retail sales as defined by NBER. Expansion periods are all the non-recession months.

Table 7: Two Way Portfolio Sorts

		Full Sample		Recessions		Expansions	
		GD-A Spread	(t-stat)	GD-A Spread	(t-stat)	GD-A Spread	(t-stat)
	Low	0.660	(2.86)	2.329	(2.79)	0.412	(1.77)
ROA	Med	0.159	(0.69)	1.704	(2.09)	-0.070	(-0.30)
	High	0.091	(0.39)	1.542	(1.94)	-0.124	(-0.51)
	Low-High	0.569	(2.72)	0.787	(1.17)	0.536	(2.45)
	Low	0.483	(2.14)	2.033	(2.69)	0.253	(1.09)
CF/A	Med	0.328	(1.41)	1.683	(2.14)	0.127	(0.53)
	High	0.037	(0.15)	1.476	(1.79)	-0.176	(-0.70)
	Low-High	0.446	(2.07)	0.557	(0.89)	0.429	(1.88)

The table reports summary statistics on 5 by 3 portfolios independently sorted on β_{GDMA} and ROA (or CF/A). The t-statistics for the return spreads are reported in parentheses. GD-A spread is the value-weighted returns on the long-short portfolio spread between the highest and the lowest β_{GDMA} quintile. ROA is the return on assets ratio and CF/A is the cashflows to assets ratio. A recession is a significant decline in economic activity which spread across the economy, lasting more than a few months, normally visible in real GDP, real income, employment, industrial production, and wholesale-retail sales as defined by NBER. Expansion periods are all the non-recession months. The sample includes data from July 1963 to December 2018.

Table 8: Estimated Parameters of the Regime-Switching Model

		Coeffiecient	(t-stat)	Tests on equality of coefficients across states
	α_{S1}	-0.017	(-3.49)	H0: α_{S1}= α_{S2}
State 1	σ_{S1}	0.043	(17.41)	χ^2= 65.89
	$\beta_{Mkt,t-1,S1}$	0.332	(4.46)	p-val = 0.00
	$\beta_{Mkt,t-2,S1}$	0.115	(1.35)	
	$\beta_{Mkt,t-3,S1}$	0.315	(4.18)	H0: $\beta_{Mkt,t-1,S1}$= $\beta_{Mkt,t-1,S2}$
				χ^2= 42.09
	α_{S2}	0.022	(7.97)	p-val = 0.00
State 2	σ_{S2}	0.030	(15.69)	
	$\beta_{Mkt,t-1,S2}$	-0.171	(-3.52)	H0: $\beta_{Mkt,t-2,S1} = \beta_{Mkt,t-2,S2}$
	$\beta_{Mkt,t-2,S2}$	-0.107	(-2.43)	χ^2= 4.32
	$\beta_{Mkt,t-3,S2}$	-0.170	(-3.03)	p-val = 0.04
	State Probabilities			H0: $\beta_{Mkt,t-3,S1} = \beta_{Mkt,t-2,S3}$
				χ^2= 26.29
	$\begin{bmatrix} p11 & p12 \\ p21 & p22 \end{bmatrix} = \begin{bmatrix} 0.48 & 0.52 \\ 0.35 & 0.65 \end{bmatrix}$			p-val = 0.04
				H0: σ_{S1}= σ_{S2}
				χ^2= 17.77
				p-val = 0.00

This table shows estimated parameters and corresponding t-statistics (in parentheses) of the regime-switching model in 5 . α_i,$\beta_{Mkt,i}$,and $\sigma_i$are the state-dependent intercepts, market betas, and the variance of residuals, for $i = 1, 2.$ $p_{i,j}$is the probability of being in state j in the current period given that the process was in state i in the previous period.

