News-Implied Linkages and Local Dependency in the Equity Market*

Shuyi Ge †, Shaoran Li ‡, Oliver Linton §

August 9, 2021

Abstract

This paper studies a heterogeneous coefficient spatial factor model that separately addresses both common factor risks (strong cross-sectional dependence) and local dependency (weak cross-sectional dependence) in the equity returns. From the asset pricing perspective, we derive the theoretical implications of no asymptotic arbitrage for the heterogeneous spatial factor model. In empirical work, it is challenging to measure granular firm-to-firm connectivity for a high-dimensional panel of equity returns. We use extensive business news to construct firms' links via which local shocks transmit, and we use those news-implied linkages as a proxy for the connectivity among firms. Empirically, we document a considerable degree of local dependency among $S\&P$ 500 stocks, and the spatial component does a great job in capturing the remaining correlations in the de-factored returns. We find that adding spatial interactions to factor models reduces mispricing and mean-squared errors. We also show that our news-implied linkages provide a comprehensive and integrated proxy for firm-to-firm connectivity, and it out-performs other existing networks in the literature.

Keywords: Spatial asset pricing model; weak and strong cross-sectional dependence; local dependency; networks; big data; large heterogeneous panel

JEL Classification: C33; C58; G10, G12

---

*The authors would like to thank Hashem Pesaran, Cheng Hsiao, Steven Kou, Donald Robertson, Emmanuel Guerre, Liudas Giraitis, Daniele Bianchi, Robin Lumsdaine, Dick van Dijk, Koen Jochmans, Julius Vainora, Michael Leung, Sampad Mohanty, Weiguang Liu, Chong Shu, Ekaterina Gavrilova, Marco Valerio Geraci, Merrick Li, Seok Young Hong, Chen Wang, Yimeng Xie, Lidan Tan, Grigory Franguridi, Chris Yoo and all the participants at the University of Cambridge Econometrics workshop and the University of California Econometrics workshop for their useful comments.

†Correspondence to the Faculty of Economics, University of Cambridge. Email address: sg751@cam.ac.uk.

‡The author is at the Faculty of Economics, University of Cambridge. Email address sl736@cam.ac.uk.

§The author is at the Faculty of Economics, University of Cambridge. Email address obl20@cam.ac.uk.

1 Introduction

Comovement in equity returns are the combined effects of exposures to common risks and local interactions. Classical asset pricing models such as the classical capital asset pricing model (CAPM) developed by Sharpe (1964), and the arbitrage pricing theory (APT) by Ross (1976) focus only on the strong pervasive component driven by a few common factors. Many studies have found those models focusing on strong dependence only are insufficient to capture all the significant interdependencies in asset returns. Local dependencies still play a non-negligible role (see for example Gabaix (2011), Acemoglu et al. (2012), Israelsen (2016), Barigozzi and Hallin (2017), Kou et al. (2018), Hale and Lopez (2019), Bailey et al. (2019a), and Barigozzi and Brownlees (2019)).

In this paper, we study a spatial factor model that separately addresses both common factor risks and local dependencies in the equity returns. The factor component and spatial component complement each other, with the former capturing strong cross-sectional dependence in equity returns and the latter capturing the weak cross-sectional dependence due to local interactions among entities. To distinguish the two sources of dependencies, imagine a group of people sitting in a room on a chilly winter day. People might catch a cold because the heater is broken (common factors) or because someone sitting close to them is ill (local interactions). The network architecture of entities, like the sitting plan in the previous example, is key to studying local interactions. Unlike spatial interactions in geographical systems, where there exists a natural network structure, for a high dimensional panel of equity returns, there is no natural network structure.

We use extensive business news data to construct firms' linkages via which local shocks transmit, and we use those news-implied linkages as a proxy for the connectivity among firms. It has been documented that common news coverage reveals information about linkages among companies, which are related to many economically important relationships like business alliances, partnerships, banking and financing, customer-supplier, and production similarity (Scherbina and Schlusche (2015), Schwenkler and Zheng (2019)). We use news data from RavenPack Equity files Dow Jones Edition for the period between the beginning of 2004 to the end of 2015. This comprehensive news dataset combines relevant content from multiple sources, including Dow Jones Newswires, Wall Street Journal, and Barron's MarketWatch, which produce the most actively monitored streams of news articles in the financial system. We identify linkages among firms by news co-mentioning.

Using the novel text-based network, we estimate spatial factor models with different sets of common risk factors. We find a considerable degree of local dependencies among $S\&P$ 500 stocks. The spatial interaction terms are highly significant after controlling for popular factors, and they continue to be significant even after adding industry-level factors. Different from most spatial econometrics literature, where spatial coefficients are assumed to be homogeneous, we adopt a heterogeneous coefficient framework from Bailey et al. (2016) and Aquaro et al. (2020). The model is very flexible, allowing us to capture general local interactions pattern among a large number of firms. Using that framework, we are able to not only investigate the average effect among all or some subgroups of firms but also gauge the individual-level effect. We document that apart from the average spatial effect measured by the mean group (MG) estimator being highly significant, at the individual level, spatial effect along news-implied linkages are also highly significant. We find that the percentage of individual contemporaneous spatial parameter being statistically significant at 5% level is over 88% across all specifications that we consider. This high significance ratio implies that the news-based link identification method is successful at detecting economically important links. The framework also allows us to examine the heterogeneity at subgroup levels. By applying mean group estimation to different industry groups, we document heterogeneity at the industry level. In particular, financial companies have the highest degree of local dependencies. We argue that the spatial factor model provides a unified way of addressing both strong and weak/local dependence in the equity returns. To investigate how well the spatial factor model captures the remaining dependence in the de-factored component, we examine the changes in correlation structure before and after adding the spatial component to the traditional factor models. We find that adding the spatial component reduces the number of non-zero pair-wise cross correlations by a huge margin, and the spatial factor model error correlation matrix is very close to diagonal. These results show that the spatial component constructed with news-implied linkages is successful at eliminating remaining correlations from the de-factored returns. We also compare the degree of mispricing and mean-squared errors for a set of factor models and their spatial augmented versions. We find that adding spatial/local interaction terms significantly reduces mispricing and mean-squared errors.

This paper contributes to three strands of literature. The first one is cross-sectional dependencies in equity returns. Cross-sectional dependence in a large panel is usually complex and reflects different types of interdependencies. Chudik et al. (2011), and Bailey et al. (2016) show that strong cross-sectional dependence (CSD) and weak cross-sectional dependence (CWD) have different economic implications and statistical behaviors, thus need to be accounted for separately. Kuersteiner and Prucha (2020) consider a short T panel with cross-sectional dependence due to both common factor risks and spatial/local interactions. While asset pricing literature has been focused on strong dependence (i.e., exposures to common risk factors), local dependence receives much less attention theoretically and empirically. Theoretically, we extend classical arbitrage pricing theory (APT), which only take strong cross-sectional dependence into account. We propose a flexible spatial factor model that addresses both strong and weak/local dependence for a large panel in a single framework. Especially, we derive the implications of no asymptotic arbitrage for our heterogeneous coefficient spatial factor model. Empirically, we show the benefit from addressing spatial/local interaction in terms of reducing mispricing errors and mean-squared errors.

Another major contribution is that we use a novel way to proxy the local connectivity among a large set of firms. Fan et al. (2011) suppose that the error covariance matrix is sparse (i.e., has lots of zeros), which represents the absence of linkages between firms beyond that contained in the common factors. They identify the location of non-zero entries by applying thresholding methods to the error sample covariance matrix. Our method uses information gathered from other sources, specifically news stories, to identify the linkages. There has been exploding research on quantifying the information embeded in unstructured data like text data. Alternative data fill the gaps in data availability induced by limited disclosure and slow update, thus complementing traditional economic datasets. For example, there has been a steep rise in the number of studies on utilizing the information from text (Garcia (2013), Scherbina and Schlusche (2015), Baker et al. (2016), Hoberg and Phillips (2016), Ke et al. (2019), Schwenkler and Zheng (2019), etc). This paper explores a comprehensive news dataset that combines relevant content from multiple sources and identifies linkages among firms by news co-mentioning. With a measure of local connectivity, we can capture correlations from both strong and the remaining weak dependence in a large panel using a single step. Without a knowledge of local connectivity, Fan et al. (2011) need a two-step procedure (they first estimate a factor model, and then use thresholding to estimate the error sample covariance matrix).

Our work also contributes to network effect or local risk spillover effect among economically linked firms.

Local risks transmit through economic linkages, and firms with links exhibit excess co-movement. There has been various proxies for firm to firm networks in the literature, including industry-based peers (Moskowitz and Grinblatt (1999), Fan et al. (2016), and Engelberg et al. (2018)), analyst co-coverage networks (Kaustia and Rantala (2013), Israelsen (2016), and Ali and Hirshleifer (2020)), customer-supplier networks (Cohen and Frazzini (2008)), geographic networks (Pirinsky and Wang (2006), Parsons et al. (2020)), etc. We show that our news-based linkages provide a comprehensive and integrated proxy for firm-level connectivity. Spatial factor model estimated with news-implied network out-performs those aforementioned networks in terms of minimizing the mispricing errors and the mean-squared errors. Even if we consider the union of all those competing networks, news-implied networks provide equally good performance. We also show that conditional on all those previously documented links, our news-implied linkages are still important channels of local risk spillovers.

The rest of the paper is organized as follows. Section 2 describes the difference between strong and weak cross-sectional dependence and introduces the spatial factor model. Section 3 develops the asset pricing implications with the presence of local interactions. Section 4 shows the estimation and inference of the heterogeneous coefficient spatial-temporal model that we use. Section 5 presents the estimation results, performances of alternative models, and comparisons with previously documented networks. Section 6 concludes. Proofs, technical details, and supplementary figures and tables are in the Appendices.

Notations: If $\{f_n\}_{n=1}^{\infty}$ and $\{g_n\}_{n=1}^{\infty}$ are both positive sequence of real numbers, then $f_n = \Theta(g_n)$ if there exist $N_0 \geq 1$ and positive finite constants C_0 and C_1, such that $inf_{n\geq N_0}(f_n/g_n) \geq C_0$, and $sup_{n\geq N_0}(f_n/g_n) \leq C_1$. For a $N\times N$ real matrix $A = (a_{ij})$, define its maximum column sum norm by $\|A\|_1 = \max_{1\leq j\leq N}\sum_{i=1}^{N} \mid a_{ij} \mid$, and its maximum row sum norm by $\|A\|_\infty = \max_{1\leq i\leq N}\sum_{j=1}^{N} \mid a_{ij} \mid$.

2 Modelling Cross-Sectional Dependence by Spatial Factor Model

2.1 Strong Dependence: Factor Model

Consider the factor model

$$\boldsymbol{r}_t - r_{ft}\mathbf{1} = \boldsymbol{\alpha} + \mathbf{B}\mathbf{f}_t + \boldsymbol{\epsilon}_t, t = 1,\ldots,T, \tag{1}$$

where $\boldsymbol{r}_t$ is the $N\times 1$ vector of equity returns at t, r_{ft} is the risk free rate at t, and $\mathbf{1}$ is $N\times 1$ vector of ones. $\mathbf{f}_t$ is the $K\times 1$ vector of common risk factors, and B is the $N\times K$ factor loadings, where β_{ik} is the loading of asset i on factor k. Let

$$\begin{aligned} &\sum_{i=1}^{N} \mid \beta_{ik} \mid = \Theta(N^{\alpha_{\beta_k}}), \ \text{ for } k = 1,\ldots,K, \\ &\|\mathbf{B}\|_1 = \max_{1\leq k\leq K}\sum_{i=1}^{N} \mid \beta_{ik} \mid = \Theta(N^{\alpha_{\mathbf{B}}}), \end{aligned} \tag{2}$$

where Θ denotes the exact order of magnitude, α_{β_k} measures how pervasive the kth factor is, and $\alpha_{\mathbf{B}} = \max_k(\alpha_{\beta_k})$ measures how pervasive the factor component $\mathbf{B}\mathbf{f}_t$ is. In the standard factor models, it is assumed that $\alpha_{\beta_k} > 0$ for $k = 1,\ldots,K$, and $\alpha_{\mathbf{B}} = 1$, so that each factor has non-diminishing effect on the system and exposures to common risk factors give rise to strong cross-sectional dependence, which is systematic and non-diversifiable.

2.2 Weak/Local Dependence: Spatial Model

Consider the canonical spatial autoregressive model with homogeneous spatial coefficient

$$\boldsymbol{r}_t = \boldsymbol{\alpha} + \psi W \boldsymbol{r}_t + \boldsymbol{\epsilon}_t, t = 1, \ldots, T, \tag{3}$$

where W is the $N \times N$ adjacency matrix that specifies the channels from which shocks transmit, where a typical entry w_{ij} gives the influence of the returns of j on that of i. The strength of spatial risk spillovers is represented by the parameter ψ.

Spatial dependence characterizes weak cross-sectional dependence where interactions are local. For demonstration, we re-write Equation 3 as follows:

$$\boldsymbol{r}_t = G(\psi)\boldsymbol{\nu}_t, \tag{4}$$

where $G(\psi) = (I_N - \psi W)^{-1}$ and $\boldsymbol{\nu}_t = \boldsymbol{\alpha} + \boldsymbol{\epsilon}_t$. Equation 4 can be interpreted as a factor model with N factors. $G(\psi)$ is the $N \times N$ factor loadings, where g_{ij} is the loading of i on factor j. All factors are weak and only have local effects if the following absolute summability condition is true

$$\sum_{i=1}^{N} \mid g_{ij} \mid \leq c \ \text{ for } j = 1, \ldots, N, \text{ where } c \text{ is a positive constant.} \tag{5}$$

The absolute summability condition Equation 5 is equivalent to a bounded column sum matrix norm condition on the Leontief inverse $G(\psi) = (I_N - \psi W)^{-1}$. As in LeSage (2008),

$$G(\psi) = (I_N - \psi W)^{-1} = I + \psi W + \psi^2 W^2 + ... = I + \sum_{j=1}^{\infty} \psi^j W^j. \tag{6}$$

The Leontief inverse take accounts of direct interaction effect and higher-order indirect effects. The assumption that the column sum of $G(\psi) = (I_N - \psi W)^{-1}$ is uniformly bounded in the number of cross-sectional units N is usually assumed in spatial econometrics (see Kelejian and Prucha (1998), Kelejian and Prucha (1999), Lee (2004)) to limit the cross-sectional correlation to a manageable degree. Some recent developments show that we may relax this assumption (Aquaro et al. (2020), Pesaran and Yang (2021)[1]). We take that assumption as a modelling assumption to distinguish strong and weak/local dependence. In particular, for weak dependence, no cross-sectional unit exerts pervasive effects on the system and the interactions are local. There will be more discussions in subsection 2.3.

2.3 Strong and Weak Dependence: Spatial Factor Model

Comovement in a large panel of equity returns arises due to both strong and weak cross-sectional dependence. Many studies have found that factor models that only focus on the non-diversifiable type of risks are insufficient to capture all the cross-sectional dependence in equity returns. We study a heterogeneous coefficient spatial factor model written as Equation 7 where the factor component and the spatial component complement each other, with the former addressing strong dependence and the latter addressing spillovers that are non-pervasive/local in nature (i.e., cross-sectional weak dependence (CWD) define in Chudik et al. (2011)).

$$\boldsymbol{r}_t = \boldsymbol{\alpha} + \mathbf{B}\mathbf{f}_t + \Psi W \boldsymbol{r}_t + \boldsymbol{\epsilon}_t, t = 1, \ldots, T, \tag{7}$$

where $\Psi = diag(\boldsymbol{\psi}) = diag(\psi_1, \ldots, \psi_N)$ is a diagonal matrix with N individual specific contemporaneous spatial coefficients on the main diagonals.

[1] They explicitly consider the case where there are dominant units that generate pervasive effect.

The spatial component has several main features. Firstly, the spillover coefficients are heterogeneous. One might reasonably suspect that the sensitivities to neighbours' risks are different from firm to firm. While the restrictive assumption that all entities share the same spatial coefficient is necessary for small T, it can be relaxed when T is big. Some recent work in the spatial literature (see LeSage and Chih (2018), Aquaro et al. (2020), and Chen et al. (2021)) consider heterogeneity in spatial parameters explicitly. We follow the framework from Aquaro et al. (2020), and they show that a heterogeneous spatial autoregressive model like Equation 7 can be consistently estimated with large T. We utilize this nice feature to explore the heterogeneity in the strength of local dependency. Moreover, we could examine the heterogeneity pattern at subgroup levels (such as industry levels) using mean-group estimation, which is a popular tool in heterogeneous panel literature. Secondly, it is possible to add weakly exogenous spatial-temporal terms $\sum_{l=1}^{L} \Psi_l W \boldsymbol{r}_{t-l}$ to Equation 7. $\Psi_l W \boldsymbol{r}_{t-l}$ corresponds to the spatial-temporal term at the lth lag for $l = 1, \dots, L$, where $\Psi_l = \text{diag}(\psi_l) = \text{diag}(\psi_{l,1}, \dots, \psi_{l,N})$ is a diagonal matrix of spatial-temporal parameters at the lth lag. These dynamic terms may account for potential market microstructure effects, which is important in our empirical application to daily individual stock returns.

Kou et al. (2018) consider a special case of this model, which they call the Spatial APT model. In particular, they consider the case where N is small, $L = 0$ (no temporal dynamics), $\psi_i = \psi$ (homogeneous spatial effects), and homoscedastic errors. Kou et al. (2018) derives the implications of the absence of arbitrage on the parameters of the model, in particular on the intercept vector $\boldsymbol{\alpha}$. In section 3, we extend their analysis by deriving the implications of no arbitrage under our framework.

In this paper, we impose the following assumptions on the spatial factor model (Equation 7):

Assumption 1 $E(\boldsymbol{\epsilon}_t) = 0,\ E(\mathbf{f}_t \boldsymbol{\epsilon}_t') = 0,\ E(\boldsymbol{\epsilon}_t \boldsymbol{\epsilon}_t') = \Omega,\ \sigma_i^2 = var(\epsilon_{it}) \leq \bar{\sigma}^2.$

Assumption 2 *Let* $\sum_{i=1}^{N} \mid \beta_{ik} \mid = \Theta(N^{\alpha_{\beta_k}}),$ *for* $k = 1, \dots, K,$ *and* $\|\mathbf{B}\|_1 = \max_{1 \leq k \leq N} \sum_{i=1}^{N} \mid \beta_{ik} \mid = \Theta(N^{\alpha_{\mathbf{B}}})$ *as Equation 2.* $\alpha_{\beta_k} > 0$ *for* $k = 1, \dots, K,$ *and* $\alpha_{\mathbf{B}} = 1.$

Assumption 3 $\sup_i \mid \psi_i \mid < \frac{1}{\|W\|_\infty},$ *for* $i = 1, \dots, N.$

Assumption 4 *Define* $G(\Psi) = (I_N - \Psi W)^{-1}$, *where* g_{ij} *is a typical entry of* G. $\sum_{i=1}^{N} \mid g_{ij} \mid \leq c$ *for* $j = 1, \dots, N$ *for some positive constant* c.

