Spatial Dependence and Real Estate Returns

STANIMIRA MILCHEVA* AND BING ZHU **

THIS VERSION: 22/12/2017

Abstract

We estimate a spatial factor model (SFM) by accounting for spatial linkages across returns of real estate companies using the physical distance between their properties. We show that the spatial factor model can better account for the cross-sectional variation across residuals as compared to a Fama-French-type factor model. Proximity across property holdings of pairs of firms can be used to model returns in addition to size, style, momentum and sector factors. Accounting for spatial linkages within an asset pricing model enhances the informational set and improves the model fit. The SFM can be used to disentangle spillover risks from market and idiosyncratic risks. The results show that the spillover risk varies considerably across regions and across time, rising sharply during the global financial crisis and being most pronounced in the US. While market risk can be low, implying good diversification potential, spillover risk may exist, neglecting of which, overstates the diversification benefits. Our results imply that investors looking for diversification should consider expose to firms with low market risk and low spillover risk, such as euro area real estate firms.

Keywords: Spatial factor model, asset pricing, listed real estate companies, spillover risk.

JEL codes: G12, C30

* University College London, 1-19 Torrington Place, London, WC1E 7HB, UK. s.milcheva@ucl.ac.uk.

** University of Reading, b.zhu@henley.reading.ac.uk.

I Introduction

While equilibrium spatial models can be used to account for unobservable characteristics of property prices by accounting for spatial linkages, they do not account for the stochastic nature of real estate returns. When analysing real estate markets, it is important to account for the demand stemming from investors and the associated risk premiums included in the real estate prices. In turn, in the investment literature, asset pricing models have been widely used to determine the required rate of return given the systematic risk but do not take into account the spatial dynamics of the underlying assets. When analyzing asset prices in isolation, the classical asset pricing models only account for the time-series variation of the asset with the factors. However, valuable information would be lost if some cross-sectional dependence exists across the residuals (Forni and Lippi, 2001; Kou et al., 2017). In a similar way, the assumption of mutually orthogonal idiosyncratic components can be violated, in particular during crisis periods. For example, the return of a pair of companies can show similar patterns and this comovement can be driven by the idiosyncratic component of the returns or their market premia. During turmoil periods, through spatial linkages, the effect of the market or idiosyncratic risk of companies in close proximity can be reinforced and lead to spillovers. Wieand (1999) develops a spatial equilibrium model under uncertainty in which the decision about the price and quantity of housing consumed is determined in a two-period asset-pricing context.[1] Some initial attempts are made to empirically show that spatial linkages matter in asset pricing, such as the papers by Fernandez (2011) and Kou et al. (2017). Fernandez (2011) estimates a 'spatial asset pricing model' using spatial matrices to account for the same information contained in the Fama and French factors instead of the using the factors directly. In other words, weights are defined using the variable value of Fama and French factors (e.g., firm size and market-to-book ratio). The empirical estimation is based on conventional spatial models, rather than spatial factor models. Fernandez (2011) reports strong evidence for the use of spatial indices instead. Kou et al. (2017) propose a spatial capital asset pricing model (S-CAPM) and a spatial arbitrage pricing theory (S-APT) using a spatial term in

[1] Wieand (1999) leaves the empirical assessment of expected returns, variances and covariances of locational payoffs for further research.

addition to the factors. They show an application of the models using regional stock indices and future contracts on S&P Case-Shiller home price indices. The weights are based on the geographic distance between regions of the respective regional indices, as this is an easy way to capture the spatial linkages between the assets. The authors find that the spatial interaction can explain the cross-sectional correlation. Our study follows on from Kou et al. (2017) and estimates a spatial factor model (SFM) for listed real estate companies accounting for the spatial dependence between single firms rather than indices. We demonstrate how a spatial factor model improves the model fit and also helps to disentangle spillover risks from market risks and idiosyncratic risks. This is in so far important as avoiding assets with high market risk is not enough for a good diversification strategy. Investors need to assess the spillover risks embodied in the total asset risk as those provide additional hurdle to diversification.

The cross-sectional dependence is captured by using techniques from spatial econometrics. The spatial approach presents the structure of a cross-sectional dependence in relation to the location and distance among units using a spatial weight matrix. The spatial weights are used to explicitly relate each unit to its neighbours. One of the challenges which also explains why a SFM has not been considered is the difficulty to associate a financial asset with a specific location. Fernandez (2011) uses firm characteristics, such as market capitalization (relative to firm size), the market-to-book, the dividend yield and the debt maturity ratios, to quantify the 'distance' between firms. Kou et al. (2017) try to overcome this problem by using regional indices instead – regional stock and house price indices. When it comes to company level data, Pirinsky and Wang (2006) were among the first to account for location in the context of an asset pricing model. They use the location of the company's headquarter. Bernile et al. (2015) construct a locational dispersion factor using the locations associated with a firm which have been mentioned in the company's financial reports. Becker et al. (2011) use the location of the large shareholders of a company to relate to the firm's performance. However, the above measures are far away from really capturing the spatial effects associated with the locations of the assets and can thus pain an incomplete picture of the link between firm performance and location and spatial interaction. We aim to address this issue by using companies whose performance is strongly related to the geographic location of their assets as is the case for real estate companies.[2] Such companies can function as funds who invest in real estate or can be run as operating companies. Most of the property companies are real estate investment trusts (REITs) which must derive a large proportion of their income (80%) from the operation of real estate assets and should pay out at least 90% of their taxable income to shareholders, and in exchange benefit from tax reductions. Since the performance of such companies is largely driven by the income (the rent) and the capital growth of the underlying assets, the location of the properties and the spatial interactions can be regarded as a key factor in company's valuation.