Table 9: Two Way Portfolio Sorts - High vs. Low Growth States

		Full Sample		Low Growth Periods		High Growth Periods	
		GD-A Spread	(t-stat)	GD-A Spread	(t-stat)	GD-A Spread	(t-stat)
	Low	0.660	(2.86)	3.449	(5.01)	0.115	(0.49)
ROA	Med	0.159	(0.69)	3.297	(4.90)	-0.454	(-1.95)
	High	0.091	(0.39)	2.948	(4.01)	-0.467	(-1.97)
	Low-High	0.569	(2.72)	0.501	(0.91)	0.582	(2.57)
	Low	0.483	(2.14)	3.418	(5.35)	-0.091	(-0.39)
CF/A	Med	0.328	(1.41)	3.642	(5.13)	-0.320	(-1.38)
	High	0.037	(0.15)	2.875	(3.85)	-0.517	(-2.08)
	Low-High	0.446	(2.07)	0.543	(0.98)	0.427	(1.83)

The table reports summary statistics on 5 by 3 portfolios independently sorted on β_{GDMA} and ROA (or CF/A). The t-statistics for the return spreads are reported in parentheses. GD-A spread is the value-weighted returns on the long-short portfolio spread between the highest and the lowest β_{GDMA} quintile. ROA is the return on assets ratio and CF/A is the cashflows to assets ratio. A recession is a significant decline in economic activity which spread across the economy, lasting more than a few months, normally visible in real GDP, real income, employment, industrial production, and wholesale-retail sales as defined by NBER. Expansions are defined as all months excluding recessions. The sample includes data from July 1963 to December 2018.

Table 10: Robustness: Geographic Agglomeration and Tradability

Model	α	(t-stat)
CAPM	0.174	(2.71)
Fama and French (1993) 3-factor model	0.164	(2.53)
Carhart (1997) 4-factor model	0.134	(2.03)
CAPM + Tian (2018) Tradability factor	0.120	(2.09)
CAPM + Durability factor	0.137	(2.23)
Pastor and Stambaugh (2003) liquidity model	0.190	(2.90)

The table reports summary statistics of the unexplained part of the GDMA portfolio returns controlling for various factor models. GDMA is the long-short value-weighted portfolio of geographically dispersed (D) minus agglomerated (A) industries in the manufacturing sector. The t-statistics are reported in parentheses. The sample includes data from January 1967 to December 2018.