Remark 1: Assumption 2 guarantees that each factor has non-diminishing effect on the system, and the exposures to common risk factors give rise to strong cross-sectional dependence. We only assume that at least one factor is strong (i.e., with $\alpha = 1$), and all other factors are not weak. This is because in practice we may want to add many factors, which have different degrees of pervasiveness. For example, Bailey et al. (2020) find that for the factors proposed in the finance literature for asset pricing, only the market factor is strong over all the windows they consider.

Remark 2: Assumption 3 is to ensure $(I_N - \Psi W)$ is invertible and $G(\Psi) = (I_N - \Psi W)^{-1}$ exists.

Remark 3: Assumption 4 assumes that the column sums of $G(\Psi) = (I_N - \Psi W)^{-1}$ is uniformly bounded in absolute values as N goes to infinity. This ensures that no cross-sectional unit exerts pervasive effects on the system and the interactions are local. From our view, the correlation beyond factors should be weak. Pervasive dependence should be addressed by adding sufficient common factors into the model. Similar assumptions are made in Fan et al. (2008) and Fan et al. (2011), where they assume that after taking out the influence of Fama-French three factors, the remaining cross-sectional dependence is weak in the way defined in Chamberlain and Rothschild (1983). But be aware that this assumption a modelling assumption, and it is not required for stationarity and consistent estimation of the model (see Aquaro et al. (2020)).

The spatial factor model provides a unified way of addressing the remaining dependence in the de-factored component. Fan et al. (2011) identify the location of significant correlations by applying thresholding methods to the factor model error sample covariance matrix. To capture both factor-driven strong dependence and remaining weak dependence in a large panel, they need a two-step procedure. Our method provides an alternative, which can be achieved in a single step. Compared with purely statistical methods, our method also has the advantage of being interpretable given that our linkages are constructed using information from business news.

To investigate how well the spatial factor model captures the remaining dependence in the de-factored component, we can examine the changes in correlation structure before and after adding the spatial component to factor models. If the spatial component is doing a good job in terms of explaining the remaining local dependence, then we should expect to see the number of pairs with non-zero pair-wise error cross correlations being reduced by adding the spatial component. In our application, we estimate the number of non-zero pair-wise cross correlations of residuals from (1) a set of factor models, and (2) their corresponding spatial augmented models. For N cross-sectional units, the problem considers testing $N(N-1)/2$ null hypotheses simultaneously. We use multiple testing procedure to control for the overall size of the tests.

Under the factor model settings (Equation 1), this task is relatively easy. Pesaran et al. (2004) establishes the asymptotic distribution of error correlation coefficient under the null $H_{0,ij} : \rho_{ij} = 0$ for panel data models as follows:

$$y_{it} = \alpha_i + \boldsymbol{\beta}_i' \boldsymbol{x}_{it} + \epsilon_{it}, \tag{8}$$

where $Var(\boldsymbol{\epsilon}_t) = \Sigma = (\sigma_{ij})$ is an $N \times N$ symmetric, positive definite matrix. Denote the correlation coefficient of ϵ_{it} and ϵ_{jt} by ρ_{ij}. To estimate the correlation coefficient of errors, one needs to first obtain residuals $\hat{\epsilon}_{it}$ as

$$\hat{\epsilon}_{it} = y_{it} - \boldsymbol{x}_{it}' (\boldsymbol{X}_i' \boldsymbol{X}_i)^{-1} \boldsymbol{X}_i' \boldsymbol{y}_i, \tag{9}$$

where $\boldsymbol{X}_i$ is the $T \times K$ matrix of regressors for unit i, and $\boldsymbol{y}_i$ is the $T \times 1$ vector of the dependent variable for unit i. The sample estimate of ρ_{ij} is given by

$$\hat{\rho}_{ij} = \frac{\hat{\boldsymbol{\epsilon}}_i' \hat{\boldsymbol{\epsilon}}_j / T}{(\hat{\boldsymbol{\epsilon}}_i' \hat{\boldsymbol{\epsilon}}_i / T)^{1/2} (\hat{\boldsymbol{\epsilon}}_j' \hat{\boldsymbol{\epsilon}}_j / T)^{1/2}} \tag{10}$$

When $\boldsymbol{x}_{it}$ are strictly exogenous, under the null $H_{0,ij} : \rho_{ij} = 0$,

$$\sqrt{T} \hat{\rho}_{ij} \to N(0,1) \text{ as } T \to \infty. \tag{11}$$

To test the hypothesis $H_{0,ij} : \rho_{ij} = 0$, for p being the chosen nominal size, we can use $\frac{1}{\sqrt{T}} \Phi^{-1}(1 - p/2)$ as the critical value, where Φ^{-1} is the inverse cdf of standard normal. However, to test $\rho_{ij} = 0$ for all $i \neq j$ jointly, we need to take the multiple testing issue into account. From Bonferroni (1935), given that there are $n_{test} = N(N-1)/2$ such tests, for the family-wise error rate (FWER) to be p, it is sufficient to set the nominal size for each individual test $p_i = p/n_{test}$ for $i = 1, \ldots, n_{test}$, so that the critical value for each $| \hat{\rho}_{ij} |$ becomes $\frac{1}{\sqrt{T}} \Phi^{-1}(1 - p/2n_{test})$.[2]

[2]There are more advanced methods of choosing threshold values, like Bailey et al. (2019b). However, the theory does not go through for testing error correlation of the spatial factor model. To have a fair comparison, we consider Bonferroni type of correction for both factor and spatial factor model.

The nice theoretical result (Equation 11) is derived under the assumptions of exogeneity of regressors. However, the spatial factor model this is not the case. For the spatial factor model we consider (Equation 7), the spatial autoregressive term $W\boldsymbol{r}_t$ is endogenous, which makes the result from Equation 11 fail. Given that, we conduct bootstrap inference for the error correlations of the spatial factor model, which proceeds in the following steps.

We first estimate the spatial factor model (Equation 7) and collect the estimated parameters $\hat{\boldsymbol{\alpha}}, \hat{\mathbf{B}}, \hat{\Psi}$. We save the residual values $R = \{\hat{\epsilon}_{11}, \dots, \hat{\epsilon}_{1T}, \dots, \hat{\epsilon}_{N1}, \dots, \hat{\epsilon}_{NT}\}$. In the next step, for $b = 1, \dots, B$, under the null of diagonal error correlation matrix, we draw $i.i.d$ $\hat{\epsilon}^b_{it}$ from R, and generate the bth bootstrap sample as

$$\boldsymbol{r}^b_t = (I - \hat{\Psi}W)^{-1}(\hat{\boldsymbol{\alpha}} + \hat{\mathbf{B}}\mathbf{f}_t + \boldsymbol{\epsilon}^b_t), t = 1, \dots, T. \tag{12}$$

We re-estimate the model using the bootstrap sample. Next, we calculate the sample correlation coefficients $\hat{\rho}^b_{ij}$ for all $i \neq j$. We save those $N(N-1)/2$ pair-wise cross-correlations for each bootstrap sample b. Finally, we can draw inference from the empirical null distribution F by computing the critical values associated with a nominal size value p as $F^{-1}(p/2)$ and $F^{-1}(1-p/2)$. Again, we need to correct for multiple testing issue here, and the critical values for each $\hat{\rho}_{ij}$ becomes $F^{-1}(p/2n_{test})$ and $F^{-1}(1-p/2n_{test})$.

3 Arbitrage Pricing Theory Under Spatial Factor Model

In this section, we derive the asset pricing implications of our heterogeneous coefficient spaital factor model (Equation 7). We follow Ingersoll Jr (1984), and consider a fixed infinite economy where a sequence of nested subsets of assets are examined. For the nth economy, a new asset is added to the $(n-1)$th economy.

$$\boldsymbol{r}^{n\prime} = (\boldsymbol{r}^{n-1\prime}, r^n). \tag{13}$$

Since APT is a cross-sectional approach, we drop the time subscript. We denote the size of economy by superscript n, and a portfolio in the nth economy is denoted as $\boldsymbol{c}^n \in R^n$. $\mathbf{1}^n$ is the vector of n ones. We consider subsequences of assets, where subsequences are indexed by v. There are asymptotic arbitrage opportunities if there is a subsequence of portfolios that satisfy the following conditions:

$$\begin{aligned} &Var(\boldsymbol{c}^{v\prime}\boldsymbol{r}^{\boldsymbol{v}}) \to 0 \;\; as \quad v \to \infty, \\ &E(\boldsymbol{c}^{v\prime}\boldsymbol{r}^{\boldsymbol{v}}) \geq \delta > 0 \;\text{ for all } v, \\ &\boldsymbol{c}^{v\prime}\mathbf{1}^v = 0 \;\text{ for all } v. \end{aligned} \tag{14}$$

Theorem 1 *Assume that the returns are generated by the heterogeneous spatial factor model (Equation 7), and Assumption 1-4 hold. If there is no arbitrage opportunities described in Equation 14, then there is a sequence of K by 1 vector of factor premiums $\boldsymbol{\lambda}^n$ and a constant λ^n_0 such that the following approximation holds*

$$\boldsymbol{\alpha}^n \approx (I_n - \Psi^n W^n)\mathbf{1}^n\lambda^n_0 + \mathbf{B}^n\boldsymbol{\lambda}^n. \tag{15}$$

Define pricing error vector as:

$$\boldsymbol{v}^n = \boldsymbol{\alpha}^n - (I_n - \Psi^n W^n)\mathbf{1}^n\lambda^n_0 - \mathbf{B}^n\boldsymbol{\lambda}^n. \tag{16}$$

The approximation Equation 15 holds in a sense that there is a positive number V such that the weighted sum of squared pricing errors is uniformly bounded,

$$(\boldsymbol{v}^n)'(\Omega^n)^{-1}(\boldsymbol{v}^n) \leq V < \infty \;\; \textit{for all } n. \tag{17}$$

Corollary 1.1 *Let λ_{max} denote the largest eigenvalue of the limit covariance matrix Ω^n, then $V = q\lambda_{max}$ for a positive number q.*

Corollary 1.2 *Suppose all factors are traded. The risk factor f_k before de-meaning is denoted as $\tilde{f}_k$, which is the payoff of the kth tradable zero-cost portfolio, where*

$$f_k = \tilde{f}_k - E(\tilde{f}_k) \ \ for\ k = 1, \ldots, K. \tag{18}$$

If there exists a risk free asset with rate r_f, the spatial factor model (Equation 7) can be written as:

$$\begin{aligned}
&\tilde{\boldsymbol{r}} = \tilde{\boldsymbol{\alpha}} + \mathbf{B}\tilde{\mathbf{f}} + \Psi W\tilde{\boldsymbol{r}} + \boldsymbol{\epsilon}, \\
&where\ \ \tilde{\boldsymbol{r}} = \boldsymbol{r} - r_f\mathbf{1}\ \ is\ the\ vector\ of\ excess\ returns, \\
&\tilde{\mathbf{f}} = (\tilde{f}_1, \ldots, \tilde{f}_K)', \\
&\tilde{\boldsymbol{\alpha}} = \boldsymbol{\alpha} - (I - \Psi W)\mathbf{1}r_f - \mathbf{B}E(\tilde{\mathbf{f}}).
\end{aligned} \tag{19}$$

Then the no arbitrage condition for an infinite economy where asset returns are generated by the spatial factor model is:

$$\tilde{\boldsymbol{\alpha}}^n \approx 0. \tag{20}$$

$\tilde{\boldsymbol{\alpha}}^n$ is the pricing error in this special case. The approximation Equation 20 holds in a sense that there is a positive number V such that the sum of squared pricing errors is uniformly bounded,

$$(\tilde{\boldsymbol{\alpha}}^n)'(\Omega^n)^{-1}(\tilde{\boldsymbol{\alpha}}^n) \leq V < \infty \ \ for\ all\ n. \tag{21}$$

And λ_0^n, $\boldsymbol{\lambda}^n$ in theorem 1 can be identified as

$$\begin{aligned}
\lambda_0^n &= r_f, \\
\boldsymbol{\lambda}^n &= E(\tilde{\mathbf{f}}).
\end{aligned} \tag{22}$$

Corollary 1.3 *In addition to the assumptions from corollary 1.2 (there exists risk free assets and all factor are traded), if we further assume that errors are uncorrelated, then for any $\delta > 0$, there is a constant N_δ such that the number of elements in $\tilde{\boldsymbol{\alpha}}$ that are bigger than δ in absolute values is uniformly bounded by N_δ,*

$$\lim_{n\to\infty} \sum_{j=1}^{n} I(\mid \tilde{\alpha}_j \mid > \delta) < N_\delta < \infty \tag{23}$$

Remark 1: corollary 1.1 implies that the correlation structure of $\Omega = E(\boldsymbol{\epsilon}_t\boldsymbol{\epsilon}_t')$ affects the bound V. The less correlation there in Ω, the smaller V is, and the better the approximation implied by Equation 15 is. Spatial factor model addresses weak/local dependence beyond strong dependence captured by factors, and we expect that there is less correlation in Ω. Compared with factor models, the spatial factor model is expected to give a better approximation.

Remark 2: corollary 1.2 points out a special case useful for empirical work. Assuming that there exists risk free asset and all factors are traded, we can just look at the spatial factor model with the dependent variable being the excess returns (Equation 19). In particular, to test the theoretical restrictions we just need to examine how close to zero the intercept vector $\tilde{\boldsymbol{\alpha}}$ is.

Remark 3: theorem 1 and the corollaries suggest some statistics that we can employ to compare the relative performance of different asset pricing models. In particular, the L_1, L_2 norms of mispricing errors, and the number of components with big mispricing errors could be useful in measuring how well the approximation is.

Remark 4: theorem 1 and the corollaries can be easily extended to spatial factor models with more than one spatial spillover channels, for example, the two-W model in Bailey et al. (2016).

Proofs of theorem 1, corollary 1.1, corollary 1.2, and corollary 1.3 are in the section A of the Appendix.

4 Estimation and Inference

The corollary 1.2 points out a special case useful for empirical work. From now on we assume there is risk free asset, and all factors are traded, and we work with the panel spatial factor model where the dependent variable is the excess returns (multi-period Equation 19).

$$\tilde{\boldsymbol{r}}_t = \tilde{\boldsymbol{\alpha}} + \mathbf{B}\tilde{\mathbf{f}}_t + \Psi W \tilde{\boldsymbol{r}}_t + \boldsymbol{\epsilon}_t, t = 1, \ldots, T, \tag{24}$$

where $\tilde{\boldsymbol{r}}_{\boldsymbol{t}} = (\tilde{r}_{1t}, \ldots, \tilde{r}_{Nt})$ is the N by 1 vector of excess returns at t, and $\tilde{\boldsymbol{f}}_{\boldsymbol{t}} = (\tilde{f}_{1t}, \ldots, \tilde{f}_{Kt})'$ is the vectors of K factors at t. W is the N by N adjacency matrix, where a typical entry is denoted as w_{ij}. Without loss of generality , we set $w_{ii} = 0$ for all i, and we assume that $w_{ij} \geq 0$. Following the convention in spatial econometrics, we further row normalize W so that $\sum_{j=1}^{N} w_{ij} = 1$ for all i. Here, $\boldsymbol{\epsilon}_{\boldsymbol{t}}$ is the vector of errors at t, which satisfies var$(\boldsymbol{\epsilon}) = \Sigma = \text{diag}(\boldsymbol{\sigma}_{\epsilon^2}) = \text{diag}(\sigma_1^2, \ldots, \sigma_N^2)$, where σ_i^2 =var(ϵ_{it}) is allowed to be heteroskedastic.

An extension of Equation 24 is to incorporate weakly exogenous spatial-temporal terms:

$$\tilde{\boldsymbol{r}}_t = \tilde{\boldsymbol{\alpha}} + \mathbf{B}\tilde{\mathbf{f}}_t + \Psi_0 W \tilde{\boldsymbol{r}}_t + \sum_{l=1}^{L} \Psi_l W \tilde{\boldsymbol{r}}_{t-l} + \boldsymbol{\epsilon}_t, t = 1, \ldots, T, \tag{25}$$

where we denote the contemporaneous spatial coefficients using Ψ_0, and L is the number of spatial-temporal terms to incorporate. In what follows, we set $L = 5$ to control for within-week dynamics. These dynamic terms may account for potential market microstructure effects, which is important in our empirical application to daily individual stock returns. This modification has been used in Eugene (1992), see also Dimson (1979).

There are mainly two classes of methods that have been developed in the literature to estimate spatial models: the maximum likelihood method (Lee (2004), Lee and Yu (2010), Shi and Lee (2017), and Aquaro et al. (2020), among others), and the IV/GMM approach (Kelejian and Prucha (1998), Kelejian and Prucha (1999), Lee (2007), and Kuersteiner and Prucha (2020) among others). In this paper, we estimate the heterogeneous coefficient spatial-temporal model (Equation 25) by the QML procedure proposed in Bailey et al. (2016) and Aquaro et al. (2020). We collect all the parameters in Equation 25 in the $(N * (K + L + 3))$ by 1 vector $\boldsymbol{\theta} = (\tilde{\boldsymbol{\alpha}}', \boldsymbol{\beta}_1', \ldots, \boldsymbol{\beta}_K', \boldsymbol{\psi}_0', \ldots, \boldsymbol{\psi}_L', \boldsymbol{\sigma}_{\epsilon^2}')'$, and denote the vector of true values by $\boldsymbol{\theta}_0$. The log-likelihood function of Equation 25 is written as follows:

$$\mathcal{L}_T(\boldsymbol{\theta}) = -\frac{NT}{2}\ln(2\pi) - \frac{T}{2}\sum_{i}^{N}\ln(\sigma_i^2) + \frac{T}{2}\ln \mid \boldsymbol{S}'(\boldsymbol{\psi_0})\boldsymbol{S}(\boldsymbol{\psi_0}) \mid - \frac{1}{2}\sum_{t=1}^{T}[\boldsymbol{S}(\boldsymbol{\psi_0})\tilde{\boldsymbol{r}}_{\boldsymbol{t}} - \boldsymbol{B}\boldsymbol{x}_{\boldsymbol{t}}]'^{-1}[\boldsymbol{S}(\boldsymbol{\psi_0})\tilde{\boldsymbol{r}}_{\boldsymbol{t}} - \boldsymbol{B}\boldsymbol{x}_{\boldsymbol{t}}], \tag{26}$$

where $\boldsymbol{S}(\boldsymbol{\psi_0}) = I_N - \Psi_0 W$, and $\tilde{\boldsymbol{r}}_{\boldsymbol{t}} = (\tilde{r}_{1t}, \ldots, \tilde{r}_{Nt})$. We stack the constant and all the weakly exogenous variables for i at t in $x_{it} = (1, \tilde{f}_{1t}, \ldots, \tilde{f}_{Kt}, \sum_{j=1}^{N} w_{ij}\tilde{r}_{jt-1}, \ldots, \sum_{j=1}^{N} w_{ij}\tilde{r}_{jt-L})$, and $\boldsymbol{x}_{\boldsymbol{t}} = (x_{1t}', \ldots, x_{Nt}')'$ is the $((1 + K + L) * N)$ by 1 vector. $\boldsymbol{B}$ is the N by $((1 + K + L) * N)$ block diagonal matrix with elements $\boldsymbol{b}_i' = (\tilde{\alpha}_i, \beta_{1,i}, \ldots, \beta_{K,i}, \psi_{1,i}, \ldots, \psi_{L,i})'$ on the main diagonal and zeros elsewhere.