We use monthly data from 1998 to 2015 and estimate a spatial factor model as a panel using a Bayesian estimation. We find that the spatial factor model is not rejected and the spatial parameters are significant. Proximity across the property holdings of real estate companies can predict higher return correlation across the firms, controlling for size, book-to-market, momentum and sector characteristics. We show that adding spatial linkages across firms' holdings within an asset pricing model enhances the informational set and improves the model fit. Moreover, we perform a variance decomposition and show that while relatively low, the spillover risks strongly increase during the global financial crisis and explains up to 21% of total asset risk. We show that spillover risk varies considerably across regions and over time. US firms are most exposed to spillover risks while UK firms are exposed to market risk implying that investors should look for companies outside of the US and UK in order to better diversify their portfolios. The euro area provides the best destination when it comes to reducing the effects of spillover risks.

II Literature Review

This paper is related to three separate strands of literature. First, by accounting for spatial linkages across returns, the study is linked to the vast literature on spatial econometrics which explains asset prices by measures of proximity such as geographic

[2] Looking at direct real estate asset returns may be suboptimal when estimating an asset pricing model as real estate has different properties compared to the traditional investment assets, such as stocks and bonds. Real estate is characterized by high transaction costs, little liquidity, indivisibility, inability of short sales, etc. Therefore, the conventional asset pricing models may not be suitable to fully capture those risks. In order to overcome some of the above problems, we use listed real estate companies since their returns are known to capture the underlying real estate market fluctuations but also provide more liquidity, reduce transaction costs and mitigate indivisibility and short-selling issues.

distance. Second, the paper is closely related to the literature on asset pricing models accounting for locational factors. Third, the paper combines above techniques in one model and hence relates to the ideas in research accounting for the correlation across residuals.

While equilibrium spatial models can explain how prices of real estate in different locations are connected, they abstract from the stochastic nature of real estate returns, from the demand for real estate stemming from investors and the associated risk premiums included in the real estate prices. In turn, in the investment literature asset pricing models have been widely used to determine the required rate of return given the systematic risk but does not take into account the spatial dynamics of the underlying assets. Wieand (1999) develops a spatial equilibrium model under uncertainty in which the decision about the price and quantity of housing consumed is determined in a two-period asset-pricing context. The bid price of a house is a function of the homeowner's portfolio risk, including the risk associated with the site, and the market risk. More recently, Ortalo-Magne and Prat (2010) construct a spatial equilibrium model with a portfolio choice in an asset pricing context. The authors show how spatial decisions about where to buy a house can be seen as an expanded portfolio model in which the cross-sectional distribution of real estate dividends (i.e. rents) is endogenous and depends on the location-specific factors and the systematic risk. A house buyer chooses a certain location when he or she is indifferent to such factors as the benefits of a location associated with access to local amenities, income perspectives, the costs of the house price, and so on. As the agent is exposed to local productivity shocks, the location choice will depend on their income minus the rent – as is the case in spatial equilibrium models – but also on the correlation of their income with that of the other residents in the same location. The decision of how to allocate funds across different countries determines the expected returns everywhere, their volatility and covariance with the other assets, and the weight of assets from each country in the global market portfolio that is relevant for the pricing of all assets in the economy.[3] In spatial econometrics, the idea is to capture the effect of a shock at a specific

[3] Ortalo-Magne and Prat (2010) argue that the country REIT index is a suitable measure to track the housing demand in the model, which is the same for all agents, as it does not include the properties which are owned by local residents for hedging purposes.

point in space, in another place (Haining, 2003). The most common spatial dependence widely studied in the literature is through geographic proximity (Fingleton, 2001 and 2008).[4] The reason is that neighboring regions often keep close economic relationships. Therefore, as Fazio (2007) and Orlov (2009) argue, geographically closer regions would have as a result stronger economic linkages too. Miao et al. (2011) explore correlations among real estate returns in 16 US metropolitan areas and find that the strongest correlation appears to be in geographically adjacent regions. A similar result has been found for stock returns by Flavin et al. (2002). Ling et al. (2017) find a significant positive impact on REITs' returns stemming from the exposure to the so called 'Gateway' markets.

Given that we use spatial linkages in an asset pricing model, our paper relates to the studies that account for locational factors as drivers of returns. On the one hand, there is the vast literature looking at international CAPM and how the inclusion of global factors or factors of other regions can add to the model.[5] The rapid pace of financial market liberalization has highlighted the role of global shocks for the pricing of local assets since foreign investors can access the market more easily. Asset pricing models which include global factors instead of local factors have been presented by Karolyi and Stulz (2003), Bekaert et al. (2009), Hou et al. (2011) and Fama and French (2012) among others. Karolyi and Stulz (2003) show that modelling asset prices to account for domestic market variations only would underestimate the returns of those assets whose residual is positively correlated with the global market portfolio. Griffin (2002) instead argues that this should not necessarily be the case as there are mainly firm-level characteristics that explain the comovements in stock returns. Hou et al. (2011) examine the effect on firm-level characteristics on the cross-sectional and time-series variation in stock returns internationally compared to global and foreign components, using various factor models. They find that the local as well as the combination of local and foreign factor models produce lower pricing errors than their purely global counterparts. On the other hand, there is some scarce research on the role of locational factors for asset returns. A study by

[4] More recent literature explores the use of other measures of proximity such as financial and economic integration (Zhu et al., 2013; Milcheva and Zhu, 2016).