Appendix

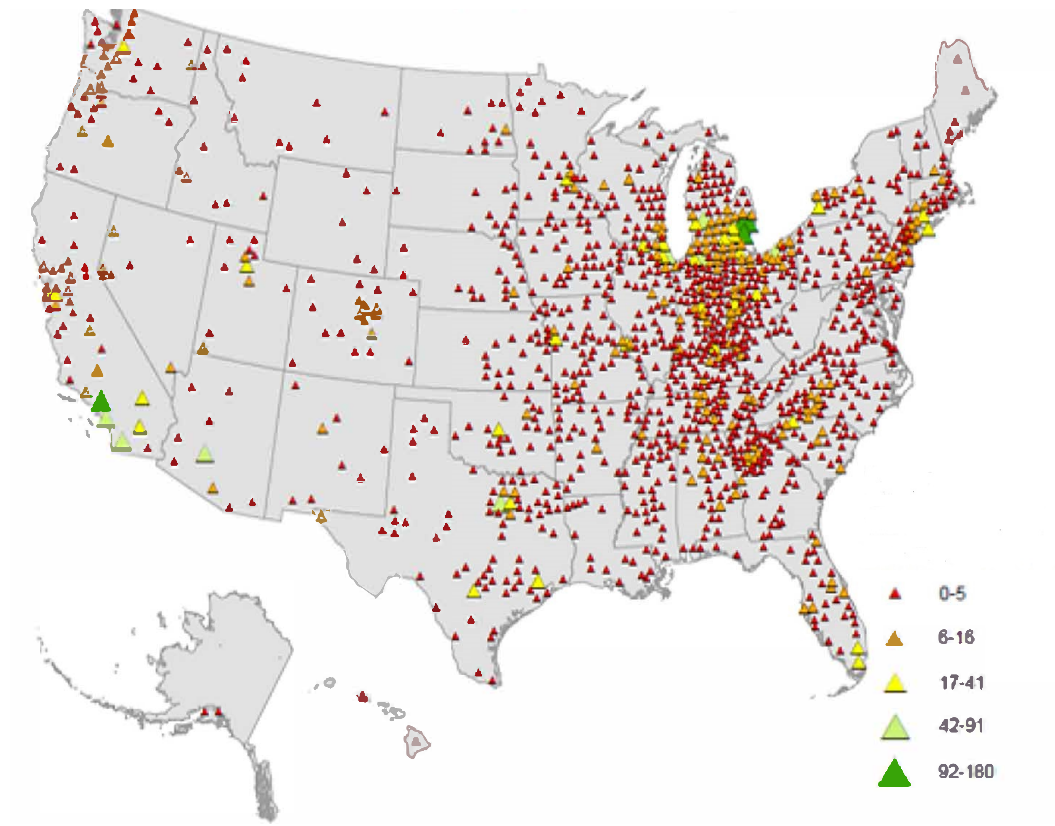

Figure A1: Locations of Motor Vehicle Parts Manufacturing

The figure shows the average motor vehicle parts manufacturing by county in 2010. The motor vehicle parts manufacturing industry is the largest sector of the motor vehicle and parts manufacturing industry. This industry consists of the manufacture of electrical and electronic equipment; engines and transmissions; brake systems; seating and interior trim; steering and suspension components; air-conditioners; and motor vehicle stampings, such as fenders, tops, body parts, trim, and molding. Figure is from the U.S. Bureau of Labor Statistics.

Table A1: Time-Series Average and Alphas - GDMA Portfolio

	GDMA Portfolio
Average Returns	0.085
(t-stat)	(1.65)
α_{CAPM}	0.127
(t-stat)	(2.48)
$\alpha_{3-\text{factor}}$	0.133
(t-stat)	(2.57)
$\alpha_{4-\text{factor}}$	0.109
(t-stat)	(2.05)

This table reports the sample average of the GDMA portfolio returns. I also report the unexplained part of the returns over the CAPM (α_{CAPM}), Fama and French (1993) 3-factor model ($\alpha_{3-Factor}$), and Carhart (1997) 4-factor model ($\alpha_{4-Factor}$). GDMA is the long-short value-weighted portfolio of geographically dispersed (D) minus agglomerated (A) industries in the manufacturing sector. The sample includes data from July 1947 to December 2018.The t-statistics are reported in parentheses.

Table A2: GDMA beta Sorts: Sub-Sample Test

	A	2	3	4	GD	GD-A
Panel A: Full Sample						
α_{CAPM}	-0.050	0.291	0.336	0.473	0.539	0.589
(t-stat)	(-0.36)	(3.18)	(5.05)	(6.73)	(6.11)	(3.13)
$\alpha_{4-\text{factor}}$	0.094	0.487	0.435	0.475	0.487	0.393
(t-stat)	(0.70)	(5.48)	(6.48)	(6.49)	(5.38)	(2.05)
Panel B: Recessions						
α_{CAPM}	-0.431	0.447	0.560	0.699	1.290	1.720
(t-stat)	(-0.96)	(1.86)	(2.28)	(3.36)	(4.25)	(2.69)
$\alpha_{4-\text{factor}}$	-0.316	0.521	0.540	0.714	1.157	1.473
(t-stat)	(-0.74)	(2.20)	(2.26)	(3.36)	(4.03)	(2.38)
Panel C: Expansions						
α_{CAPM}	0.004	0.266	0.315	0.439	0.412	0.408
(t-stat)	(0.03)	(2.67)	(4.76)	(5.86)	(4.59)	(2.11)
$\alpha_{4-\text{factor}}$	0.138	0.502	0.430	0.424	0.388	0.251
(t-stat)	(0.98)	(5.20)	(6.35)	(5.40)	(4.16)	(1.27)

The table reports the summary statistics of the unexplained part of the value-weighted portfolios sorted on β_{GDMA} for the sub-sample for which agglomeration can only be computed using $\beta'_{GDMA}s$. GDMA is the long-short value-weighted portfolio of geographically dispersed (GD) minus agglomerated (A) industries in the manufacturing sector. The t-statistics for the return spreads are reported in parentheses. “GD-A” is the return difference between the highest and lowest β_{GDMA}sorted portfolios. I report the unexplained part of the returns over the CAPM (α_{CAPM}) and Carhart (1997) 4-factor model ($\alpha_{4-Factor}$). The sample includes data from July 1947 to December 2018. A recession is a significant decline in economic activity which spread across the economy, lasting more than a few months, normally visible in real GDP, real income, employment, industrial production, and wholesale-retail sales as defined by NBER. Expansion periods are all the non-recession months.