The quasi maximum likelihood estimator $\hat{\boldsymbol{\theta}}_{QMLE}$ maximizes Equation 26. The error terms need not be Gaussian, but when they are, $\hat{\boldsymbol{\theta}}_{QMLE}$ is the maximum likelihood estimator of $\boldsymbol{\theta}$. Note that conditional on $\boldsymbol{\psi_0}$, the system is linear, so that we can concentrate out the parameters $\boldsymbol{B}, \boldsymbol{\sigma}_{\epsilon^2}$ to reduce the dimensionality and hence computational burden.

Aquaro et al. (2020) provides sufficient conditions for consistency and asymptotic normality of $\hat{\boldsymbol{\theta}}_{QMLE}$ in the case where T is large and N is fixed, and consistency and asymptotic normality for the mean group estimators in the case where both T and N are large but $\sqrt{N}/T \to 0$. Further details regarding the identification conditions and inference can be found in the section B of the appendix.

5 Data

We consider daily returns of $S\&P$ 500 stocks for our application. All the stock market related data are from the Center for Research in Security Prices (CRSP). Daily factor returns and industry classification information are obtained from Kenneth French's website. Accounting data are from the merged CRSP/Compustat database. Data used to construct alternative networks are described in details in subsection 6.3.

The news data are obtained from RavenPack Equity files Dow Jones Edition for the period January 2004 to December 2015. This comprehensive news dataset combines relevant content from multiple sources, including Dow Jones Newswires, Wall Street Journal, and Barron's MarketWatch, which produce the most actively monitored streams of news articles in the financial system. Each unique news story (identified by unique story ID) tags the companies mentioned in the news by their unique and permanent entity identifier codes (RP_ENTITY_ID), by which we link to stock identifier TICKER and PERMNO.

Inspired by Scherbina and Schlusche (2015) and Schwenkler and Zheng (2019), we identify links by news co-mentioning. To be more specific, if a piece of business news reports two and only two companies together, then the two firms have some business relationship/link. Although news that mentions more than two companies together may carry potential information about links, they provide noisier information. We also remove news with topics including analyst recommendations, rating changes, and index movements as these types of news might stack multiple companies together when they actually do not have real links. Table 10 provides descriptive statistics for RavenPack Equity files Dow Jones Edition dataset during the sample period. Since our comprehensive news dataset combines several sources, given a similar length of sample period, the number of unique news stories is more than ten times larger than that from Scherbina and Schlusche (2015) and more than eight hundred times than that from Schwenkler and Zheng (2019). For link identification purposes, we only use sample news (1) are not about topics mentioned above, (2) tag $S\&P$ 500 companies, and (3) mention exactly two companies, which is a subsample of $1,637,256$ unique news stories.

From all the links identified using this methodology, some links are transitory while some are more long-lasting. To gauge the persistency of links, we split full sample news data into 12 yearly link identification windows. Table 11 is the frequency distribution table of the number of yearly link identification windows that a pair gets identified as economic neighbours for all possible pairs (i, j) in our sample. 72.80% of the pairs never get co-mentioned during the sample period. For all the linked pairs (i, j) identified throughout the sample period, 49.6% of them are only mentioned in one yearly window. We consider those pairs as temporarily linked. They could get co-mentioned multiple times within a yearly window. But out of that one-year window, they are never mentioned together. To further reduce noise, we say a pair (i, j) has persistent economic relationships if they are identified in more than a certain number $(1 \leq m \leq 11)$ of yearly identification windows. For the construction of full sample adjacency matrix W, we set w_{ij} to the number of times i and j are co-mentioned throughout the sample if the pair (i, j) gets co-mentioned in more than m yearly identification windows (i.e., their link is persistent), and to zero otherwise.

Table 12 presents the number of identified pairs aggregated at industry level for threshold $m = 1$. Results for higher threshold values are shown in Table 13, and Table 14. We classify stocks into Fama-French 12 industries based on their Standard Industrial Classification (SIC) code. Compared with companies from other industries, financial companies, hi-tech companies, and manufacturing companies are more connected. Another important feature is that there are a lot of intra-industry links. Except for some industries with very few stock like Durables and Telecommunication, whose statistics should be interpreted with care, other industries all have a high percentage of intra-industry links. Comparing tables of adjacency matrices with different threshold values m, we can tell that although higher threshold values reduce the absolute number of identified pairs, the relative industry level network remains very robust.

6 Results

6.1 Main Results

For full sample estimation, we keep $S\&P$ 500 stocks that have no missing observations for the period 2004 to 2015, which leaves us $N = 394$ stocks. Adjacency matrix W contains all the persistent links (for different thresholds m) identified throughout the sample. As a convention in spatial econometrics, we then apply row-normalization to W so that $\sum_j^N w_{ij} = 1$ for all $i = 1, \ldots, N$. We investigate several models under the general framework Equation 25:

- Model 1: Spatial CAPM model

$$\tilde{\boldsymbol{r}}_t = \tilde{\boldsymbol{\alpha}} + \boldsymbol{\beta_1} f_{MRT,t} + \sum_{l=0}^{L} \Psi_l W \tilde{\boldsymbol{r}}_{t-l} + \boldsymbol{\epsilon}_t. \tag{27}$$

- Model 2: Spatial factor model with Fama-French three factors

$$\tilde{\boldsymbol{r}}_t = \tilde{\boldsymbol{\alpha}} + \boldsymbol{\beta_1} f_{MRT,t} + \boldsymbol{\beta_2} f_{SMB,t} + \boldsymbol{\beta_3} f_{HML,t} + \sum_{l=0}^{L} \Psi_l W \tilde{\boldsymbol{r}}_{t-l} + \boldsymbol{\epsilon}_t. \tag{28}$$

- Model 3: Spatial factor model with Fama-French five factors

$$\tilde{\boldsymbol{r}}_t = \tilde{\boldsymbol{\alpha}} + \boldsymbol{\beta_1} f_{MRT,t} + \boldsymbol{\beta_2} f_{SMB,t} + \boldsymbol{\beta_3} f_{HML,t} + \boldsymbol{\beta_4} f_{RMW,t} + \boldsymbol{\beta_5} f_{CMA,t} + \sum_{l=0}^{L} \Psi_l W \tilde{\boldsymbol{r}}_{t-l} + \boldsymbol{\epsilon}_t. \tag{29}$$

- Model 4: Spatial factor model with Fama-French five factors plus Momentum factor

$$\tilde{\boldsymbol{r}}_t = \tilde{\boldsymbol{\alpha}} + \boldsymbol{\beta_1} f_{MRT,t} + \boldsymbol{\beta_2} f_{SMB,t} + \boldsymbol{\beta_3} f_{HML,t} + \boldsymbol{\beta_4} f_{RMW,t} + \boldsymbol{\beta_5} f_{CMA,t} + \boldsymbol{\beta_6} f_{MOM,t} + \sum_{l=0}^{L} \Psi_l W \tilde{\boldsymbol{r}}_{t-l} + \boldsymbol{\epsilon}_t. \tag{30}$$

The parameters $(N * (K + L + 3))$ in Equation 25 are estimated using quasi maximum likelihood (QML). Given the huge amount of parameters in the model, here we only report some important summary statistics of the estimates in Table 1.[3] For a heterogeneous coefficient panel model, what is often of interest to empirical researchers is the average estimates across all entities (or all entities within a sub-group). If we assume that individual-specific coefficients are randomly distributed around their common means as follows:

$$\begin{aligned} &\beta_{k,i} = \lambda_k + \zeta_{k,i}, \psi_{l,i} = \psi_l + \varsigma_{l,i} \quad \text{for } k = 1, \ldots, K,\ l = 1, \ldots, L, \text{ and } i = 1, \ldots, N. \\ &\boldsymbol{\eta}_i = (\boldsymbol{\zeta}'_i, \boldsymbol{\varsigma}'_i)' \sim IID(\boldsymbol{0}, \Omega_\eta). \end{aligned} \tag{31}$$

[3] Full estimation results can be requested from the author.

The common mean parameters β_k and ψ_l for $k = 1, \ldots, K$, $l = 1, \ldots, L$ are the the objects of interest and they can be consistently estimated with the following mean group (MG) estimator given N and T are large, with $\sqrt{N}/T \to 0$. [4].

$$\hat{\beta}_{k1}^{MG} = \frac{1}{N}\sum_{i=1}^{N} \hat{\beta}_{k1,i} \;\; and \;\; \hat{\psi}_{l}^{MG} = \frac{1}{N}\sum_{i=1}^{N} \hat{\psi}_{l,i}. \tag{32}$$

For a heterogeneous coefficient spatial model, we can only identify the spatial coefficients of those units with at least one link. Spatial coefficients of those units with zero link cannot be identified, and we need to restrict them to be zeros. If we apply the full sample adjacency matrix W discussed above with threshold value $m = 1$, only $N_0 = 7$ out of $N = 394$ companies do not have any long-run links. $N_p = N - N_0 = 387$ units have unrestricted spatial coefficients. In contrast, individual-specific factor coefficients and intercepts are identified for all units, with $N_p = N = 394$.

We estimate Model 1 - Model 4 over the full sample period. We report the mean group (MG) estimates with their standard errors and the percentages of companies with statistically significant parameters at 5% level for models estimated with threshold value $m = 1$ in Table 1. Results for alternative thresholds m are reported in Table 15 and Table 16.

	(1) factor component							(2) spatial-temporal component					
	α	β_1	β_2	β_3	β_4	β_5	β_6	ψ_0	ψ_1	ψ_2	ψ_3	ψ_4	ψ_5
(1) Spatial CAPM													
MG	0.015	0.564						0.446	0.002	-0.008	0.001	-0.003	0.004
	(0.001)	(0.022)						(0.020)	(0.003)	(0.002)	(0.002)	(0.002)	(0.001)
%sig	9.1%	90.4%						89.9%	51.9%	28.2%	20.7%	30.2%	21.7%
N_p	394	394						387	387	387	387	387	387
(2) Spatial factor model (Fama-French three factors)													
MG	0.013	0.529	0.129	-0.137				0.489	0.002	-0.008	-0.001	-0.001	0.003
	(0.001)	(0.021)	(0.014)	(0.022)				(0.019)	(0.003)	(0.002)	(0.001)	(0.002)	(0.001)
%sig	6.1%	86.0%	75.1%	82.7%				89.7%	54.0%	27.1%	21.4%	29.7%	18.9%
N_p	394	394	394	394				387	387	387	387	387	387
(3) Spatial factor model (Fama-French five factors)													
MG	0.011	0.544	0.144	-0.137	0.140	0.179		0.493	0.007	-0.007	-0.001	-0.002	0.002
	(0.001)	(0.021)	(0.014)	(0.023)	(0.022)	(0.021)		(0.019)	(0.003)	(0.002)	(0.001)	(0.002)	(0.001)
%sig	5.6%	86.8%	73.9%	82.5%	74.6%	73.6%		89.1%	52.2%	27.6%	22.2%	27.9%	18.6%
N_p	394	394	394	394	394	394		387	387	387	387	387	387
(4) Spatial factor model (Fama-French five factors plus Momentum)													
MG	0.015	0.471	0.124	-0.116	0.074	0.113	-0.010	0.545	0.002	-0.014	0.002	-0.008	0.005
	(0.001)	(0.022)	(0.014)	(0.021)	(0.022)	(0.021)	(0.007)	(0.019)	(0.003)	(0.002)	(0.001)	(0.002)	(0.001)
%sig	4.8%	86.5%	73.3%	79.2%	75.1%	74.1%	59.4%	88.1%	52.7%	27.4%	23.3%	27.9%	18.6%
N_p	394	394	394	394	394	394	394	387	387	387	387	387	387

Table 1: **QML estimation results of Equation 27 to Equation 30 using full sample.**

Note: threshold $m = 1$. For each panel , the first row gives the mean group (MG) estimates for each parameter with their standard errors in the parenthesis. The row of each panel gives the percentages of unrestricted units with statistically significant parameters at 5% level, and the last row gives the number of unrestricted units N_p for each parameter.

[4]See Pesaran and Smith (1995) for proofs of the consistency when individual-specific coefficients are independently distributed. Recent development by Chudik and Pesaran (2019) proves the consistency under weakly correlated individual-specific estimators. In both cases, T and N are required to be large with $\sqrt{N}/T \to 0$. Intuitively, big T is required for the consistent estimation of individual-specific coefficients, and N needs to be big enough for the consistent estimation of the means. To see how the MG estimators behave in the context of heterogeneous spatial-temporal model, see Aquaro et al. (2020)

Contemporaneous local dependency parameter ψ_0 is highly statistically significant under all specifications. Among 387 unrestricted contemporaneous spatial coefficients $\psi_{0,i}$, more than 88% of them are individually significant under all cases. This high significance ratio implies that the new-implied linkage identification method is very successful at detecting relevant links. If our data contain a lot of spurious links, we will be more likely to see the spatial parameters to be insignificant for many individuals. Local dependencies also exhibit strong economic importance: the mean group (MG) estimates of ψ_0 are around $0.45-0.55$ over the four models we consider, which is comparable to the average strength of the market factor, with the mean group (MG) estimates of market beta lying between $0.47-0.56$ across models.

Dynamic spatial dependency terms are also statistically significant, although smaller in economic magnitude. For the first lag ψ_1, there are more than 50% of $\psi_{1,i}$ are individually significant under all cases. For further lags, there are always more than 20% of $\psi_{l,i}$ $(l = 2, \ldots, 5)$ are individually significant.

Adding more common risk factors does not weaken local dependencies. The magnitudes of mean group estimates and the percentages of companies with statistically significant parameters at 5% level do not change with the number of factors we include. Interestingly, the magnitude of average contemporaneous local dependency captured by ψ_0^{MG} is the largest for Model 4, while the magnitude of average exposures to market factor is the smallest.

A large proportion of the news-implied links that we identify are intra-industry links. It has been documented widely that stocks within the same industry exhibit excess co-movement beyond common risk factors at market level (Moskowitz and Grinblatt (1999), Fan et al. (2016), Engelberg et al. (2018)). In order to control for industry factors as an additional source of co-movement, we further augment Equation 27 to Equation 30 with industry factors.

$$\tilde{\boldsymbol{r}}_t = \tilde{\boldsymbol{\alpha}} + \mathbf{B}\tilde{\mathbf{f}}_\mathbf{t} + \boldsymbol{\beta}_I \tilde{f}_{IND,t} + \sum_{l=0}^{L} \Psi_l W \tilde{\boldsymbol{r}}_{t-l} + \boldsymbol{\epsilon}_t. \tag{33}$$

We use Fama French 12 equal-weighted industry portfolios. We choose to use broad industry classification and equal weighting. This is because we are dealing with $S\&P$ 500 stocks, and we do not want industry returns to be dominated by several large stocks within that industry.

Table 2 reports the results for models with industry factors. Industry factors are highly significant in all cases, and the mean group (MG) estimates of industry beta are between $0.41-0.45$. The introduction of the industry factor largely weakens the effect of the market factor, with the average market beta being reduced to $0.20-0.23$. On the other hand, the magnitudes of local dependencies are only slightly reduced by the introduction of the industry factor. This shows that our results are not driven by exposure to common industry-level shocks but by granular interactions. Using other equal-weighted industry factors does not affect this finding.

	(1) factor component							(2) spatial-temporal component						
	α	β_1	β_2	β_3	β_4	β_5	β_6	β_I	ψ_0	ψ_1	ψ_2	ψ_3	ψ_4	ψ_5
(1) Spatial CAPM+ Industry factor														
MG	0.009	0.231						0.407	0.387	-0.015	-0.024	-0.008	-0.008	0.002
	(0.002)	(0.025)						(0.021)	(0.018)	(0.003)	(0.002)	(0.002)	(0.002)	(0.001)
%sig	7.9%	81.0%						84.5%	87.3%	49.6%	46.0%	27.9%	30.5%	21.7%
N_P	394	394						394	387	387	387	387	387	387
(2) Spatial factor model (Fama-French three factors+ Industry factor)														
MG	0.007	0.197	-0.120	-0.156				0.451	0.415	-0.018	-0.019	-0.009	-0.007	-0.000
	(0.002)	(0.026)	(0.018)	(0.018)				(0.026)	(0.018)	(0.003)	(0.002)	(0.002)	(0.002)	(0.001)
%sig	8.1%	79.2%	72.8%	82.0%				81.7%	86.6%	51.7%	39.8%	27.1%	28.4%	19.6%
N_p	394	394	394	394				394	387	387	387	387	387	387
(3) Spatial factor model (Fama-French five factors+ Industry factor)														
MG	0.005	0.212	-0.108	-0.164	0.106	0.199		0.445	0.422	-0.014	-0.018	-0.009	-0.008	-0.001
	(0.002)	(0.025)	(0.018)	(0.019)	(0.019)	(0.017)		(0.025)	(0.018)	(0.003)	(0.002)	(0.002)	(0.002)	(0.001)
%sig	6.9%	77.7%	71.6%	81.5%	69.0%	70.8%		82.0%	86.6%	46.8%	36.7%	28.2%	28.4%	20.4%
N_p	394	394	394	394	394	394		394	387	387	387	387	387	387
(4) Spatial factor model (Fama-French five factors plus Momentum+ Industry factor)														
MG	0.005	0.221	-0.104	-0.155	0.102	0.194	0.003	0.435	0.420	-0.013	-0.017	-0.009	-0.008	-0.001
	(0.002)	(0.026)	(0.018)	(0.018)	(0.019)	(0.017)	(0.007)	(0.025)	(0.018)	(0.003)	(0.002)	(0.002)	(0.002)	(0.001)
%sig	6.3%	79.4%	71.1%	80.2%	69.0%	67.8%	54.1%	80.7%	85.5%	46.0%	37.2%	27.6%	27.9%	20.4%
N_p	394	394	394	394	394	394	394	394	387	387	387	387	387	387

Table 2: **QML estimation results of industry factors augmented models (Equation 33) using full sample.**

Note: W is constructed using threshold $m = 1$. Estimation results for alternative threshold values are reported in Table 17 and Table 18.