[5] Merton (1973), Solnik (1974), Grauer et al. (1976), Sercu (1980), Stulz (1981), and Errunza and Losq (1985) present an international asset pricing model accounting for other regions to which the local assets can be related.

Pirinsky and Wang (2006) look at the role of the location of the headquarters of the company for its stock returns in an asset pricing context. They find strong comovement across returns of companies located in the same geographic area. Hong et al. (2008) explain the difference in stock prices of companies located in different areas by the aggregate book value of the firms located and the aggregate risk tolerance of investors in that region. Bernile et al. (2015) develop a more sophisticated measure of local firm exposure using the company's annual 10-K fillings financial reports. They construct the measure of firm location by using the number of times a state is mentioned in a given report. They find that including a local market index, measured as described above, explains more of the time-series variation of the returns as compared to only using national market factors. The authors show that international portfolio decision and performance are affected by the information asymmetries created through spatially distributed information. Therefore investors prefer to invest in companies with greater local economic exposure and, overall, such assets perform better.

Among the most widely used asset pricing models is the CAPM which takes into account the price of systematic risk but ignores firm-specific risk. Laghi and Di Marcantonio (2016) extend the CAPM to quantify idiosyncratic risk related to firm characteristics such as firm size and value, operating costs, financial structure, etc. The authors show that the CAPM systematically underestimates the cost of equity of firms and accounting for above firm-specific features can capture unsystematic risks. Forni and Lippi (2001) argue that using a theoretical framework to model a large set of cross sections of time series data is hardly possible. They introduce the generalized dynamic factor model to account for serial correlation within and across individual processes and allow for non-orthogonal idiosyncratic terms. One way which has been used in the literature to model the dependence across variables is the use of vector autoregressive (VAR) models. However, such models are only suitable if we are dealing with a small number of time series variables. Forni and Lippi (2001) argue that using a theoretical framework to model a large set of cross sections of time series data is hardly possible. Pesaran and Tosetti (2011) expand the dynamic factor model in Forni and Lippi (2001) to a panel data model with common factors in which the idiosyncratic errors display spatial dependence. By including the spatial dimension, they account for both time-specific weak and strong cross-section dependence across returns.[6] They show that the part of the returns that display weak dependence can be fully diversified away when the portfolios are constructed using the spatial weights. The part of the asset returns that is strongly dependent can only be diversified away when the portfolios are constructed using the factor loadings as weights. Jourdain and Sbai (2012) present a model in which the local volatility of the index is first calibrated followed by the dynamics of each stock in a stochastic volatility model in a second step.

III Methodology

In finance, factor models are widely used to determine a theoretically appropriate required excess rate of return of an asset from the risk-free rate, if that asset is to be added to an already well-diversified portfolio, given the asset's sensitivity to macroeconomic factors. A factor model is given by:

$$\tilde{r}_{t,i} = \alpha_i + \sum_{k=1}^{K} \beta_{i,k} f_{t,k} + u_{t,i}, \tag{1}$$

with $\tilde{r}_{t,i}$, the excess return of asset i (i = 1,2.., N) in period t (t = 1,2.., T) calculated as $\tilde{r}_{t,i} = r_{t,i} - r_t^{rf}$ with r_t^{rf}, the risk-free rate in period t. $f_{t,k}$ denotes the kth common factor such as the excess market return, etc., with k = 1,2.., K. $u_{t,i}$ is the error term with $u_{t,i} \sim N(0, \nu_i^2)$. α_i represents the individual asset alphas and $\beta_{i,k}$ is the sensitivity of the ith asset to the kth factor. The classic factor model explains the time-series variation in asset returns. To capture the effect of the spatial linkages on returns in a spatial factor model, we need to account for the cross-sectional variation across returns adding a spatial term so that:

$$\tilde{r}_{t,i} = \alpha_i + \rho \sum_{j=1, j \neq i}^{N} w_{t,j,i} \tilde{r}_{t,i} + \sum_{k=1}^{K} \beta_{i,k} f_{t,k} + e_{t,i}, \tag{2}$$

where $w_{j,i.t}$ is the weight based on the 'distance' between each two assets j and i in year t, and $w_{j,i.t} = 0$ for $i = j$. $\sum_{j=1, j \neq i}^{N} w_{j,i.t} \tilde{r}_{t,i}$ is the weighted sum of the contemporaneous excess returns of the remaining firms in the sample. ρ is the spatial autoregressive coefficient capturing the degree of comovement across the returns. $e_{t,i}$ is the error term,

[6] Weak dependence at a given point in time is when the weighted average of returns converges to the expected quadratic mean once the number of assets (the cross section dimension) is increased. Otherwise, strong dependence exists.