So far, we have shown the mean group (MG) estimation results for whole sample companies. It is also interesting to gauge heterogeneity at sub-group levels. It is reasonable to suspect that the mean sensitivities to local risk spillovers are different for different industry groups. To explore this heterogeneity, here we adopt the random coefficient assumptions at the industry level. Subscript g denotes industry membership, and we classify stocks into six broad industries[5].

$$\begin{aligned} &\beta_{k,i,g} = \beta_{k,g} + \zeta_{k,i,g}, \quad \psi_{l,i,g} = \psi_{l,g} + \varsigma_{l,i,g} \\ &\text{for } k = 1, \dots, K,\ l = 1, \dots, L \text{ ,and } i = 1, \dots, N,\ g = 1, \dots, G. \\ &\boldsymbol{\eta}_{i,g} = (\boldsymbol{\zeta}'_{i,g}, \boldsymbol{\varsigma'}'_{i,g})' \sim IID(\mathbf{0}, \Omega_\eta). \end{aligned} \tag{34}$$

The industry-level common mean parameters for industry g can be consistently estimated when N_g, the number of cross-sectional units within that industry is large.

$$\hat{\beta}^{MG}_{k,g} = \frac{1}{N_g} \sum_{i \in \mathbb{N}_g} \hat{\beta}_{k,i,g} \ \ and \ \ \hat{\psi}^{MG}_{l,g} = \frac{1}{N_g} \sum_{i \in \mathbb{N}_g} \hat{\psi}_{l,i,g}. \tag{35}$$

We report the mean group (MG) estimates by industry for the spatial factor model with Fama-French five factors plus the momentum factor Equation 30 and its counterpart with the industry factor in Table 3 and

[5] We adopt broad industry classification to guarantee that there are a large number of stocks within each industry since mean group estimation requires large N to be consistent. We build the industry classification on top of the Fama-French five industry definitions where they classify all stocks according to their SIC code into five broad groups: "Consumer", "Health", "Hi-tech", "Manufacturing" and "Others". For the first four categories, we keep the same definitions as Fama and French. Since there are a large proportion of financial companies in the $S\&P500$ universe, it would be interesting to separate financial firms from those in the "Others" category. Among the stocks that fall into "Others", we categorize the stocks with a SIC in the range 6000 − 6799 as "Finance" and put the remaining in the "Others" category.

Table 4, respectively. Both tables reveal that our main conclusions that equity returns are affected by that of their economic neighbours are very robust to the industry disaggregation. Local dependencies are highly significant for all six industries. The industrial mean group (MG) estimates of ψ_0 are between $0.36 - 0.58$ for the Spatial factor model with Fama-French five factors and the momentum factor (Equation 30). After controlling for the industry factor, the estimates still range from $0.31 - 0.56$.

Financial companies have the largest exposures to their neighbours' shocks. And this high level of sensitivity to local shocks cannot be explained by exposures to common industry shocks as the estimates of spatial parameters stay unchanged with the introduction of the industry factor. After controlling for the industry factor, the mean group (MG) estimates of ψ_0 for the financial industry is still as large as 0.56 (0.05). By contrast, the introduction of the industry factor reduces the estimated local dependencies for the consumer industry, health industry, and manufacturing industry by a larger margin. Apart from the large contemporaneous spatial coefficient, it is also worth noticing that financial companies also have a stronger momentum spillover effect. The percentage of significant spatial-temporal coefficients[6] for financial companies are much larger than that for other industries at any lag.

	(1) factor component							(2) spatial-temporal component					
	α	β_1	β_2	β_3	β_4	β_5	β_6	ψ_0	ψ_1	ψ_2	ψ_3	ψ_4	ψ_5
Panel A: Finance													
MG	0.017	0.471	0.073	0.410	-0.002	-0.086	-0.101	0.584	-0.040	-0.013	-0.008	-0.007	0.003
	(0.002)	(0.055)	(0.037)	(0.068)	(0.058)	(0.050)	(0.019)	(0.044)	(0.010)	(0.005)	(0.004)	(0.005)	(0.004)
%sig	5.9%	73.5%	67.6%	86.8%	55.9%	54.4%	64.7%	88.2%	75.0%	47.1%	36.8%	47.1%	29.4%
N_p	68	68	68	68	68	68	68	68	68	68	68	68	68
Panel B: Consumer													
MG	0.012	0.525	0.186	-0.230	0.303	0.401	-0.048	0.412	0.019	-0.012	-0.002	0.001	-0.000
	(0.003)	(0.044)	(0.033)	(0.023)	(0.027)	(0.030)	(0.015)	(0.040)	(0.005)	(0.004)	(0.004)	(0.004)	(0.003)
%sig	5.3%	85.3%	80.0%	76.0%	82.7%	89.3%	53.3%	86.7%	48.0%	26.7%	28.0%	28.0%	17.3%
N_p	75	75	75	75	75	75	75	75	75	75	75	75	75
Panel C: Health													
MG	0.026	0.495	0.025	-0.396	-0.149	0.161	0.085	0.364	0.016	-0.003	0.013	0.004	0.005
	(0.007)	(0.066)	(0.044)	(0.038)	(0.053)	(0.046)	(0.022)	(0.058)	(0.007)	(0.004)	(0.004)	(0.005)	(0.005)
%sig	21.4%	92.9%	64.3%	96.4%	64.3%	60.7%	67.9%	89.3%	39.3%	7.1%	17.9%	10.7%	10.7%
N_p	28	28	28	28	28	28	28	28	28	28	28	28	28
Panel D: Hi-tech													
MG	0.016	0.651	0.163	-0.410	-0.292	0.280	-0.038	0.407	0.009	-0.009	0.005	-0.002	0.007
	(0.004)	(0.045)	(0.032)	(0.028)	(0.047)	(0.042)	(0.013)	(0.043)	(0.006)	(0.004)	(0.004)	(0.004)	(0.003)
%sig	5.7%	91.4%	65.7%	92.9%	68.6%	65.7%	45.7%	85.5%	55.1%	27.5%	17.4%	18.8%	14.5%
N_p	70	70	70	70	70	70	70	69	69	69	69	69	69
Panel E: Manufacturing													
MG	0.002	0.538	0.129	-0.169	0.424	0.121	0.049	0.580	0.019	-0.000	-0.002	-0.001	-0.002
	(0.002)	(0.045)	(0.024)	(0.022)	(0.026)	(0.047)	(0.013)	(0.038)	(0.004)	(0.002)	(0.002)	(0.003)	(0.002)
%sig	0.0%	89.4%	80.5%	66.4%	93.8%	81.4%	69.9%	91.6%	49.5%	20.6%	14.0%	27.1%	17.8%
N_p	113	113	113	113	113	113	113	107	107	107	107	107	107
Panel F: Others													
MG	0.007	0.634	0.288	-0.217	0.206	0.258	-0.050	0.418	0.013	-0.007	-0.010	-0.010	-0.000
	(0.005)	(0.070)	(0.043)	(0.052)	(0.046)	(0.053)	(0.021)	(0.067)	(0.006)	(0.004)	(0.004)	(0.005)	(0.004)
%sig	2.5%	90.0%	72.5%	72.5%	60.0%	80.0%	50.0%	85.0%	37.5%	27.5%	30.0%	25.0%	17.5%
N_p	40	40	40	40	40	40	40	40	40	40	40	40	40

Table 3: **QML estimation results of Spatial factor model with Fama-French five factors and the momentum factor (Equation 30). Parameters summarized by industry.**

[6] We need to interpret the mean group estimates of these spatial-temporal parameters with care. The individual parameters $\psi_{l,i,g}$ are quite dispersed for $l \geq 1$, with some firms having significantly positive spatial-temporal terms and some having significantly negative ones. That is why the mean group estimates for these spatial-temporal parameters may not look very statistically significant, although high percentages of individual coefficients are significant — there is simply too much heterogeneity.

	(1) factor component								(2) spatial-temporal component					
	α	β_1	β_2	β_3	β_4	β_5	β_6	β_I	ψ_0	ψ_1	ψ_2	ψ_3	ψ_4	ψ_5
Panel A: Finance														
MG	0.010	0.298	-0.068	0.313	0.036	-0.055	-0.071	0.302	0.560	-0.041	-0.018	-0.013	-0.009	0.001
	(0.003)	(0.063)	(0.058)	(0.054)	(0.055)	(0.048)	(0.018)	(0.062)	(0.046)	(0.010)	(0.005)	(0.005)	(0.005)	(0.004)
%sig	7.4%	89.7%	72.1%	88.2%	58.8%	54.4%	67.6%	66.2%	85.3%	75.0%	50.0%	39.7%	42.6%	33.8%
N_p	68	68	68	68	68	68	68	68	68	68	68	68	68	68
Panel B: Consumer														
MG	0.003	0.084	-0.195	-0.226	0.235	0.318	0.015	0.576	0.358	-0.010	-0.025	-0.014	-0.008	-0.008
	(0.003)	(0.056)	(0.028)	(0.022)	(0.024)	(0.028)	(0.012)	(0.053)	(0.037)	(0.004)	(0.004)	(0.004)	(0.004)	(0.003)
%sig	4.0%	74.7%	70.7%	77.3%	74.7%	76.0%	52.0%	90.7%	81.3%	40.0%	38.7%	30.7%	28.0%	16.0%
N_p	75	75	75	75	75	75	75	75	75	75	75	75	75	75
Panel C: Health														
MG	0.011	0.280	-0.200	-0.262	0.116	0.228	0.071	0.361	0.305	-0.014	-0.022	0.000	-0.005	-0.004
	(0.007)	(0.054)	(0.039)	(0.039)	(0.040)	(0.044)	(0.022)	(0.036)	(0.059)	(0.007)	(0.004)	(0.004)	(0.005)	(0.005)
%sig	10.7%	78.6%	67.9%	96.4%	42.9%	67.9%	64.3%	96.4%	92.9%	28.6%	14.3%	10.7%	7.1%	14.3%
N_p	28	28	28	28	28	28	28	28	28	28	28	28	28	28
Panel D: Hi-tech														
MG	0.017	0.628	0.147	-0.396	-0.301	0.261	-0.035	-0.002	0.410	0.006	-0.009	0.004	-0.004	0.007
	(0.004)	(0.043)	(0.040)	(0.029)	(0.049)	(0.043)	(0.013)	(0.034)	(0.041)	(0.006)	(0.004)	(0.004)	(0.004)	(0.004)
%sig	7.1%	85.7%	64.3%	90.0%	72.9%	62.9%	40.0%	58.6%	84.1%	42.0%	27.5%	15.9%	18.8%	10.1%
N_p	70	70	70	70	70	70	70	70	69	69	69	69	69	69
Panel E: Manufacturing														
MG	-0.002	-0.028	-0.246	-0.188	0.249	0.189	0.053	0.732	0.407	-0.014	-0.016	-0.012	-0.009	-0.002
	(0.002)	(0.048)	(0.029)	(0.020)	(0.020)	(0.030)	(0.011)	(0.047)	(0.036)	(0.004)	(0.003)	(0.002)	(0.003)	(0.002)
%sig	6.2%	75.2%	81.4%	69.0%	76.1%	69.9%	54.9%	94.7%	87.9%	47.7%	40.2%	27.1%	29.0%	24.3%
N_p	113	113	113	113	113	113	113	113	107	107	107	107	107	107
Panel F: Others														
MG	-0.000	0.299	0.031	-0.225	0.248	0.257	-0.019	0.376	0.432	-0.006	-0.017	-0.016	-0.014	-0.003
	(0.005)	(0.086)	(0.044)	(0.050)	(0.047)	(0.053)	(0.021)	(0.054)	(0.062)	(0.006)	(0.004)	(0.004)	(0.005)	(0.004)
%sig	5.0%	72.5%	55.0%	75.0%	67.5%	77.5%	50.0%	75.0%	85.0%	22.5%	37.5%	35.0%	30.0%	17.5%
N_p	40	40	40	40	40	40	40	40	40	40	40	40	40	40

Table 4: **QML estimation results of spatial factor model with Fama-French five factors, the momentum factor, and the industry factor. Parameters summarized by industry.**

Next, we examine how the spatial factor model captures the remaining dependence in the de-factored returns. Using the method described in subsection 2.3, we compute the number of non-zero pair-wise cross correlations of residuals from (1) Factor Model with FF5+MOM+IF that use Fama-French five factors, and momentum factor + 12 Industry factor, and (2) Spatial Factor Model with FF5+MOM+IF that use Fama-French five factors, and momentum factor + 12 Industry factor. [7]

To test $H_{0,ij} : \rho_{ij} = 0$ for $n_{test} = N(N-1)/2$ pairs of (i, j), for a given family-wise error rate (FWER) p, the critical values are $\pm\frac{1}{\sqrt{T}}\Phi^{-1}(1-p/2n_{test})$ for the factor model, and $F^{-1}(p/2n_{test})$ and $F^{-1}(1-p/2n_{test})$ for the spatial factor model. F is the empirical null distribution from $B = 500$ bootstrap samples. Figure 1 shows the histogram of bootstrapped $\hat{\rho}^b_{ij}$ for all $i \neq j, b = 1, \ldots, 500$ for the spatial factor model. Table 5 presents the degree of cross-sectional dependence in the factor model and its spatial-augmented version under different family-wise error rates. First of all, bootstrap results show that the limiting distribution of $\hat{\rho}_{ij}$ under the null is indeed altered by the addition of the spatial term, and we need different critical values for testing. The table shows that adding the spatial component reduces the number of non-zero pair-wise cross correlations by a huge margin.[8] The spatial component constructed with news-implied linkages is successful at eliminating correlations from the de-factored returns.

[7] Here we only present the results for the models with most factors that are supposed to have least residual correlations among all. We can do more different factor models and their spatial augmented versions, at the cost of bootstrap inference for each spatial factor model.

[8] This is not only because we need larger critical values under the spatial factor specification. Even if we do not consider the distortion brought by the spatial component and still use limiting distributions from factor model residual correlation coefficients, the percentage of non-zero pair-wise cross correlations is still reduced by half.

	Critical values	# Non-zero pair-wise cross correlations	Density
(1) $p = 0.05$			
Factor Model with FF5+MOM+IF	-0.091,0.091	8516	5.50%
Spatial Factor Model with FF5+MOM+IF	-0.243,0.245	416	0.27%
(2) $p = 0.1$			
Factor Model with FF5+MOM+IF	-0.088,0.088	9490	5.87%
Spatial Factor Model with FF5+MOM+IF	-0.225,0.229	478	0.31%

Table 5: Degree of Cross-Sectional Dependence in the Residuals.

Note: Density gives the percentage of non-zero pair-wise cross correlations (i.e., density=number of non-zero pair-wise cross-correlations/N(N-1)).

6.2 Model Performance: Degree of Mispricing

Next, we compare the performance of factor models with their spatial versions by several measures of the degree of mispricing. We are interested in whether adding the spatial component to the factor model could reduce mispricing.

If an asset pricing model completely captures expected returns, then the intercept $\tilde{\boldsymbol{\alpha}}$ is approximately zero. We are dealing with a high-dimensional system. Testing $H_0 : \tilde{\boldsymbol{\alpha}} = 0$ for large N using traditional tests like Gibbons et al. (1989) which are designed for cases where the number of test assets are small will have low power problem. Also, this test for whether exact arbitrage pricing holds is stronger than what is implied by the approximate no arbitrage condition (Equation 20 and Equation 21).

So instead of using standard GRS test, we employ other three statistics to compare the relative performance of different models (1) the percentage of individually significant $\tilde{\alpha}_i$ (Pesaran and Yamagata (2012)); (2) average L_1 norm of intercepts $A(|\ \tilde{\alpha}_i\ |)$; (3) average L_2 norm of intercepts $A(\tilde{\alpha}_i^2)$ (Fama and French (2015)). Those three statistics are implied by theorem 1 and the corollaries to be useful in measuring how well the approximation (Equation 20) is. In addition to those three measures of mispricing, we also report the mean-squared errors (MSE) of different models.

We compare the relative performances of the following factor models, their spatial augmented versions.

- Model 1.1: CAPM model (CAPM)
- Model 1.2: CAPM model + Industry factor (CAPM+IF)
- Model 1.3: Spatial CAPM model (CAPM(S))
- Model 1.4: Spatial CAPM model + Industry factor (CPAM+IF(S))
- Model 2.1: Factor model with Fama-French three factors (FF3)
- Model 2.2: Factor model with Fama-French three factors + Industry factor (FF3+IF)
- Model 2.3: Spatial factor model with Fama-French three factors (FF3(S))
- Model 2.4: Spatial factor model with Fama-French three factors + Industry factor (FF3+IF(S))
- Model 3.1: Factor model with Fama-French five factors (FF5)

- Model 3.2: Factor model with Fama-French five factors + Industry factor (FF5+IF)
- Model 3.3: Spatial factor model with Fama-French five factors (FF5(S))
- Model 3.4: Spatial factor model with Fama-French five factors + Industry factor (FF5+IF(S))
- Model 4.1: Factor model with Fama-French five factors, and momentum factor (FF5+MOM)
- Model 4.2: Factor model with Fama-French five factors, and momentum factor + Industry factor (FF5+MOM+IF)
- Model 4.3: Spatial factor model with Fama-French five factors, and momentum factor (FF5+MOM(S))
- Model 4.4: Spatial factor model with Fama-French five factors, and momentum factor + Industry factor (FF5+MOM+IF(S))

	% of significant $\tilde{\alpha}_i$	$A(\mid \tilde{\alpha}_i \mid)$	$A(\tilde{\alpha}_i^2)$	Mean-squared error (MSE)
Model 1.1: CAPM	7.61%	2.51	10.91	3.26
Model 1.2: CAPM+IF	7.61%	2.47	11.19	2.99
Model 1.3: CAPM(S)	9.13%	2.43	10.41	2.86
Model 1.4: CAPM+IF(S)	7.87%	2.36	10.41	2.73
Model 2.1: FF3	7.64%	2.49	10.66	3.26
Model 2.2: FF3+IF	7.35%	2.35	10.69	2.86
Model 2.3: FF3(S)	6.09%	2.31	9.39	2.75
Model 2.4: FF3+IF(S)	8.12%	2.28	9.90	2.65
Model 3.1: FF5	5.84%	2.45	11.17	2.99
Model 3.2: FF5+IF	6.60%	2.39	11.42	2.82
Model 3.3: FF5(S)	5.58%	2.21	9.39	2.72
Model 3.4: FF5+IF(S)	6.68%	2.24	9.90	2.62
Model 4.1: FF5+MOM	6.09%	2.49	11.18	2.97
Model 4.2: FF5+MOM+IF	6.35%	2.39	11.34	2.80
Model 4.3: FF5+MOM(S)	4.80%	2.18	9.14	2.70
Model 4.4: FF5+MOM+IF(S)	6.45%	2.19	9.64	2.61

Table 6: Summary of Model Performance.

Note: Each panel shows the performance statistics of a factor model, its spatial augmented version, and its spatial and industry factor augmented version. Note: $\tilde{\boldsymbol{\alpha}}$ used to compute $A(\mid \tilde{\alpha}_i \mid)$ and $A(\tilde{\alpha}_i^2)$ are in basis point, and $\boldsymbol{\epsilon}$ used to compute MSE errors are all in percentage point. For each column, the best statistic is highlighted in red.

Table 6 shows that for all factor models except the CAPM, adding spatial interactions improves all performance measures. Spatial CAPM fails to reduce the percentage of individually significant intercepts. This is because spatial models are designed for modelling local interactions. If there are not enough common risk factors to capture the strong dependence in equity returns, adding the spatial component which deals with weak dependence is not going to be helpful. For factor models with Fama-French three-factor, five-factor, and five-factor plus momentum factor, adding spatial interactions all provides noticeable improvement on reducing the mispricing component and mean squared errors. Interestingly, although adding the industry factor can further reduce mean-squared error, it cannot help to reduce the mispricing [9].

For each measure, the best-performing statistic is highlighted in red. Model 4.3, the spatial factor model with Fama-French five factors, and momentum factor appears to have the best performance in terms of reducing pricing errors. And model 4.4, which is model 4.3 with the industry factor, has the smallest mean-squared errors.

6.3 Comparisons with alternative networks

In this section, we gauge whether the news-implied links carry additional information on top of existing linkage datasets. We first show that spatial factor models estimated with W constructed using other existing linkage datasets under-perform that estimated with news-implied W. We then show that conditional on other existing linkages, local risk spillovers via our news-implied links continue to be significant.

We consider the following competing networks:

- Industry-based adjacency matrices based on industry classification of different granularities including 4-digit SIC codes, 3-digit SIC codes, and 2-digit SIC codes classifications. This is motivated by Moskowitz and Grinblatt (1999), Engelberg et al. (2018) and Fan et al. (2016). For each classification criteria, we build block-diagonal matrices where companies within the same industry are fully connected.
- IBES analyst co-coverage networks. It has been documented that shared analyst coverage is a strong proxy for fundamental linkages between firms and reflects firm similarities along many dimensions (Ali and Hirshleifer (2020), Israelsen (2016), Kaustia and Rantala (2013)). We use the Institutional Brokers Estimate System (IBES) detail history files to construct the analyst co-coverage-based adjacency matrix. For each year in the sample, we consider a stock is covered by an analyst if the analyst issues at least one FY1 or FY2 earnings forecast for the stock during the year. And we consider two stocks as linked if there are common analysts during the year, weighted by the number of common analysts. We then add up the yearly adjacency matrices to get the full sample adjacency matrix.
- Customer-supplier links (Cohen and Frazzini (2008)) from Andrea Frazzini's data library. The strength of links is weighted by sales.
- Geographic links (Pirinsky and Wang (2006) and Parsons et al. (2020)). We obtain location information from CRSP Compustat merged files. We then merge the sample firms with the Metropolitan Statistical Areas (MSA) data using the ZIP-FIPS-MSA data from the US Department of Labor, which maps zip codes to MSAs. We follow Pirinsky and Wang (2006), and consider firms whose headquarters are in the same MSA as linked.
- The union of above mentioned links. We let a typical entry w_{ij} in this matrix to be one if the pair (i, j) is linked in any of the above networks, and zero otherwise.

[9] In unreported tables, if we replace equal-weighted industry portfolios with value-weighted industry portfolios, the introduction of industry factor does further bring down three statistics of mispricing. And model 4.4, the spatial factor model with Fama-French five factors, and momentum factor plus Industry factor has the best performance in all dimensions. However, as we have argued earlier, value-weighted portfolios might cause endogeneity issues given we are working with large companies. Using other equal weighting industry factors does not change the results.

Table 7 shows the performance of competing networks. Individually, spatial APT models estimated with new-implied linkages out-perform other existing networks. And even if we consider the union of all alternative networks, the news-implied network still does not lose the battle. For the four statistics that we consider, spatial factor model with Fama-French five factors and momentum factor estimated using W_{news} (last line of panel (1)) performs the best among all candidates along two dimensions. Spatial factor model with Fama-French five factors, momentum factor, and industry factor estimated using W_{union} (second last line of panel (2)) also performs the best among all candidates along two dimensions. Models with adjacency matrices capturing multiple channels out-perform those with adjacency matrices focusing on one particular channel. This seems to support the fact that there are multiple channels of local risk spillovers. There is no reason to focus on one particular channel like intra-industry channel, customer-supplier channel, etc. Our news-implied linkages provide a comprehensive and integrated measure of firm-level relatedness, and it can be seen as a nice proxy of firm-to-firm connectivity.

	% of significant $\tilde{\alpha}_i$	$A(\mid \tilde{\alpha}_i \mid)$	$A(\tilde{\alpha}_i^2)$	Mean-squared error (MSE)
Panel (1): Spatial factor model with Fama-French five factors, and momentum factor				
$W_{2-digit-SIC}$	7.61%	2.22	9.64	2.68
$W_{3-digit-SIC}$	6.85%	2.37	10.91	2.82
$W_{4-digit-SIC}$	6.35%	2.40	11.68	2.88
W_{IBES}	6.85%	2.29	9.90	2.73
$W_{Customer-Supplier}$	6.09%	2.48	11.17	2.96
$W_{Geographic}$	7.11%	2.67	12.44	2.97
W_{Union}	6.09%	2.17	9.14	2.64
W_{News}	4.80%	2.18	9.14	2.70
Panel (2): Spatial factor model with Fama-French five factors, and momentum factor + Industry factor				
$W_{2-digit-SIC}$	6.09%	2.23	10.14	2.65
$W_{3-digit-SIC}$	5.58%	2.31	10.91	2.77
$W_{4-digit-SIC}$	5.58%	2.35	11.93	2.83
W_{IBES}	6.09%	2.22	9.90	2.68
$W_{Customer-Supplier}$	6.85%	2.39	11.17	2.81
$W_{Geographic}$	6.60%	2.45	11.93	2.81
W_{Union}	5.83%	2.16	9.39	2.61
W_{News}	6.45%	2.19	9.64	2.61

Table 7: Summary of Model Performance using competing networks.
Note: Panel (1) shows the performance of competing adjacency matrices under the spatial factor model with Fama-French five factors, and momentum factor. Panel (2) shows the performance of competing adjacency matrices under the spatial factor model with Fama-French five factors, and momentum factor plus industry factor. $\tilde{\boldsymbol{\alpha}}$ used to compute $A(\mid \tilde{\alpha}_i \mid)$ and $A(\tilde{\alpha}_i^2)$ are in basis point, and $\boldsymbol{\epsilon}$ used to compute MSE errors are all in percentage point. For each column, the best statistic is highlighted in red.

Next, we examine that whether our news-implied linkages carry new information on top of existing linkages documented? To do that, we estimate the two-W spatial factor models below, with W_1 being our news-implied networks and W_2 being a set of other candidate matrices

$$\tilde{\boldsymbol{r}}_t = \tilde{\boldsymbol{\alpha}} + \mathbf{B}\tilde{\mathbf{f}}_\mathbf{t} + \sum_{l=0}^{L} \Psi_{1,l} W_1 \tilde{\boldsymbol{r}}_{t-l} + \sum_{l=0}^{L} \Psi_{2,l} W_2 \tilde{\boldsymbol{r}}_{t-l} + \boldsymbol{\epsilon}_t. \tag{36}$$

$$\tilde{\boldsymbol{r}}_t = \tilde{\boldsymbol{\alpha}} + \mathbf{B}\tilde{\mathbf{f}}_\mathbf{t} + \boldsymbol{\beta}_I f_{IND,t} + \sum_{l=0}^{L} \Psi_{1,l} W_1 \tilde{\boldsymbol{r}}_{t-l} + \sum_{l=0}^{L} \Psi_{2,l} W_2 \tilde{\boldsymbol{r}}_{t-l} + \boldsymbol{\epsilon}_t. \tag{37}$$

	(1) W_1						**(2)** W_2					
	$\psi_{1,0}$	$\psi_{1,1}$	$\psi_{1,2}$	$\psi_{1,3}$	$\psi_{1,4}$	$\psi_{1,5}$	$\psi_{2,0}$	$\psi_{2,1}$	$\psi_{2,2}$	$\psi_{2,3}$	$\psi_{2,4}$	$\psi_{2,5}$
Panel(1): $W_1 = W_{news}$ **and** $W_2 = W_{2-digit-SIC}$												
MG	0.299	0.005	-0.002	0.001	0.001	0.000	0.309	-0.000	-0.005	-0.004	-0.002	0.001
	(0.017)	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)	(0.020)	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)
%sig	75.2%	32.3%	19.6%	21.2%	20.2%	19.6%	82.8%	27.1%	16.4%	18.3%	18.8%	17.8%
N_p	387	387	387	387	387	387	377	377	377	377	377	377
Panel(2): $W_1 = W_{news}$ **and** $W_2 = W_{3-digit-SIC}$												
MG	0.296	0.007	-0.003	-0.000	-0.002	-0.001	0.271	0.000	-0.003	-0.003	0.001	0.002
	(0.019)	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)	(0.026)	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)
%sig	75.5%	35.9%	19.9%	20.9%	25.3%	17.8%	83.4%	26.8%	19.0%	17.5%	17.8%	15.1%
N_p	387	387	387	387	387	387	332	332	332	332	332	332
Panel(3): $W_1 = W_{news}$ **and** $W_2 = W_{4-digit-SIC}$												
MG	0.318	0.010	-0.003	-0.000	-0.003	0.000	0.264	-0.002	-0.004	-0.002	0.002	0.002
	(0.019)	(0.004)	(0.003)	(0.002)	(0.003)	(0.003)	(0.029)	(0.005)	(0.003)	(0.003)	(0.003)	(0.003)
%sig	75.2%	39.3%	19.9%	20.7%	24.3%	19.6%	81.6%	26.7%	18.8%	17.7%	17.7%	16.3%
N_p	387	387	387	387	387	387	288	288	288	288	288	288
Panel(4): $W_1 = W_{news}$ **and** $W_2 = W_{IBES}$												
MG	0.365	0.004	-0.005	-0.001	-0.003	-0.001	0.197	0.001	-0.001	-0.002	0.000	0.003
	(0.019)	(0.004)	(0.002)	(0.003)	(0.003)	(0.002)	(0.014)	(0.003)	(0.002)	(0.002)	(0.003)	(0.002)
%sig	83.5%	39.8%	20.9%	19.1%	22.5%	19.4%	80.0%	26.2%	16.8%	17.6%	15.6%	17.6%
N_p	387	387	387	387	387	387	340	340	340	340	340	340
Panel(5): $W_1 = W_{news}$ **and** $W_2 = W_{Customer-Supplier}$												
MG	0.483	0.006	-0.007	-0.001	-0.003	0.001	0.082	0.005	-0.004	-0.003	0.008	0.003
	(0.019)	(0.003)	(0.002)	(0.002)	(0.002)	(0.001)	(0.021)	(0.004)	(0.004)	(0.004)	(0.005)	(0.004)
%sig	87.1%	50.6%	26.1%	23.3%	27.1%	17.6%	69.5%	13.6%	11.9%	11.9%	10.2%	13.6%
N_p	387	387	387	387	387	387	59	59	59	59	59	59
Panel(6): $W_1 = W_{news}$ **and** $W_2 = W_{Geographic}$												
MG	0.459	0.006	-0.006	0.002	-0.000	0.001	0.093	-0.002	-0.001	-0.005	-0.005	0.001
	(0.018)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)	(0.019)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
%sig	86.3%	36.7%	21.7%	19.9%	21.2%	15.8%	59.9%	22.5%	18.2%	12.1%	14.0%	13.4%
N_p	387	387	387	387	387	387	307	307	307	307	307	307
Panel(7): $W_1 = W_{news}$ **and** $W_2 = W_{Union}$												
MG	0.305	0.007	-0.001	0.003	0.003	-0.000	0.374	-0.004	-0.005	-0.005	-0.005	0.001
	(0.016)	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)	(0.020)	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)
%sig	75.5%	26.6%	17.8%	20.2%	18.9%	16.3%	83.8%	23.6%	17.3%	18.3%	17.0%	17.5%
N_p	387	387	387	387	387	387	394	394	394	394	394	394

Table 8: **QML estimation results of two-W spatial factor model with Fama-French five factors, and the Momentum factor (Equation 36).**

Note: we only report spatial parameters here. W_{news} is constructed using threshold $m = 1$.

	(1) W_1						(2) W_2					
	$\psi_{1,0}$	$\psi_{1,1}$	$\psi_{1,2}$	$\psi_{1,3}$	$\psi_{1,4}$	$\psi_{1,5}$	$\psi_{2,0}$	$\psi_{2,1}$	$\psi_{2,2}$	$\psi_{2,3}$	$\psi_{2,4}$	$\psi_{2,5}$
Panel(1): $W_1 = W_{news}$ **and** $W_2 = W_{2-digit-SIC}$												
MG	0.280	-0.002	-0.007	-0.003	-0.003	-0.001	0.240	-0.011	-0.009	-0.007	-0.004	-0.001
	(0.017)	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)	(0.019)	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)
%sig	74.7%	26.9%	21.2%	22.7%	21.2%	21.2%	82.0%	28.9%	15.4%	18.6%	17.8%	18.3%
N_p	387	387	387	387	387	387	377	377	377	377	377	377
Panel(2): $W_1 = W_{news}$ **and** $W_2 = W_{3-digit-SIC}$												
MG	0.267	-0.004	-0.011	-0.006	-0.006	-0.003	0.228	-0.007	-0.005	-0.004	-0.001	0.001
	(0.018)	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)	(0.025)	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)
%sig	72.4%	31.3%	24.0%	24.8%	22.7%	18.1%	80.4%	28.3%	19.3%	18.1%	16.0%	16.0%
N_p	387	387	387	387	387	387	332	332	332	332	332	332
Panel(3): $W_1 = W_{news}$ **and** $W_2 = W_{4-digit-SIC}$												
MG	0.282	-0.003	-0.010	-0.007	-0.007	-0.002	0.222	-0.010	-0.006	-0.003	0.000	0.001
	(0.018)	(0.004)	(0.003)	(0.002)	(0.003)	(0.003)	(0.028)	(0.005)	(0.003)	(0.003)	(0.003)	(0.003)
%sig	71.3%	31.8%	25.3%	24.3%	22.0%	19.4%	78.8%	29.5%	20.1%	18.1%	18.1%	17.0%
N_p	387	387	387	387	387	387	288	288	288	288	288	288
Panel(4): $W_1 = W_{news}$ **and** $W_2 = W_{IBES}$												
MG	0.318	-0.007	-0.013	-0.007	-0.006	-0.002	0.160	-0.006	-0.004	-0.003	-0.002	0.002
	(0.018)	(0.004)	(0.002)	(0.003)	(0.003)	(0.002)	(0.015)	(0.003)	(0.002)	(0.002)	(0.003)	(0.002)
%sig	79.6%	36.2%	23.8%	20.9%	22.2%	19.4%	79.7%	27.6%	17.1%	17.6%	15.9%	19.1%
N_p	387	387	387	387	387	387	340	340	340	340	340	340
Panel(5): $W_1 = W_{news}$ **and** $W_2 = W_{Customer-Supplier}$												
MG	0.418	-0.014	-0.017	-0.009	-0.008	-0.001	0.053	0.002	-0.002	-0.004	0.005	0.002
	(0.018)	(0.003)	(0.002)	(0.002)	(0.002)	(0.001)	(0.014)	(0.004)	(0.004)	(0.004)	(0.005)	(0.004)
%sig	84.2%	45.0%	37.2%	27.1%	27.4%	18.9%	71.2%	15.3%	13.6%	11.9%	6.8%	15.3%
N_p	387	387	387	387	387	387	59	59	59	59	59	59
Panel(6): $W_1 = W_{news}$ **and** $W_2 = W_{Geographic}$												
MG	0.399	-0.007	-0.014	-0.003	-0.004	-0.001	0.074	-0.009	-0.005	-0.008	-0.008	-0.000
	(0.018)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)	(0.016)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
%sig	83.7%	33.6%	24.8%	23.5%	21.4%	16.5%	59.3%	25.7%	19.2%	13.0%	12.7%	14.3%
N_p	387	387	387	387	387	387	307	307	307	307	307	307
Panel(7): $W_1 = W_{news}$ **and** $W_2 = W_{Union}$												
MG	0.292	0.002	-0.006	-0.000	0.001	-0.001	0.305	-0.016	-0.010	-0.010	-0.009	-0.001
	(0.016)	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)	(0.018)	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)
%sig	76.0%	25.6%	18.1%	18.9%	19.4%	17.1%	83.0%	25.4%	17.5%	18.8%	17.0%	18.0%
N_p	387	387	387	387	387	387	394	394	394	394	394	394

Table 9: **QML estimation results of two-W spatial factor model with Fama-French five factors, the momentum factor, and the industry factor (Equation 37).**

Note: we only report spatial parameters here. W_{news} is constructed using threshold $m = 1$.

Table 8 and Table 9 shows the estimation results for Equation 36 and Equation 37, respectively. Although the magnitude of local dependencies among news-implied peers is weakened by the introduction of other networks, our new-implied links are still important channels of risk spillovers. The mean group (MG) estimates of $\psi_{1,0}$ are around $0.27 - 0.48$, with more than 70% of parameters being individually significant across different specifications.

Even if we condition on the union of all alternative linkages (i.e., $W_2 = W_{union}$), the magnitude of local dependencies among news-implied peers is still quite large. For the specification without industry factor, average $\psi_{2,0}$ is larger than average $\psi_{1,0}$. However, the introduction of industry factor reduces $\hat{\psi}_{2,0}^{MG}$ while leaving $\hat{\psi}_{1,0}^{MG}$ unchanged, making the local spillover effect via news-implied network equally strong as the effect via the union of all other alternative networks. The results confirm that the novel dataset carries additional information on top of existing networks. The statistically and economically significant local dependencies among the news-implied peers cannot be explained by other existing networks.

7 Conclusion

This paper studies a heterogeneous coefficient spatial factor model, which addresses both strong cross-sectional dependence and a very flexible form of weak cross-sectional dependence in equity returns. Theoretically, it extends classical asset pricing models like CAPM and APT, which only consider the strong form of cross-sectional dependence. We characterize how local dependence affects asset returns under the assumption of no asymptotic arbitrage. Empirically, we focus on the weak/local dependency in equity returns, which is an area less explored in empirical financial studies due to data availability issues. Utilizing the novel business news-implied linkage data, we construct the channels through which the local shocks transmit. We adopt a flexible heterogeneous coefficient spatial-temporal model, and we find that stocks linked via business news co-mentioning exhibit excess co-movement beyond that is predicted by standard asset pricing models like CAPM and APT. Exposures to common risk factors and local interactions are two distinct mechanism that jointly explain the co-movement in asset returns. It is important for investors and policy makers to separately analyse the two types of dependencies to fully understand what type of risk are they exposed to.