which is heterogeneous across the assets, $u_{t,i} \sim N(0, \sigma_i^2)$. The spatial factor model can be represented as:

$$\begin{bmatrix} \tilde{r}_{t,1} \\ \vdots \\ \tilde{r}_{t,N} \end{bmatrix} = \begin{bmatrix} \alpha_{t,1} \\ \vdots \\ \alpha_{t,N} \end{bmatrix} + \rho \begin{bmatrix} 0 & \cdots & w_{t,1,N} \\ \vdots & \ddots & \vdots \\ w_{t,N,1} & \cdots & 0 \end{bmatrix} \begin{bmatrix} \tilde{r}_{t,1} \\ \vdots \\ \tilde{r}_{t,N} \end{bmatrix} + \begin{bmatrix} \beta_{1,1} & \cdots & \beta_{1,K} \\ \vdots & \ddots & \vdots \\ \beta_{N,1} & \cdots & \beta_{N,K} \end{bmatrix} \begin{bmatrix} f_{t,1} \\ \vdots \\ f_{t,K} \end{bmatrix} + \begin{bmatrix} e_{t,1} \\ \vdots \\ e_{t,N} \end{bmatrix}. \quad (3)$$

Equation (3) can also be expressed as:

$$\tilde{r}_t = \alpha + \rho W_{t,N} \tilde{r}_t + \mathrm{B} f_t + e_t, \quad (4)$$

where $\tilde{r}_t$ is an N×1 vector of excess returns of the N assets in period *t*. f_t is a K×1 vector of K global factors, which are the same for all returns. α is an N×1 coefficient vector of the asset alphas. B is an N×K matrix of the asset betas. e_t is an N ×1 vector of error terms.

Since the model in Equation (4) has the dependent variable both on the left and right-hand side, we rewrite Equation (4) into its reduced form such as:

$$(I_N - \rho W_{t,N}) \tilde{r}_t = \alpha + \mathrm{B} f_t + e_t, \quad (5)$$

We define $(I_N - \rho W_{t,N})^{-1} = V$ so that Equation (5) can be rewritten as:

$$\tilde{r}_t = V\alpha + V\mathrm{B} f_t + V e_t, \quad (6)$$

Since $V_t = (I_N - \rho W_{t,N})^{-1} = I_N + \rho W_{t,N} + \rho^2 W_{t,N}^2 + \rho^3 W_{t,N}^3 + \cdots$, Equation (6) implies a spatial multiplier effect on the asset excess returns (see Anselin, 2006 and LeSage and Pace, 2009). Not only the 'first-order neighbors', $\rho W_{t,N}$, but also 'neighbors' neighbors' are affected through the spatial multiplier effect, $\rho^2 W_{t,N}^2, \rho^3 W_{t,N}^3$, etc. In the end, the shock can have a feedback effect on the company of the origin of the spatial shock.

With regards to the estimation, we use Bayesian estimation with heteroskedastic error terms following LeSage (1997). The Bayesian estimation is formalized in the Appendix.

In equilibrium, under the assumption of $cov(f_{k,t}, e_t)$=0, the variance of the returns in the spatial factor model can be decomposed into the market risk of the asset and the idiosyncratic risk (Kou et al., 2017):

$$\Sigma = \bar{V} B \Psi B' \bar{V}' + \bar{V} \Xi \bar{V}', \quad (7)$$

where Σ is the covariance matrix of the returns which is the same as the covariance in the factor model. Ξ is the covariance matrix of the error terms in the spatial factor model. $\bar{V}$ =

$\frac{1}{T}\sum_{t=1}^{T} V_t = \frac{1}{T}\sum_{t=1}^{T}\left(I_t - \rho W_{t,N}\right)^{-1}$.

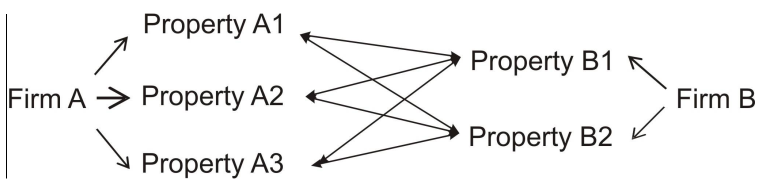
The spillover risk Θ is the part of the total asset risk in (7) which is associated with the spatial term V_t and is given by:

$$\Theta = \Sigma - B\Psi B' + \Xi, \quad (8)$$

The spatial weight matrix

In order to construct the spatial weight matrix, we use listed property companies. They provide a suitable setting for the spatial factor model as such companies extract a large proportion of their income (80–90%) from real estate, mostly through rents. This enables us to use the locations of the underlying assets of the company and to construct spatial weights for each of the firms. The location of the properties of each company is identified using the SNL database. Figure 1 shows an example of how the distance between asset A and B is calculated as the average across the individual distances between the properties held by the two firms. Let us assume that company A is invested in three properties, A1, A2, A3. and company B has two properties, B1, B2. Then, the weight between firm A and firm B will be an average of the weights across all combinations of above buildings – a total of 6 linkages.

Figure 1: Construction of the spatial weights between each pair of firms



When any firms hold more than one property, the distance is measured as the average distance:

$$D_{i,j,t} = \frac{1}{N1_t N2_t}\sum_{i=1:N1_t, j=1:N2_t} d_{i,j} \quad (9)$$

In the next step, we convert the D matrix to a corresponding continuity matrix C whose elements $c_{i,j}$ are defined as (Asgharian et al., 2013):

$$c_{i,j,t} = \begin{cases} \frac{max_{j,t} - D_{i,j,t}}{max_{j,t} - min_{j,t}} & if\ i \neq j \\ 0 & otherwise \end{cases}, \qquad (10)$$

The shorter the average distance between the two firms is, the larger the value $c_{i,j,t}$ has. The weight matrix *W* is then obtained from *C* through row standardisation, such that for each i, $\sum_j w_{i,j,t} = 1$.