One interesting question for future work is whether the spatial factor model can be applied for portfolio construction problem. With the presence of both factor-driven strong dependence and the remaining weak dependence, literature on high-dimensional equity returns covariance matrix usually consider the following estimator

$$\hat{\Sigma}_y = \hat{\mathbf{B}} c\hat{o}v(\boldsymbol{f})_t \hat{\mathbf{B}}' + \hat{\Sigma}_\epsilon, \tag{38}$$

where $\hat{\Sigma}_\epsilon$ is a regularised sparse error covariance matrix, and the estimation of Σ_y is achieved in two steps. The spatial factor model we study in this paper implies the following covariance structure

(39)

which can be estimated in a single step.

Another interesting future work is the formal testing of no arbitrage $H_0 : \boldsymbol{\alpha} = \mathbf{0}$ for the spatial factor model when N is large. Pesaran and Yamagata (2012), Pesaran and Yamagata (2017) consider testing for alpha in factor models with large N, how to extend the theory is non-trivial and needs thorough analysis.

References

D. Acemoglu, V. M. Carvalho, A. Ozdaglar, and A. Tahbaz-Salehi. The network origins of aggregate fluctuations. *Econometrica*, 80(5):1977–2016, 2012.

U. Ali and D. Hirshleifer. Shared analyst coverage: Unifying momentum spillover effects. *Journal of Financial Economics*, 136(3):649–675, 2020.

M. Aquaro, N. Bailey, and M. H. Pesaran. Estimation and inference for spatial models with heterogeneous coefficients: an application to us house prices. *Journal of Applied Econometrics*, 2020.

N. Bailey, S. Holly, and M. H. Pesaran. A two-stage approach to spatio-temporal analysis with strong and weak cross-sectional dependence. *Journal of Applied Econometrics*, 31(1):249–280, 2016.

N. Bailey, G. Kapetanios, and M. H. Pesaran. Exponent of cross-sectional dependence for residuals. *Sankhya B*, 81(1):46–102, 2019a.

N. Bailey, M. H. Pesaran, and L. V. Smith. A multiple testing approach to the regularisation of large sample correlation matrices. *Journal of Econometrics*, 208(2):507–534, 2019b.

N. Bailey, G. Kapetanios, and M. H. Pesaran. Measurement of factor strenght: Theory and practice. 2020.

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

M. Barigozzi and C. Brownlees. Nets: Network estimation for time series. *Journal of Applied Econometrics*, 34 (3):347–364, 2019.

M. Barigozzi and M. Hallin. A network analysis of the volatility of high dimensional financial series. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 66(3):581–605, 2017.

C. E. Bonferroni. Il calcolo delle assicurazioni su gruppi di teste. *Studi in onore del professore salvatore ortu carboni*, pages 13–60, 1935.

G. Chamberlain and M. Rothschild. Arbitrage, factor structure, and mean-variance analysis on large asset markets. *Econometrica: Journal of the Econometric Society*, pages 1281–1304, 1983.

J. Chen, Y. Shin, and C. Zheng. Estimation and inference in heterogeneous spatial panels with a multifactor error structure. *Journal of Econometrics*, 2021.

A. Chudik and M. H. Pesaran. Mean group estimation in presence of weakly cross-correlated estimators. *Economics Letters*, 175:101–105, 2019.

A. Chudik, M. H. Pesaran, and E. Tosetti. Weak and strong cross-section dependence and estimation of large panels, 2011.

L. Cohen and A. Frazzini. Economic links and predictable returns. *The Journal of Finance*, 63(4):1977–2011, 2008.

E. Dimson. Risk measurement when shares are subject to infrequent trading. *Journal of Financial Economics*, 7(2):197–226, 1979.

J. Engelberg, A. Ozoguz, and S. Wang. Know thy neighbor: Industry clusters, information spillovers, and market efficiency. *Journal of Financial and Quantitative Analysis*, 53(5):1937–1961, 2018.

F. Eugene. The cross-section of expected stock returns. *Journal of Finance*, 47(2):427–465, 1992.

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

J. Fan, Y. Fan, and J. Lv. High dimensional covariance matrix estimation using a factor model. *Journal of Econometrics*, 147(1):186–197, 2008.

J. Fan, Y. Liao, and M. Mincheva. High dimensional covariance matrix estimation in approximate factor models. *Annals of statistics*, 39(6):3320, 2011.

J. Fan, A. Furger, and D. Xiu. Incorporating global industrial classification standard into portfolio allocation: A simple factor-based large covariance matrix estimator with high-frequency data. *Journal of Business & Economic Statistics*, 34(4):489–503, 2016.

X. Gabaix. The granular origins of aggregate fluctuations. *Econometrica*, 79(3):733–772, 2011.

D. Garcia. Sentiment during recessions. *The Journal of Finance*, 68(3):1267–1300, 2013.

M. R. Gibbons, S. A. Ross, and J. Shanken. A test of the efficiency of a given portfolio. *Econometrica: Journal of the Econometric Society*, pages 1121–1152, 1989.

G. Hale and J. A. Lopez. Monitoring banking system connectedness with big data. *Journal of Econometrics*, 212(1):203–220, 2019.

G. Hoberg and G. Phillips. Text-based network industries and endogenous product differentiation. *Journal of Political Economy*, 124(5):1423–1465, 2016.

J. E. Ingersoll Jr. Some results in the theory of arbitrage pricing. *The Journal of Finance*, 39(4):1021–1039, 1984.

R. D. Israelsen. Does common analyst coverage explain excess comovement? *Journal of Financial and Quantitative Analysis*, 51(4):1193–1229, 2016.

M. Kaustia and V. Rantala. Common analyst-based method for defining peer firms. *Available at SSRN*, 2013.

Z. T. Ke, B. T. Kelly, and D. Xiu. Predicting returns with text data. Technical report, National Bureau of Economic Research, 2019.

H. H. Kelejian and I. R. Prucha. A generalized spatial two-stage least squares procedure for estimating a spatial autoregressive model with autoregressive disturbances. *The Journal of Real Estate Finance and Economics*, 17(1):99–121, 1998.

H. H. Kelejian and I. R. Prucha. A generalized moments estimator for the autoregressive parameter in a spatial model. *International economic review*, 40(2):509–533, 1999.

S. Kou, X. Peng, and H. Zhong. Asset pricing with spatial interaction. *Management Science*, 64(5):2083–2101, 2018.

G. M. Kuersteiner and I. R. Prucha. Dynamic spatial panel models: Networks, common shocks, and sequential exogeneity. *Econometrica*, 88(5):2109–2146, 2020.

L.-F. Lee. Asymptotic distributions of quasi-maximum likelihood estimators for spatial autoregressive models. *Econometrica*, 72(6):1899–1925, 2004.

L.-f. Lee. Gmm and 2sls estimation of mixed regressive, spatial autoregressive models. *Journal of Econometrics*, 137(2):489–514, 2007.

L.-f. Lee and J. Yu. Estimation of spatial autoregressive panel data models with fixed effects. *Journal of econometrics*, 154(2):165–185, 2010.

J. P. LeSage. An introduction to spatial econometrics. *Revue d'économie industrielle*, (123):19–44, 2008.

J. P. LeSage and Y.-Y. Chih. A bayesian spatial panel model with heterogeneous coefficients. *Regional Science and Urban Economics*, 72:58–73, 2018.

T. J. Moskowitz and M. Grinblatt. Do industries explain momentum? *The Journal of finance*, 54(4):1249–1290, 1999.

C. A. Parsons, R. Sabbatucci, and S. Titman. Geographic lead-lag effects. *The Review of Financial Studies*, 33 (10):4721–4770, 2020.

H. Pesaran et al. General diagnostic tests for cross-sectional dependence in panels. *University of Cambridge, Cambridge Working Papers in Economics*, 435, 2004.

M. H. Pesaran and R. Smith. Estimating long-run relationships from dynamic heterogeneous panels. *Journal of econometrics*, 68(1):79–113, 1995.

M. H. Pesaran and T. Yamagata. Testing capm with a large number of assets. In *AFA 2013 San Diego Meetings Paper*, 2012.

M. H. Pesaran and T. Yamagata. Testing for alpha in linear factor pricing models with a large number of securities. 2017.

M. H. Pesaran and C. F. Yang. Estimation and inference in spatial models with dominant units. *Journal of Econometrics*, 221(2):591–615, 2021.

C. Pirinsky and Q. Wang. Does corporate headquarters location matter for stock returns? *The Journal of Finance*, 61(4):1991–2015, 2006.

S. Ross. The arbitrage theory of capital asset pricing. *Journal of Economic Theory*, 13(3):341–360, 1976.

A. Scherbina and B. Schlusche. Economic linkages inferred from news stories and the predictability of stock returns. *Available at SSRN 2363436*, 2015.

G. Schwenkler and H. Zheng. The network of firms implied by the news. *Available at SSRN 3320859*, 2019.

W. F. Sharpe. Capital asset prices: A theory of market equilibrium under conditions of risk. *The journal of finance*, 19(3):425–442, 1964.

W. Shi and L.-f. Lee. Spatial dynamic panel data models with interactive fixed effects. *Journal of Econometrics*, 197(2):323–347, 2017.

Appendices

A Proofs of Theorems and Corollaries

A.1 Proof of Theorem 1

This proof is heavily borrowed from Kou et al. (2018) and Ingersoll Jr (1984). Under Assumption 3, $(I-\Psi W)$ is invertible and we denote the inverse as $G(\psi)=(I-\Psi W)^{-1}$. We rewrite the spatial factor model (Equation 7) as

$$\boldsymbol{r} = G\boldsymbol{\alpha} + G\mathbf{B}\mathbf{f} + G\boldsymbol{\epsilon}. \tag{40}$$

We let

$$\dot{\boldsymbol{\alpha}} = G\boldsymbol{\alpha}, \ \ \dot{\mathbf{B}} = G\mathbf{B}, \ \ \dot{\boldsymbol{\epsilon}} = G\boldsymbol{\epsilon}. \tag{41}$$

The spatial factor model can be written as a reduced-form factor model

$$\boldsymbol{r} = \dot{\boldsymbol{\alpha}} + \dot{\mathbf{B}}\mathbf{f} + \dot{\boldsymbol{\epsilon}}. \tag{42}$$

In particular, the covariance matrix of the reduced form error is

$$\dot{\Omega} = E(\dot{\boldsymbol{\epsilon}}\dot{\boldsymbol{\epsilon}}') = G\Omega G'. \tag{43}$$

We follow Ingersoll Jr (1984), and factor the positive definite covariance matrix $\dot{\Omega}$ as $\dot{\Omega} = CC'$, where C is a nonsingular matrix. Now consider a subsequences of assets. For the nth economy, consider the orthogonal projection of the vector $(C^n)^{-1}\dot{\boldsymbol{\alpha}}^n$ into the space spanned by $(C^n)^{-1}\mathbf{1}^n$ and the columns of $(C^n)^{-1}\dot{\mathbf{B}}^n$ as follows:

$$(C^n)^{-1}\dot{\boldsymbol{\alpha}}^n = (C^n)^{-1}\mathbf{1}^n\lambda_0^n + (C^n)^{-1}\dot{\mathbf{B}}^n\boldsymbol{\lambda}^n + \boldsymbol{u}^n. \tag{44}$$

By the nature of orthogonal projection,

$$0 = (\dot{\mathbf{B}}^n)'((C^n)')^{-1}\boldsymbol{u}^n = (\mathbf{1}^n)'((C^n)')^{-1}\boldsymbol{u}^n. \tag{45}$$

Given the pricing error $\boldsymbol{v}^n$ we defined in Equation 16, the reduced form pricing error from the reduced form factor model Equation 42 is $\dot{\boldsymbol{v}}^n = G\boldsymbol{v}^n$

$$\dot{\boldsymbol{v}}^n = \dot{\boldsymbol{\alpha}}^n - \mathbf{1}^n\lambda_0^n - \dot{\mathbf{B}}^n\boldsymbol{\lambda}^n = G^n(\boldsymbol{\alpha}^n - (G^n)^{-1}\mathbf{1}^n\lambda_0^n - \mathbf{B}^n\boldsymbol{\lambda}^n) = G^n\boldsymbol{v}^n. \tag{46}$$

The reduced form pricing error $\dot{\boldsymbol{v}}^n$ can be written as $\dot{\boldsymbol{v}}^n = C^n\boldsymbol{u}^n$ directly from Equation 55. Using the orthogonal conditions Equation 56 and the factorization $\dot{\Omega}^n = C^n(C^n)'$, we have:

$$(\dot{\mathbf{B}}^n)'(\dot{\Omega}^n)^{-1}\dot{\boldsymbol{v}}^n = (\mathbf{1}^n)'(\dot{\Omega}^n)^{-1}\dot{\boldsymbol{v}}^n = 0. \tag{47}$$

Consider a zero cost portfolio $\boldsymbol{c}^n = (\dot{\Omega}^n)^{-1}\dot{\boldsymbol{v}}^n[(\dot{\boldsymbol{v}}^n)'(\dot{\Omega}^n)^{-1}\dot{\boldsymbol{v}}^n]^{-1}$

$$(\mathbf{1}^n)'\boldsymbol{c}^n = (\mathbf{1}^n)'(\dot{\Omega}^n)^{-1}\dot{\boldsymbol{v}}^n[(\dot{\boldsymbol{v}}^n)'(\dot{\Omega}^n)^{-1}\dot{\boldsymbol{v}}^n]^{-1} = 0, \tag{48}$$

with expected return

$$E((\boldsymbol{c}^n)'\boldsymbol{r}^n) = (\boldsymbol{c}^n)'\dot{\boldsymbol{\alpha}}^n = [(\dot{\boldsymbol{v}}^n)'(\dot{\Omega}^n)^{-1}\dot{\boldsymbol{v}}^n]^{-1}(\dot{\boldsymbol{v}}^n)'(\dot{\Omega}^n)^{-1}(\mathbf{1}^n\lambda_0^n + \dot{\mathbf{B}}^n\boldsymbol{\lambda}^n + \dot{\boldsymbol{v}}^n) = 1, \tag{49}$$

and variance

$$
\begin{aligned}
Var((\boldsymbol{c}^n)'\boldsymbol{r}^n) &= (\boldsymbol{c}^n)'Var(\boldsymbol{r}^n)\boldsymbol{c}^n \\
&= [(\dot{\boldsymbol{v}}^n)'(\dot{\Omega}^n)^{-1}\dot{\boldsymbol{v}}^n]^{-1}(\dot{\boldsymbol{v}}^n)'(\dot{\Omega}^n)^{-1}(\dot{\mathbf{B}}^n(\dot{\mathbf{B}}^n)' + \dot{\Omega}^n)(\dot{\Omega}^n)^{-1}\dot{\boldsymbol{v}}^n[(\dot{\boldsymbol{v}}^n)'(\dot{\Omega}^n)^{-1}\dot{\boldsymbol{v}}^n]^{-1} \\
&= [(\dot{\boldsymbol{v}}^n)'(\dot{\Omega}^n)^{-1}\dot{\boldsymbol{v}}^n]^{-1} = [(G^n\boldsymbol{v}^n)'(G^n\Omega^n(G^n)')^{-1}G^n\boldsymbol{v}^n]^{-1} \\
&= [(\boldsymbol{v}^n)'(\Omega^n)^{-1}(\boldsymbol{v}^n)]^{-1}
\end{aligned}
\tag{50}
$$

If the weighted sum of squared pricing errors $(\boldsymbol{v}^n)'(\Omega^n)^{-1}(\boldsymbol{v}^n)$ is not uniformly bounded (i.e., Equation 17 is violated), then the variance of this portfolio would go to zero along some subsequence, and the asymptotic arbitrage opportunity described in Equation 14 exists.

A.2 Proof Corollary 1.1

This is a direct result of Theorem 3 from Chamberlain and Rothschild (1983).

A.3 Proof Corollary 1.2

Again, we look at the reduced form factor model

$$
\boldsymbol{r} = \dot{\boldsymbol{\alpha}} + \dot{\mathbf{B}}\mathbf{f} + \dot{\boldsymbol{\epsilon}}, \tag{51}
$$

where

$$
\dot{\boldsymbol{\alpha}} = G\boldsymbol{\alpha}, \quad \dot{\mathbf{B}} = G\mathbf{B}, \quad \dot{\boldsymbol{\epsilon}} = G\boldsymbol{\epsilon}. \tag{52}
$$

$$
\dot{\Omega} = E(\dot{\boldsymbol{\epsilon}}\dot{\boldsymbol{\epsilon}}') = G\Omega G'. \tag{53}
$$

The risk factors are given by

$$
f_k = \tilde{f}_k - E(\tilde{f}_k) \text{ for } k = 1, \ldots, K. \tag{54}
$$

We factor the positive definite covariance matrix $\dot{\Omega}$ as $\dot{\Omega} = CC'$, where C is a nonsingular matrix. Now consider a subsequences of assets. For the nth economy, consider the orthogonal projection of the vector $(C^n)^{-1}(\dot{\boldsymbol{\alpha}}^n - r_f\mathbf{1}^n)$ onto the space spanned by columns of $(C^n)^{-1}\dot{\mathbf{B}}^n$ as follows:

$$
(C^n)^{-1}(\dot{\boldsymbol{\alpha}}^n - r_f\mathbf{1}^n) = (C^n)^{-1}\dot{\mathbf{B}}^n\boldsymbol{\lambda}^n + \boldsymbol{u}^n. \tag{55}
$$

By the nature of orthogonal projection,

$$
(\dot{\mathbf{B}}^n)'((C^n)')^{-1}\boldsymbol{u}^n = 0. \tag{56}
$$

We define

$$
\boldsymbol{v}^n = \boldsymbol{\alpha}^n - r_f(I_n - \Psi^n W^n)\mathbf{1}^n - \mathbf{B}^n E(\tilde{\boldsymbol{f}}) = \boldsymbol{\alpha}^n - r_f(G^n)^{-1}\mathbf{1}^n - \mathbf{B}^n E(\tilde{\boldsymbol{f}}). \tag{57}
$$

Define the the reduced form pricing error

$$
\dot{\boldsymbol{v}}^n = \dot{\boldsymbol{\alpha}}^n - r_f\mathbf{1}^n - \dot{\mathbf{B}}^n\boldsymbol{\lambda}^n. \tag{58}
$$

Under the assumption that factors are traded, $\boldsymbol{\lambda}^n = E(\tilde{\boldsymbol{f}})$. And we have $\dot{\boldsymbol{v}}^n = G^n\boldsymbol{v}^n$ as:

$$
\dot{\boldsymbol{v}}^n = \dot{\boldsymbol{\alpha}}^n - r_f\mathbf{1}^n - \dot{\mathbf{B}}^n\boldsymbol{\lambda}^n = G^n(\boldsymbol{\alpha}^n - r_f(G^n)^{-1}\mathbf{1}^n - \mathbf{B}^n E(\tilde{\boldsymbol{f}})) = G^n\boldsymbol{v}^n. \tag{59}
$$