IV Data

The data regarding the individual companies is collected from SNL Financial. In order to account for the distance between individual companies we will use the location of the properties they are invested in. We collect data for 120 listed property companies from US, UK and the euro area during the period from 1998M7 to 2015M3 which focus on the domestic market and report their property location. The idea is that we want to have a more condensed measure of the location of the underlying assets. This is especially important in the US where a property in Europe can lead to overestimating the weights local properties. Hence why, we exclude real estate companies which invest abroad.[7] Figure 2 shows the average number of properties held by US, UK and euro area real estate firms over time. US firms hold a significantly larger amount of properties than UK and euro area firms, with the number of underlying properties varying between 123 and 153 between 1998 and 2015. For UK and euro area firms, the number of properties started with 20 and 40 respectively in 1998 and grew by over 50% up until 2014.

Figure 2: Average number of properties held by real estate firms

[7] For US and UK, only firms that invest in their domestic market are included. For the euro area, firms investing only in Europe are included.

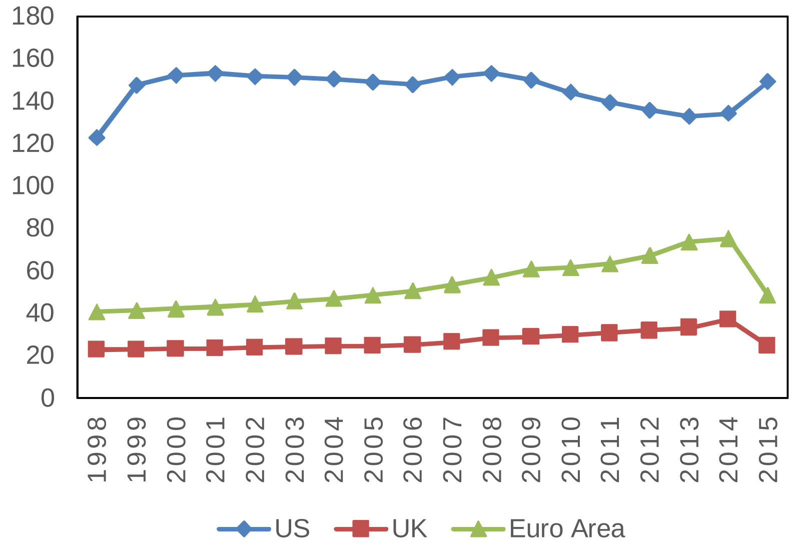


Table 1 shows the descriptive statistics of the listed real estate companies. On average, each US firm holds 139 properties, each British firm holds 27 properties, and each euro area company holds 54 properties. Over half of the companies (68) are based in the US, 21 firms are in the UK and 31 firms are in the euro area. The average total return of all firms is around 0.6% with a standard deviation between 11.7% and 9.5%. The highest volatility is observed in US.

Table 1: Descriptive statistics for the real estate companies (country averages 1998–2015)

Countries	Total return, mean	Total return, std. dev.	Total return, max	Total return, min	Average number of properties in a firm
US	0.0062	0.117	1.984	-2.413	139
UK	0.0063	0.095	0.954	-1.379	27
Euro Area	0.0064	0.102	1.008	-1.031	54

Explanatory variables include three Fama-French factors and the fourth factor is the Carhart momentum factor all obtained from Ken French's website (French, 2016)[8]. The four factors include a market return index (MR), the difference between the returns on

[8] See http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/

diversified portfolios of small stocks and big stocks (SMB), the difference between the returns on diversified portfolios of high book-to-market (value) stocks and low book-to-market (growth) stocks (HML), and the difference between the month *t* returns on diversified portfolios of the winners and losers of the past year (WML). In addition, we follow Ling et al. (2017) and include the EPRA/NAREIT return index to account for the real estate market performance.

Summary statistics of the factors are presented in Table 2. The average excess return of the market index over the sample period is 0.45% for US and 0.39% for Europe[9]. We can also see that the risk premium due to momentum effects is as high as 0.29% for US and 0.35% for Europe. Europe offers a large premium for value stocks of nearly 0.42% on average across the sample period. Instead, in the US, small stocks trade at a higher premium than in Europe (0.36%). The real estate index for US shows a higher average return of 0.37% and also a higher volatility of 8% than for the UK and the euro area. In the UK and euro area, the average REIT return is about 0.16%.

Table 2: Descriptive statistics for the factors (averages 1998–2015)

	Total return, mean	Total return, std. dev.	Total return, max	Total return, min
Fama and French Factors				
US				
MR	0.0045	0.0468	0.1135	-0.1723
SMB	0.0036	0.0335	0.1918	-0.1536
HML	0.0016	0.0343	0.1391	-0.1311
WML	0.0029	0.0303	0.1219	-0.1757
RE (US)	0.0037	0.0800	0.3833	-0.4611
Europe[10]				
MR	0.0039	0.0547	0.1367	-0.2202
SMB	0.0017	0.0226	0.0877	-0.0734

[9] The Fama and French factors only exist for Europe and the indices do not distinguish between the euro area and UK.

[10] The factors are for Europe, including both euro area countries and the UK.

HML	0.0042	0.0276	0.1133	-0.0912
WML	0.0035	0.0161	0.0569	-0.0451
RE (UK)	0.0016	0.0623	0.1714	-0.3198
RE (Euro Area)	0.0016	0.0503	0.1257	–0.2867

V Results

Table 3 shows the results for three separate regressions – US, UK and the euro area. If we use the Fama-French factors, the spatial coefficients are significant and range between 0.27 and 0.70. The significance of the spatial coefficients highlights the importance of the spatial linkages across listed real estate stocks for company returns. By adding the spatial component, the adjusted R-squared increases in all three specifications. In the US, the increase is the strongest, from 33% in the factor model to 39% in the spatial factor model. In the UK, the increase is from 35% to 37%. The lowest increase in observed in the euro area which may be because the properties' locations are more widespread and the comovements are smaller. We can see from the Bayesian Information Criterion (BIC) that the spatial factor model is preferred to the factor model in all three cases as it has a lower BIC. Table 3 also reports the average absolute value of the intercept, or the alpha, which is the average of the absolute alphas of the individual returns. In addition to the alpha, we include its average standard error and a GRS F-test by Gibbons et al. (1989) for the joint significance of the individual alphas. The average absolute value of alpha decreases in the spatial factor model. The standard error of the intercept is also smaller in the spatial factor model. In particular, in the results for euro area, the GRS becomes insignificant when the spatial term is added which shows that the SFM can lead to a better fit. In all spatial models, the GRS tests suggest insignificant intercepts.

Table 3: Model fit of the spatial factor model versus the factor model

Note: The model is estimated from 1998M7 to 2015M3. The dependent variable is the log difference of the excess return of real estate stocks in each month. ρ *is the coefficient for spatial dependence. Std(*ρ*) is the standard deviation of* ρ*. Adj. R2 is adjusted R-square, the average coefficient of determination in the panel model. BIC stands for the Bayesian Information Criteria.* $|\alpha|$ *stands for the absolute mean of the individual alphas. Std(*α*) stands for the average standard deviation of alpha. The GRS is a test for the joint significance of the firm alphas. For the US, the critical value at 10%, 5% and 1% significance level is 1.30,1.40, and*

1.61 respectively. For the UK, the critical value at 10%, 5% and 1% significance level is 1.40,1.54, and 1.83 respectively. For the euro area, the critical value at 10%, 5% and 1% significance level is 1.38,1.52, and 1.80 respectively.

	ρ	Std(ρ)	Adj. R2	BIC	| α|	Std(α)	GRS
			Panel A: US				
SFM	0.7002***	0.0124	0.3879	-2.1700	0.0038	0.0039	0.5823
FM			0.3329	-0.8395	0.0038	0.0071	0.5916
			Panel B: UK				
SFM	0.3477***	0.0233	0.3468	-2.3150	0.0040	0.0040	1.1160
FM			0.3237	-1.4180	0.0045	0.0057	1.4320*
			Panel C: Euro Area				
SFM	0.2724***	0.0257	0.2175	-2.0140	0.0058	0.0046	1.2140
FM			0.2124	-0.9963	0.0090	0.0071	1.8020***

Table 4 reports the average coefficients across firms based on the Fama and French factors and the real estate factor. Real estate companies can be driven by factors specific to the real estate market such as location, sector, etc. This is reflected in the low and insignificant betas of the Fama and French factors and the high and significant coefficient for real estate factor. One of the reasons for the low synchronicity between real estate companies and the market, as argued by Chung et al (2011), can be due to spatial uniqueness of the underlying assets. If we compare the results with and without the spatial term, we can see that the beta of the real estate factor in the spatial factor model is lower than that for the factor model.

Table 4: Estimation of factor models

Note: RM stays for the average coefficient associated with the index return; SMB is the average coefficient of the return differential of small-minus-big portfolios; HML is the average coefficient of a return differential of high-minus-low portfolios; MOM is the average coefficient of the momentum return index. RE is the EPRA/NAREIT index comprising of listed real estate companies. Average t-statistics are provided in brackets.

	RM	SMB	HML	MOM	RE
			US		
SFM	0.0598	-0.0184	-0.1610	0.0652	0.2229

	[0.6308]	[-0.1537]	[-1.2066]	[0.3315]	[3.7654]
FM	0.1669	-0.0944	-0.3969	0.1814	0.7386
	[0.9796]	[-0.3841]	[-1.6314]	[0.5177]	[8.5535]
			UK		
SFM	0.0646	0.1352	-0.0030	0.0445	0.3797
	[0.7841]	[0.7078]	[0.0477]	[0.2071]	[5.3270]
FM	0.1268	0.2480	0.0076	0.1266	0.6362
	[1.1988]	[1.0279]	[0.0340]	[0.3171]	[7.1537]
			Euro Area		
SFM	0.0119	0.0875	-0.0708	-0.0622	0.5196
	[0.1080]	[0.4351]	[-0.3034]	[-0.1568]	[5.0925]
FM	0.0201	0.0894	-0.1673	-0.1284	0.7792
	[0.1502]	[0.2942]	[-0.6078]	[-0.2596]	[5.5248]