Given there exists risk free asset, consider a zero-cost portfolio which take a long position $\boldsymbol{c}^n$ in the risky assets and short position $(\boldsymbol{c}^n)'\mathbf{1}^n$ in the risk free asset, where $\boldsymbol{c}^n = (\dot{\Omega}^n)^{-1}\dot{\boldsymbol{v}}^n[(\dot{\boldsymbol{v}}^n)'(\dot{\Omega}^n)^{-1}\dot{\boldsymbol{v}}^n]^{-1}$. This portfolio generates expected return

$$\begin{aligned}
&E((\boldsymbol{c}^n)'(\boldsymbol{r}^n - r_f\mathbf{1}^n)) \\
&= E((\boldsymbol{c}^n)'(\dot{\boldsymbol{\alpha}}^n - r_f\mathbf{1}^n)) + E((\boldsymbol{c}^n)'\dot{\mathbf{B}}^n\boldsymbol{f}) + E((\boldsymbol{c}^n)'\dot{\boldsymbol{\epsilon}}^n) \\
&= E((\boldsymbol{c}^n)'(\dot{\boldsymbol{\alpha}}^n - r_f\mathbf{1}^n)) = E((\boldsymbol{c}^n)'(\dot{\mathbf{B}}^n\boldsymbol{\lambda}^n + \dot{\boldsymbol{v}}^n)) \\
&= [(\dot{\boldsymbol{v}}^n)'(\dot{\Omega}^n)^{-1}\dot{\boldsymbol{v}}^n]^{-1}(\dot{\boldsymbol{v}}^n)'(\dot{\Omega}^n)^{-1}\dot{\boldsymbol{v}}^n = 1,
\end{aligned} \tag{60}$$

and variance

$$\begin{aligned}
&Var((\boldsymbol{c}^n)'(\boldsymbol{r}^n - r_f\mathbf{1}^n)) \\
&= (\boldsymbol{c}^n)'Var(\dot{\mathbf{B}}^n\boldsymbol{f} + \dot{\boldsymbol{\epsilon}}^n)\boldsymbol{c}^n \\
&= [(\dot{\boldsymbol{v}}^n)'(\dot{\Omega}^n)^{-1}\dot{\boldsymbol{v}}^n]^{-1}(\dot{\boldsymbol{v}}^n)'(\dot{\Omega}^n)^{-1}(\dot{\mathbf{B}}^n(\dot{\mathbf{B}}^n)' + \dot{\Omega}^n)(\dot{\Omega}^n)^{-1}\dot{\boldsymbol{v}}^n[(\dot{\boldsymbol{v}}^n)'(\dot{\Omega}^n)^{-1}\dot{\boldsymbol{v}}^n]^{-1} \\
&= [(\dot{\boldsymbol{v}}^n)'(\dot{\Omega}^n)^{-1}\dot{\boldsymbol{v}}^n]^{-1} = [(G^n\boldsymbol{v}^n)'(G^n\Omega^n(G^n)')^{-1}G^n\boldsymbol{v}^n]^{-1} \\
&= [(\boldsymbol{v}^n)'(\Omega^n)^{-1}(\boldsymbol{v}^n)]^{-1}
\end{aligned} \tag{61}$$

For pricing errors $\boldsymbol{v}^n = \boldsymbol{\alpha}^n - r_f(I_n - \Psi^n W^n)\mathbf{1}^n - \mathbf{B}^n E(\tilde{\boldsymbol{f}})$ defined in Equation 57, if the weighted sum of squared pricing errors $(\boldsymbol{v}^n)'(\Omega^n)^{-1}(\boldsymbol{v}^n)$ is not uniformly bounded (i.e., Equation 17 is violated), then the variance of this portfolio would go to zero along some subsequence, and the asymptotic arbitrage opportunity described in Equation 14 exists.

When there exists risk free rate r_f, and if we write the risk factor f_k before de-meaning as $\tilde{f}_k$, then the spatial factor model (Equation 7) can be written as:

$$\begin{aligned}
&\tilde{\boldsymbol{r}} = \tilde{\boldsymbol{\alpha}} + \mathbf{B}\tilde{\mathbf{f}} + \Psi W\tilde{\boldsymbol{r}} + \boldsymbol{\epsilon}, \\
&\text{where } \ \tilde{\boldsymbol{r}} = \boldsymbol{r} - r_f\mathbf{1} \ \text{ is the vector of excess returns,} \\
&\tilde{\mathbf{f}} = (\tilde{f}_1, \dots, \tilde{f}_K)', \\
&\tilde{\boldsymbol{\alpha}} = \boldsymbol{\alpha} - (I - \Psi W)\mathbf{1}r_f - \mathbf{B}E(\tilde{\mathbf{f}}).
\end{aligned} \tag{62}$$

Comparing this pricing errors $\boldsymbol{v}^n = \boldsymbol{\alpha}^n - r_f(I_n - \Psi^n W^n)\mathbf{1}^n - \mathbf{B}^n E(\tilde{\boldsymbol{f}})$ with Equation 62, we can tell $\boldsymbol{v}^n = \tilde{\boldsymbol{\alpha}}^n$, and the asymptotic no arbitrage condition is equivalent to

$$\tilde{\boldsymbol{\alpha}}^n \approx 0. \tag{63}$$

Comparing $\boldsymbol{v}^n = \tilde{\boldsymbol{\alpha}}^n = \boldsymbol{\alpha}^n - r_f(I_n - \Psi^n W^n)\mathbf{1}^n - \mathbf{B}^n E(\tilde{\boldsymbol{f}})$ with Equation 16, we can tell

$$\begin{aligned}
&\lambda_0^n = r_f, \\
&\boldsymbol{\lambda}^n = E(\tilde{\mathbf{f}}).
\end{aligned} \tag{64}$$

A.4 Proof Corollary 1.3

We first rewrite the spatial-factor model with the dependent variable being the excesss returns (Equation 62) as

$$(I - \Psi W)\tilde{\boldsymbol{r}} = G^{-1}\tilde{\boldsymbol{r}} = \tilde{\boldsymbol{\alpha}} + \mathbf{B}\tilde{\mathbf{f}} + \boldsymbol{\epsilon}. \tag{65}$$

Suppose the asset returns from an infinite economy are generated by:

$$(G^n)^{-1}\tilde{\boldsymbol{r}}^n = \tilde{\boldsymbol{\alpha}}^n + \mathbf{B}^n\tilde{\mathbf{f}} + \boldsymbol{\epsilon}^n \tag{66}$$

For any fixed $\delta > 0$, assume $I(\mid \tilde{\alpha}_j^n \mid > \delta) = 1$ for $j = 1, \dots, N(n, \delta)$. For each of those $N(n, \delta)$ elements, we can construct a zero cost portfolio as following way.

Take the jth element for example. Denote the jth column of Identify matrix I_n by $\boldsymbol{e}_j$. If $\tilde{\alpha}_j^n > \delta$, consider a zero-cost portfolio which takes a long position $\boldsymbol{e}_j'(G^n)^{-1}$ in excess returns $\tilde{\boldsymbol{r}}^n$, and short position $\boldsymbol{e}_j'\mathbf{B}^n$ in the zero-cost traded factors $\tilde{\mathbf{f}}$. If $\tilde{\alpha}_j^n < -\delta$, consider a zero-cost portfolio which takes a short position $\boldsymbol{e}_j'(G^n)^{-1}$ in excess returns $\tilde{\boldsymbol{r}}^n$, and long position $\boldsymbol{e}_j'\mathbf{B}^n$ in the zero-cost traded factors $\tilde{\mathbf{f}}$. The portfolio is a zero-cost one because the long and short position are both zero-cost. The portfolio has expected return $\mid \tilde{\alpha}_j^n \mid > \delta$, and variance $\sigma_j^2 < \bar{\sigma}^2$.

We can construct $N(n, \delta)$ such portfolios. Consider a new portfolio that takes equal weight in these $N(n, \delta)$ portfolios. This new portfolio is zero-cost, with expected return $\frac{\sum_{j=1}^{N(n,\delta)} |\tilde{\alpha}_j^n|}{N(n,\delta)} > \delta > 0$. For uncorrelated errors, the variance of this portfolio is smaller than $\frac{\bar{\sigma}^2}{N(n,\delta)}$. If Equation 23 fails, and $N(n, \delta)$ is diverging, then we have asymptotic arbitrage.

B Identification and Inference of the Heterogeneous Spatial-Temporal Model

Aquaro et al. (2020) studies the conditions under which $\boldsymbol{\theta}_0$ is identified, and establishes consistency and asymptotic normality of the estimator. Write the $(N*(K+L+3))$ by 1 vector $\boldsymbol{\theta} = (\tilde{\boldsymbol{\alpha}}', \boldsymbol{\beta}_1', \dots, \boldsymbol{\beta}_K', \boldsymbol{\psi}_0', \dots, \boldsymbol{\psi}_L', \boldsymbol{\sigma}_{\epsilon^2}')' = (\boldsymbol{b}', \boldsymbol{\psi}_0', \boldsymbol{\sigma}_{\epsilon^2}')'$, where $\boldsymbol{b} = (\tilde{\boldsymbol{\alpha}}', \boldsymbol{\beta}_1', \dots, \boldsymbol{\beta}_K', \boldsymbol{\psi}_1' \dots, \boldsymbol{\psi}_L')'$ is a $(N*(K+L+1))$ by 1 vector that contains all the parameters associated with weakly exogenous variables $\boldsymbol{x}_t$. The following assumptions are made:

Assumption 5 *The parameter vector* $\boldsymbol{\theta} = (\boldsymbol{b}', \boldsymbol{\psi}_0', \boldsymbol{\sigma}_{\epsilon^2}')'$ *belongs to* $\Theta = \Theta_b \times \Theta_{\boldsymbol{\psi}_0} \times \Theta_\sigma \subset R^{N*(K+L+1)} \times R^N \times R^N$*, a subset of the* $(N*(K+L+3))$ *dimensional Euclidean space* $\mathbb{R}^{N*(K+L+3)}$. Θ *is a closed and bounded (compact) set, and* $\boldsymbol{\theta}_0$ *is an interior point of* Θ.

Assumption 6 *The error terms* $\{\epsilon_{it}, i = 1, \dots, N; t = 1, \dots, T\}$ *are independently distributed over* i *and* t*. For filtration* $\mathcal{F}_t = (\boldsymbol{x}_t, \boldsymbol{x}_{t-1}, \boldsymbol{x}_{t-2}, ...)$, $E(\epsilon_{it} \mid \mathcal{F}_t) = 0$, $E(\epsilon_{it}^2 \mid \mathcal{F}_t) = \sigma_{i0}^2$*, for* $i = 1, \dots, N$*, so there is no conditional heteroskedasticity.* $inf_i\sigma_{i0}^2 > c > 0$ *and* $sup_i\sigma_{i0}^2 < \bar{\sigma}^2 < \infty$*, and* $E(\mid \epsilon_{it} \mid^p \mid \mathcal{F}_t) = E((\mid \epsilon_{it} \mid^p) = \bar{\omega}_{ip} < \bar{c}$*, for all* i *and* t*, where* $1 \leq p \leq 4 + \varepsilon$*, for some* $\varepsilon > 0$.

Assumption 7 *(a)* $\boldsymbol{x}_t$ *are stationary processes, that satisfy the moment condition* $sup_{i,t,l}E(\mid x_{it,l} \mid^{2+g}) < \bar{c}$*, for some* $g > 0$*,* $i = 1, \dots, N$*,* $t = 1, \dots, T$*,* $l = 1, \dots, (K + L + 1)$.
(b) $E(\boldsymbol{x}_t\boldsymbol{x}_t') = \Sigma_{xx}$*, where entry* $\Sigma_{ij} = E(\boldsymbol{x}_{it}\boldsymbol{x}_{jt}')$ *exists for all* i *and* j*, such as* $sup_{i,j}\|\Sigma_{ij}\| < \bar{c}$*, and* Σ_{ii} *is a* $k \times k$ *non-singular matrix with* $inf_i[\lambda_{min}(\Sigma_{ii})] > c > 0$*, and* $sup_i[\lambda_{min}(\Sigma_{ii})] < \bar{c} < \infty$.
(c) $\frac{1}{T}\sum_{t=1}^T \boldsymbol{x}_t\boldsymbol{x}_t' \xrightarrow{a.s} \Sigma_{xx}$ *as* $T \to \infty$.

Assumption 8 *(a) The adjacency matrix* W *is known, with zeros on the diagonal.*

(b) The adjacency matrix W has bounded row sum norm, and $\|W\|_\infty < c < \infty$, and

$$\sup_{\psi_i \in \Theta_{\psi}} | \psi_i | < \frac{1}{\|W\|_\infty}. \tag{67}$$

Definition 1 *The set $\mathcal{N}_c(\boldsymbol{\sigma}_0^2)$ in the closed neighbourhood of $\boldsymbol{\sigma}_0^2$ if:*

$$\mathcal{N}_c(\boldsymbol{\sigma}_0^2) = \{\boldsymbol{\sigma}_0^2 \in \Theta_\sigma, | \sigma_{i0}^2/\sigma_i^2 - 1 | < c_i, \text{for } i = 1, \ldots, N\}, \tag{68}$$

for some $c_i > 0$, where Θ_σ is a compact subset of $\mathcal{R}^N$.

Assumption 9 *The $(N * (K + L + 3))$ by 1 vector $\boldsymbol{\theta} = (\boldsymbol{b}', \boldsymbol{\psi}_0', \boldsymbol{\sigma}_{\epsilon^2}')'$ belongs to $\Theta_c = \Theta_b \times \Theta_{\boldsymbol{\psi}_0} \times \mathcal{N}_c(\boldsymbol{\sigma}_0^2)$. Θ_b and $\Theta_{\boldsymbol{\psi}_0}$ are compact subsets of $\mathcal{R}^{N*(K+L+1)}$ and $\mathcal{R}^N$, respectively, and $\mathcal{N}_c(\boldsymbol{\sigma}_0^2)$ is defined in definition 1, and Θ_c is a subset of the $(N * (K + L + 3))$ dimensional Euclidean space, $\mathcal{R}^{N*(K+L+3)}$.*

The identification results are given by the following proposition:

Proposition 2 *Suppose that Assumptions 1-5 hold, consider a heterogeneous coefficient spatial-temporal model given by Equation 25 and log-likelihood function given by Equation 26. For fixed N, K and L, the $(N*(K+L+3))$ dimensional true parameter vector $\boldsymbol{\theta}_0$ is almost surely locally identified on Θ_c.*

The main inference results are given by the following proposition:

Proposition 3 *Suppose that Assumptions 1-5 hold, consider a heterogeneous coefficient spatial-temporal model given by Equation 25. For fixed N, K and L, the $(N*(K+L+3))$ dimensional QML estimator of $\boldsymbol{\theta}_0$ is denoted as $\hat{\boldsymbol{\theta}}_{QMLE}$, which is almost surely locally consistent for $\boldsymbol{\theta}_0$ on Θ_c, and has the following asymptotic distribution:*

$$\sqrt{T}(\hat{\boldsymbol{\theta}}_{QMLE} - \boldsymbol{\theta}_0) \xrightarrow{d} N(\boldsymbol{0}, \boldsymbol{V}_\theta), \tag{69}$$

where $\boldsymbol{V}_\theta$ is the asymptotic covariance matrix, which has a standard sandwich form:

$$\boldsymbol{V}_\theta = \boldsymbol{H}^{-1}(\boldsymbol{\theta}_0)\boldsymbol{J}(\boldsymbol{\theta}_0, \gamma)\boldsymbol{H}^{-1}(\boldsymbol{\theta}_0), \tag{70}$$

where $\boldsymbol{H}(\boldsymbol{\theta}_0) = \lim_{T\to\infty} E_0(-\frac{1}{T}\frac{\partial^2 \ell_T(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}'})$ is the Hessian, and $\boldsymbol{J}(\boldsymbol{\theta}_0, \gamma)$ is the asymptotic variance of the score, which depends on the distribution of the errors. In the case of Gaussian errors, $\gamma = 2$, and $\boldsymbol{H}(\boldsymbol{\theta}_0) = \boldsymbol{J}(\boldsymbol{\theta}_0, 2)$.

Remark: theorem 2 and theorem 3 describe the identification results and asymptotic distribution for each individual parameter in the $(N * (K + L + 3))$ by 1 vector. When $T \to \infty$, estimation and inference can be conducted for any N.

C Supplementary Figures and Tables

Number of unique news stories	88, 316, 898
Number of stories remaining after removing topics including analyst recommendations, ratings changes, and index movements	87, 841, 641
Of these:	
Number of stories tag sample companies	8, 341, 848
Of these:	
Number of stories that mention only one company	5, 507, 772 (66.03%)
Number of stories that mention exactly two companies	1, 637, 256 (19.63%)
Number of stories that mention more than two companies	1, 196, 820 (14.34%)

Table 10: Descriptive statistics for RavenPack Equity files Dow Jones Edition for the period January 2004 to December 2015.

Number of yearly window a pair gets identified	Frequency	Percentage	Cumulative Percentage
0	217024	72.80%	72.80%
1	40178	13.48%	86.28%
2	13302	4.46%	90.74%
3	7116	2.39%	93.13%
4	4522	1.52%	94.65%
5	3236	1.09%	95.74%
6	2506	. 0.84%	96.58%
7	2022	0.68%	97.26%
8	1804	0.61%	97.87%
9	1508	0.51%	98.38%
10	1350	0.45%	98.83%
11	1232	0.41%	99.24%
12	2316	0.78%	100%

Table 11: Frequency distribution table of the number of yearly link identification windows that a pair gets identified as economic neighbours for all possible pairs (i, j) in our sample. Note: A pair identified in k yearly windows could get multiple co-mentions within each window.

	Finance	Durbs	Energy	Hi-tec	Health	Manuf	Nondur	Other	Shops	Tel	Utilities
Finance	1840	81	256	777	315	529	273	573	568	116	235
	0.33	0.01	0.05	0.14	0.06	0.10	0.05	0.10	0.10	0.02	0.04
Durbs	81	12	14	67	16	72	27	49	45	10	13
	0.20	0.03	0.03	0.17	0.04	0.18	0.07	0.12	0.11	0.02	0.03
Energy	256	14	372	147	42	115	51	153	83	20	172
	0.18	0.01	0.26	0.10	0.03	0.08	0.04	0.11	0.06	0.01	0.12
Hi-tec	777	67	147	1376	227	419	182	439	403	126	86
	0.18	0.02	0.03	0.32	0.05	0.10	0.04	0.10	0.09	0.03	0.02
Health	315	16	42	227	370	111	71	134	143	28	19
	0.21	0.01	0.03	0.15	0.25	0.08	0.05	0.09	0.10	0.02	0.01
Manuf	529	72	115	419	111	470	134	287	211	43	62
	0.220	0.03	0.05	0.17	0.05	0.19	0.05	0.12	0.09	0.02	0.03
Nondur	273	27	51	182	71	134	196	152	244	42	25
	0.20	0.02	0.04	0.13	0.05	0.10	0.14	0.11	0.17	0.03	0.02
Other	573	49	153	439	134	287	152	344	295	63	138
	0.22	0.02	0.06	0.17	0.05	0.11	0.06	0.13	0.11	0.02	0.05
Shops	568	45	83	403	143	211	244	295	698	73	40
	0.20	0.02	0.03	0.14	0.05	0.08	0.09	0.11	0.25	0.03	0.01
Telcm	116	10	20	126	28	43	42	63	73	18	22
	0.21	0.02	0.04	0.22	0.05	0.08	0.07	0.11	0.13	0.03	0.04
Utilities	235	13	172	86	19	62	25	138	40	22	366
	0.20	0.01	0.15	0.07	0.02	0.05	0.02	0.12	0.03	0.02	0.31

Table 12: **Links aggregated at industry level**. Note: The adjacency matrix is construct using threshold $m = 1$. we use Fama-French 12 industry classification. For each panel, the first row gives the number of intra or inter industry pairs indentified, and the second gives the proportion to total number of links firms in that industry have.