VI Robustness checks

In order to see if our results depend on the choice of the factors, we perform robustness estimations to test whether the spatial matrix does a good job in capturing the comovement across the assets and make sure that this comovement is not associated with global variations driving all assets at the same time. If global shocks are the predominant reason for the cross-country comovement, the estimated spatial coefficient ρ can be very large, no matter what kind of weight matrix we choose. Therefore, we instead use a randomly generated weight matrix to assess whether the spatial dependence is captured by the physical distance between the properties or is due to common shocks. We estimate the equation:

$$\tilde{r}_{t,i} = \alpha_i + \rho^{random} \sum_{j=1,j\neq i}^{N} w_{t,j,i}^{random} \tilde{r}_{t,i} + \sum_{k=1}^{K} \beta_{i,k} f_{t,k} + e_{t,i}. \quad (11)$$

The estimation is rerun 200 times. The spatial coefficient ρ^{random} for the random matrix can be used to derive the confidence interval for the economic significance of ρ. A considerably larger spatial coefficient for the geographic distance matrix than that for the random matrix can reveal significant comovement effects not associated with global shocks. The results in Table 6 show that for all three regions, the confidence intervals for the spatial coefficient are lower than those for the baseline model. Thus, the variations in the stock returns can indeed be driven by geographic locations of the underlying properties and not by strong global comovements in the returns. This result also shows that the measure of distance of the underlying asset outperforms the majority of randomly generated weight matrices best capturing the spatial comovement.

Table 6: Robustness analysis controlling for unobserved global factors for the baseline specification

Note: The model is estimated from 1998M7 to 2015M3. The dependent variable is the log difference of the excess return of real estate stocks in each month. ρ is the contemporaneous spatial coefficient.

	US	UK	Euro Area
Random weight matrix ρ	[0.6643, 0.6815]	[0.2763; 0.3223]	[0.2115;0.2713]

The second issue is about spatial validation. Spatial models always assume that the co-movement between countries should depend on the strength of their linkages. Under this assumption, weights are constructed based on the strength of the linkages. However, this assumption has not been formally tested. In other words, it still needs to be checked whether countries with weaker financial or trade linkages do indeed have a lower degree of comovement than those countries with stronger linkages. Given this concern, we apply a distance decay model. This model explicitly checks whether the comovement decreases when the distance increases. We construct the weights according to the proportion of the properties locating within a certain range. We include five matrices into the regression. Matrix one is based on the range of within 50km. That means, each weight between a pair of firms reflects the proportion of properties of one firm that locates within 50km to any of the properties held by the other firm. Matrix two is defined in the same way as matrix one with the only difference being that the weight is defined based on the proportion of properties held by one firm that locates with a distance between 50 and 150km to any of the properties held by the other firm. In the same way, we define matrix three, but the distance is 150 –300km. Matrix four accounts for the proportion of the properties between two firms that locate between 300 – 450km. Matrix five accounts for the proportion of the properties between two firms that locate more than 450km away. For each matrix, the weight is defined as the proportion of the underlying properties within the respective distance. For the rest of the firms, the weights are set to zero.

Table 7 reports the results. For US, the matrix based on 50 km bandwidth has the highest coefficient, which is 0.449. The matrix based on the bandwidth between 50km and 150km has a coefficient of 0.103, which is much smaller than the weight matrix based on 50km bandwidth, both statistically and economically. When the distance is more than 150km, the impact is below 10%. The decrease in the spatial dependence intensity with different bandwidths implies that the comovement in the excess returns declines with the distance of the properties. The same conclusion can also be drawn in UK and euro area. The coefficient decreases significantly when the bandwidth increases. When the range is larger than 50km, the coefficient drops from 0.449 to 0.103 for the US, from 0.207 to 0.133 in the UK, and from 0.152 to 0.06 in the euro area.

Table 7: Robustness analysis controlling for unobserved global factors for the baseline specification

Note: The model is estimated from 1998M7 to 2015M3. The dependent variable is the log difference of the excess return of real estate stocks in each month. ρ is the contemporaneous spatial coefficient. ρ_50 represent the coefficient for weight matrix based on a bandwidth of within 50km. ρ_50_150 is for the weight matrix based on the bandwidth of between 50 and 150km. ρ_150_300 is for the weight matrix based on the bandwidth of between 150 and 300km. ρ_300_450 is for the weight matrix based on the bandwidth of between 300 and 450km. ρ_450 is for the weight matrix based on weights according to proportion of properties locating more than 450km.

	US	UK	Euro Area
ρ_50	0.449	0.207	0.152
	[0.400,0.494]	[0.178,0.239]	[0.114,0.193]
ρ_50_150	0.103	0.133	0.060
	[0.067,0.144]	[0.089,0.172]	[0.031,0.089]

ρ_150_300	0.035	0.018	0.039
	[0.014,0.055]	[-0.007,0.039]	[0.009,0.067]
ρ_300_450	0.073	0.016	0.075
	[0.044,0.098]	[-0.021,0.051]	[0.038,0.114]
ρ_450	0.030	0.010	0.017
	[0.013,0.049]	[-0.014,0.033]	[-0.011,0.047]

The above models are estimated using a Bayesian estimation which accounts for heteroskedasticity in the error terms. Other estimators include the Maximum Likelihood (ML)[11] estimator. Additionally, we estimate another specification of the spatial error model as suggested by Pesaran and Tosetti (2011) accounting for spatial dependence across the residuals. The spatial error model is given as:

$$\tilde{r}_{t,i} = \alpha_i + \sum_{k=1}^{K} \beta_{i,k} f_{t,k} + e_{t,i}, \quad (12)$$

with the error terms expressed as:

$$e_{t,i} = \lambda \sum_{j=1,j\neq i}^{N} w_{i,j,t} e_{t,i} + \varepsilon_{t,i}. \quad (13)$$

Compared with the spatial error model in Equation (13), the spatial lag model in Equation (4), which we use as our baseline, can take into account the comovement in both the error terms and the common factors. The results are reported in Table 8. The spatial coefficients based on spatial error model are not different from the spatial lag model. Moreover, the results using the Bayesian estimator are not considerably different from the other estimators.