	Finance	Durbs	Energy	Hi-tec	Health	Manuf	Nondur	Other	Shops	Tel	Utilities
Finance	1496	65	193	566	233	377	193	451	397	84	173
	0.35	0.02	0.05	0.13	0.06	0.09	0.05	0.11	0.09	0.02	0.04
Durbs	65	12	8	41	8	43	14	36	29	8	7
	0.24	0.04	0.03	0.15	0.03	0.16	0.05	0.13	0.11	0.03	0.03
Energy	193	8	294	87	18	61	29	95	44	11	110
	0.20	0.01	0.31	0.09	0.02	0.06	0.03	0.10	0.05	0.01	0.12
Hi-tec	566	41	87	1040	123	254	103	311	254	85	40
	0.19	0.01	0.03	0.36	0.04	0.09	0.04	0.11	0.09	0.03	0.01
Health	233	8	18	123	288	64	40	86	75	17	4
	0.24	0.01	0.02	0.13	0.30	0.07	0.04	0.09	0.08	0.02	0
Manuf	377	43	61	254	64	264	84	205	109	26	20
	0.25	0.03	0.04	0.17	0.04	0.18	0.06	0.14	0.07	0.02	0.01
Nondur	193	14	29	103	40	84	144	97	158	30	12
	0.21	0.02	0.03	0.11	0.04	0.09	0.16	0.11	0.17	0.03	0.01
Other	451	36	95	311	86	205	97	256	177	52	84
	0.24	0.02	0.05	0.17	0.05	0.11	0.05	0.14	0.10	0.03	0.05
Shops	397	29	44	254	75	109	158	177	536	48	19
	0.22	0.02	0.02	0.14	0.04	0.06	0.09	0.10	0.29	0.03	0.01
Telcm	84	8	11	85	17	26	30	52	48	18	13
	0.21	0.02	0.03	0.22	0.04	0.07	0.08	0.13	0.12	0.05	0.03
Utilities	173	7	110	40	4	20	12	84	19	13	290
	0.22	0.01	0.14	0.05	0.01	0.03	0.02	0.11	0.02	0.02	0.38

Table 13: **Links aggregated at industry level**. Note: The adjacency matrix is construct using threshold $m = 2$. we use Fama-French 12 industry classification. For each panel, the first row gives the number of intra or inter industry pairs indentified, and the second gives the proportion to total number of links firms in that industry have.

	Finance	Durbs	Energy	Hi-tec	Health	Manuf	Nondur	Other	Shops	Tel	Utilities
Finance	1250	50	153	415	187	289	160	380	315	61	136
	0.37	0.01	0.05	0.12	0.06	0.09	0.05	0.11	0.09	0.02	0.04
Durbs	50	8	7	29	5	31	10	30	20	5	4
	0.25	0.04	0.04	0.15	0.03	0.16	0.05	0.15	0.10	0.03	0.02
Energy	153	7	246	54	11	42	19	72	22	8	60
	0.22	0.01	0.35	0.08	0.02	0.06	0.03	0.10	0.03	0.01	0.09
Hi-tec	415	29	54	832	82	172	63	235	164	73	23
	0.19	0.01	0.03	0.39	0.04	0.08	0.03	0.11	0.08	0.03	0.01
Health	187	5	11	82	246	44	26	67	44	9	2
	0.26	0.01	0.02	0.11	0.34	0.06	0.04	0.09	0.06	0.01	0
Manuf	289	31	42	172	44	186	55	156	62	16	10
	0.27	0.03	0.04	0.16	0.04	0.17	0.05	0.15	0.06	0.02	0.01
Nondur	160	10	19	63	26	55	112	75	114	21	5
	0.24	0.02	0.03	0.10	0.04	0.08	0.17	0.11	0.17	0.03	0.01
Other	380	30	72	235	67	156	75	210	126	30	63
	0.26	0.02	0.05	0.16	0.05	0.11	0.05	0.15	0.09	0.02	0.04
Shops	315	20	22	164	44	62	114	126	394	32	10
	0.24	0.02	0.02	0.13	0.03	0.05	0.09	0.10	0.30	0.02	0.01
Telcm	61	5	8	73	9	16	21	30	32	16	8
	0.22	0.02	0.03	0.26	0.03	0.06	0.08	0.11	0.11	0.06	0.03
Utilities	136	4	60	23	2	10	5	63	10	8	214
	0.25	0.01	0.11	0.04	0	0.02	0.01	0.12	0.02	0.01	0.40

Table 14: **Links aggregated at industry level**. Note: The adjacency matrix is construct using threshold $m = 3$. we use Fama-French 12 industry classification. For each panel, the first row gives the number of intra or inter industry pairs indentified, and the second gives the proportion to total number of links firms in that industry have.

	(1) factor component							(2) spatial-temporal component					
	α	β_1	β_2	β_3	β_4	β_5	β_6	ψ_0	ψ_1	ψ_2	ψ_3	ψ_4	ψ_5
(1) Spatial CAPM													
MG	0.015	0.598						0.416	0.002	-0.008	0.001	-0.003	0.004
	(0.001)	(0.021)						(0.019)	(0.003)	(0.002)	(0.001)	(0.002)	(0.001)
%sig	8.4%	91.6%						88.9%	51.2%	27.6%	20.9%	29.5%	21.2%
N_p	394	394						387	387	387	387	387	387
(2) Spatial factor model (Fama-French three factors)													
MG	0.014	0.570	0.129	-0.127				0.450	0.002	-0.008	-0.001	-0.001	0.002
	(0.001)	(0.021)	(0.014)	(0.022)				(0.019)	(0.003)	(0.002)	(0.001)	(0.002)	(0.001)
%sig	7.1%	90.1%	75.1%	81.5%				89.1%	52.7%	27.4%	20.4%	28.9%	19.1%
N_p	394	394	394	394				387	387	387	387	387	387
(3) Spatial factor model (Fama-French five factors)													
MG	0.011	0.586	0.144	-0.127	0.142	0.177		0.453	0.007	-0.007	-0.001	-0.002	0.002
	(0.001)	(0.021)	(0.014)	(0.023)	(0.022)	(0.021)		(0.018)	(0.003)	(0.002)	(0.001)	(0.002)	(0.001)
%sig	5.8%	90.6%	74.4%	82.2%	75.1%	73.9%		88.9%	51.4%	27.4%	20.7%	27.4%	18.3%
N_p	394	394	394	394	394	394		387	387	387	387	387	387
(4) Spatial factor model (Fama-French five factors plus Momentum)													
MG	0.012	0.593	0.146	-0.137	0.139	0.184	-0.022	0.444	0.006	-0.007	-0.001	-0.002	0.001
	(0.001)	(0.021)	(0.014)	(0.021)	(0.022)	(0.021)	(0.007)	(0.019)	(0.003)	(0.002)	(0.001)	(0.002)	(0.001)
%sig	4.8%	90.1%	74.1%	79.9%	75.4%	74.4%	59.1%	87.6%	51.7%	27.6%	22.2%	27.9%	18.3%
N_p	394	394	394	394	394	394	394	387	387	387	387	387	387

Table 15: **QML estimation results of heterogeneous spatial-temporal model using full sample, with W constructed using threshold $m = 2$**

	(1) factor component							(2) spatial-temporal component					
	α	β_1	β_2	β_3	β_4	β_5	β_6	ψ_0	ψ_1	ψ_2	ψ_3	ψ_4	ψ_5
(1) Spatial CAPM													
MG	0.015	0.628						0.392	0.002	-0.008	0.001	-0.003	0.004
	(0.001)	(0.021)						(0.019)	(0.003)	(0.002)	(0.001)	(0.002)	(0.001)
%sig	8.9%	92.9%						87.8%	49.7%	27.9%	20.3%	29.2%	22.7%
N_p	394	394						384	384	384	384	384	384
(2) Spatial factor model (Fama-French three factors)													
MG	0.014	0.606	0.128	-0.117				0.420	0.002	-0.008	-0.000	-0.000	0.002
	(0.001)	(0.021)	(0.014)	(0.022)				(0.018)	(0.003)	(0.002)	(0.001)	(0.002)	(0.001)
%sig	7.6%	92.1%	75.1%	82.0%				87.8%	51.8%	27.3%	20.6%	29.7%	20.6%
N_p	394	394	394	394				384	384	384	384	384	384
(3) Spatial factor model (Fama-French five factors)													
MG	0.012	0.621	0.144	-0.116	0.141	0.173		0.423	0.006	-0.006	-0.001	-0.002	0.001
	(0.001)	(0.020)	(0.014)	(0.023)	(0.022)	(0.021)		(0.018)	(0.003)	(0.002)	(0.001)	(0.002)	(0.001)
%sig	6.3%	92.6%	73.3%	81.7%	75.1%	72.3%		87.5%	50.8%	28.4%	20.6%	27.9%	19.8%
N_p	394	394	394	394	394	394		384	384	384	384	384	384
(4) Spatial factor model (Fama-French five factors plus Momentum)													
MG	0.012	0.629	0.145	-0.130	0.138	0.182	-0.027	0.414	0.005	-0.007	-0.001	-0.002	0.001
	(0.001)	(0.021)	(0.014)	(0.021)	(0.022)	(0.021)	(0.007)	(0.018)	(0.003)	(0.002)	(0.001)	(0.002)	(0.001)
%sig	5.8%	92.1%	73.6%	78.7%	75.1%	73.3%	61.4%	87.0%	51.0%	28.4%	22.1%	28.4%	20.1%
N_p	394	394	394	394	394	394	394	384	384	384	384	384	384

Table 16: **QML estimation results of heterogeneous spatial-temporal model using full sample, with W constructed using threshold $m = 3$**

	(1) factor component							(2) spatial-temporal component						
	α	β_1	β_2	β_3	β_4	β_5	β_6	β_I	ψ_0	ψ_1	ψ_2	ψ_3	ψ_4	ψ_5
(1) Spatial CAPM+ Industry factor														
MG	0.009	0.257						0.410	0.361	-0.014	-0.023	-0.008	-0.007	0.002
	(0.002)	(0.024)						(0.021)	(0.018)	(0.003)	(0.002)	(0.002)	(0.002)	(0.001)
%sig	7.9%	82.5%						84.5%	86.6%	49.4%	44.4%	27.9%	29.5%	22.0%
N_p	394	394						394	387	387	387	387	387	387
(2) Spatial factor model (Fama-French three factors)+ Industry factor														
MG	0.007	0.230	-0.122	-0.147				0.456	0.381	-0.018	-0.018	-0.008	-0.006	-0.000
	(0.002)	(0.025)	(0.018)	(0.018)				(0.026)	(0.018)	(0.003)	(0.002)	(0.002)	(0.002)	(0.001)
%sig	7.9%	80.5%	72.8%	80.2%				81.7%	84.8%	51.2%	39.0%	28.4%	28.7%	20.2%
N_p	394	394	394	394				394	387	387	387	387	387	387
(3) Spatial factor model (Fama-French five factors)+ Industry factor														
MG	0.006	0.246	-0.110	-0.155	0.107	0.198		0.449	0.386	-0.013	-0.017	-0.009	-0.008	-0.001
	(0.002)	(0.025)	(0.018)	(0.019)	(0.019)	(0.017)		(0.025)	(0.017)	(0.003)	(0.002)	(0.002)	(0.002)	(0.001)
%sig	6.9%	80.2%	72.1%	79.7%	69.3%	71.1%		82.5%	85.5%	46.5%	37.7%	27.6%	27.9%	19.9%
N_p	394	394	394	394	394	394		394	387	387	387	387	387	387
(4) Spatial factor model (Fama-French five factors plus Momentum)														
MG	0.006	0.257	-0.106	-0.148	0.103	0.194	-0.001	0.439	0.383	-0.013	-0.017	-0.009	-0.008	-0.001
	(0.002)	(0.026)	(0.018)	(0.018)	(0.019)	(0.017)	(0.007)	(0.025)	(0.018)	(0.003)	(0.002)	(0.002)	(0.002)	(0.001)
%sig	6.6%	80.5%	72.3%	78.9%	68.8%	68.8%	54.6%	80.7%	85.5%	45.7%	38.2%	26.9%	27.4%	18.9%
N_p	394	394	394	394	394	394	394	394	387	387	387	387	387	387

Table 17: **QML estimation results of Industry factors augmented models Equation 33 using full sample, with W constructed using threshold $m = 2$**

	(1) factor component							(2) spatial-temporal component						
	α	β_1	β_2	β_3	β_4	β_5	β_6	β_I	ψ_0	ψ_1	ψ_2	ψ_3	ψ_4	ψ_5
(1) Spatial CAPM+ Industry factor														
MG	0.009	0.282						0.414	0.340	-0.014	-0.023	-0.007	-0.007	0.002
	(0.002)	(0.024)						(0.021)	(0.017)	(0.003)	(0.002)	(0.002)	(0.002)	(0.001)
%sig	7.9%	83.2%						84.3%	84.6%	49.7%	44.5%	27.6%	30.2%	22.7%
N_p	394	394						394	384	384	384	384	384	384
(2) Spatial factor model (Fama-French three factors)+ Industry factor														
MG	0.008	0.259	-0.125	-0.138				0.459	0.354	-0.017	-0.017	-0.008	-0.006	-0.000
	(0.002)	(0.025)	(0.018)	(0.018)				(0.026)	(0.017)	(0.003)	(0.002)	(0.002)	(0.002)	(0.001)
%sig	8.1%	82.5%	73.3%	80.7%				82.0%	83.3%	50.3%	39.3%	27.9%	28.9%	20.3%
N_p	394	394	394	394				394	384	384	384	384	384	384
(3) Spatial factor model (Fama-French five factors)+ Industry factor														
MG	0.006	0.275	-0.113	-0.146	0.106	0.195		0.454	0.358	-0.013	-0.017	-0.009	-0.007	-0.001
	(0.002)	(0.025)	(0.018)	(0.019)	(0.019)	(0.017)		(0.025)	(0.017)	(0.003)	(0.002)	(0.001)	(0.002)	(0.001)
%sig	7.1%	82.7%	72.3%	79.7%	69.8%	70.3%		82.0%	83.9%	47.1%	38.0%	28.1%	27.6%	20.6%
N_p	394	394	394	394	394	394		394	384	384	384	384	384	384
(4) Spatial factor model (Fama-French five factors plus Momentum)														
MG	0.006	0.257	-0.106	-0.148	0.103	0.194	-0.001	0.439	0.383	-0.013	-0.017	-0.009	-0.008	-0.001
	(0.002)	(0.026)	(0.018)	(0.018)	(0.019)	(0.017)	(0.007)	(0.025)	(0.018)	(0.003)	(0.002)	(0.002)	(0.002)	(0.001)
%sig	6.6%	80.5%	72.3%	78.9%	68.8%	68.8%	54.6%	80.7%	85.5%	45.7%	38.2%	26.9%	27.4%	18.9%
N_p	394	394	394	394	394	394	394	394	387	387	387	387	387	387

Table 18: **QML estimation results of Industry factors augmented models Equation 33 using full sample, with W constructed using threshold $m = 3$**

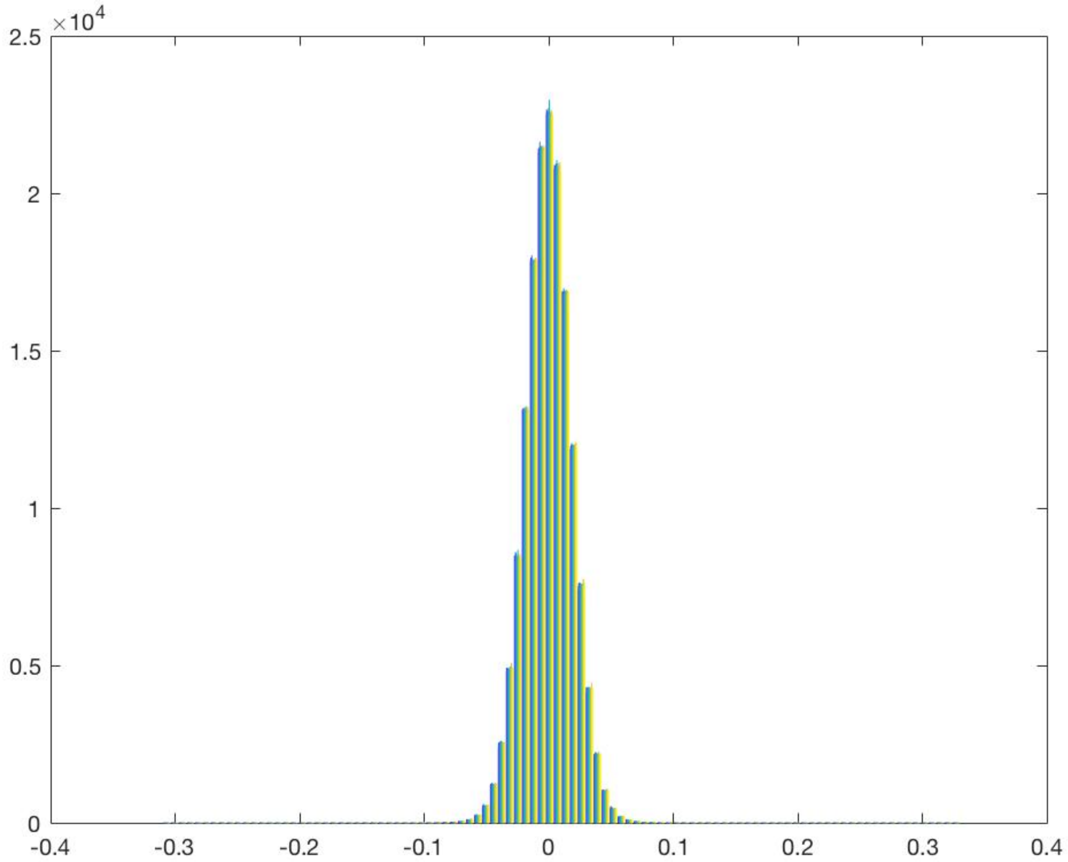

Figure 1: Histogram of bootstrapped $\hat{\rho}_{ij}^b$ for all $i \neq j, b = 1, \ldots, 500$.