Table 8: Robustness analysis using alternative estimators and model speciation

Note: The model is estimated from 1998M7 to 2015M3. The dependent variable is the log difference of the excess return of real estate stocks in each month. ρ *is the coefficient for spatial dependence. Std(*ρ*) is the standard deviation of* ρ*. Adj. R2 is the adjusted R-square, the average coefficient of determination in the panel model. BIC stands for the Bayesian Information Criteria.* |α| *stands for the mean of absolute alphas*

[11] IV and GMM estimators are also widely used in solving spatial panel models (e.g., Baltagi, and Liu, 2011). However, in a factor model, because the factor variables are the same for the cross section, the conventional instrument variables are not applicable. Finding suitable instrument variables remains for further research.

of the individual firms. Std(α) stands for the average standard deviation of the alphas. The GRS is a test for the joint significance of the firm alphas. ML stands for the maximum likelihood estimator. For US, the critical value at 10%, 5% and 1% significance level is 1.30,1.40, and 1.61 respectively. For UK, the critical value at 10%, 5% and 1% significance level is 1.40,1.54, and 1.83 respectively. For EU, the critical value at 10%, 5% and 1% significance level is 1.38,1.52, and 1.80 respectively.

	ρ	Std(ρ)	Adj. R2	BIC	|α|	Std(α)	GRS
Panel A US							
ML – Spatial factor model	0.6570***	0.0825	0.3867	-2.2750	0.0039	0.0067	0.5810
Bayesian - Spatial error model	0.6879***	0.0163	0.3879	-1.0820	0.0037	0.0061	0.5276
Panel B UK							
ML – Spatial factor model	0.3650***	0.0882	0.3475	-2.8230	0.0041	0.0055	1.4210*
Bayesian - Spatial error model	0.3600***	0.0320	0.3473	-1.4980	0.0044	0.0053	1.9730***
Panel C Euro Area							
ML – Spatial factor model	0.2020***	0.1387	0.2169	-2.3620	0.0081	0.0069	1.7970*
Bayesian - Spatial error model	0.2430***	0.0377	0.2173	-1.0320	0.0084	0.0068	2.2040***

To sum up, the results remain robust. We can show that the spatial matrix does a good job in capturing the comovement across the firms and the comovement is not associated with global variations driving all assets at the same time. Variations in the returns are driven by geographic locations of the properties and not by global shocks to the returns. Moreover, we see that the use of an alternative estimator, such as ML, and alternative model specification, such as a spatial error model, do not change the findings.

VII Spillover Effects

One application of the spatial factor model is that it can serve to disentangle the spillover risk from the overall asset risk. We define the spillover risk as the variation in returns which is due to the spatial comovement. We calculate the total asset risk, the idiosyncratic risk and the systematic risk using the variances of the returns, the residuals and the market factors, respectively. The total asset risk as estimated under the spatial factor model, is by definition, equivalent to the total asset risk as estimated under the factor model. However, a part of this risk can be attributed to spatial comovement across the returns and can hence be associated with spillover effects. The spillovers are larger for companies whose underlying assets are spatially 'closer' to each other in the broader sense as captured by the different weight matrices. Based on the estimated spillover intensity, we decompose the variance of the returns into three parts: market risk, idiosyncratic risk and spillover risk.

Table 9 shows the variance decomposition. The total variance is reported in the parentheses. The spillover risk is largest in the US with a share of 21% of total asset risk. In Europe, it seems that spillover risks play a less dominant role with 14% in the UK and 7% in the euro area. It is surprising that the US has a high proportion of spillover risk than a smaller and more compact country like the UK where a large proportion of the property holdings of the firms are concentrated in one city, London. However, the lower spillover risk in the UK as compared to the US may be explained with the sectoral specialisation of real estate companies. Most UK firms are specialised in one property sector and hence, little spillover through the location may occur. On the contrary, the majority of the US firms adopt a diversified portfolio strategy meaning that they invest in a variety of sectors. Hence there can be more correlation across the firms through the location of their properties. In the US, the average market risk is lowest across the three regions with 6% as compared to 17% in the UK and 14% in the euro area. The euro area has the highest share of idiosyncratic risk with 78% of total asset risk attributed to firm-specific variation. This may be due to the fact that firms' properties in the euro area are further apart hence and less exposed to spillover risks. In the UK the idiosyncratic risk is lowest with 70% suggesting that there is less room for diversification among UK firms and a large exposure to market risks. US lies in-between with a share of an asset-specific risk of 72%.

Table 9: Variance decomposition for the entire sample period, 1998-2015

This table reports the percentage variance of real estate equity return trigged by market risk, idiosyncratic risk and spillover risk based on Equation (8). The absolute value of the variance is reported in parenthesis.

	Market Risk	Idiosyncratic Risk	Spillover Risk
Panel A: US	6.27%	72.70%	21.02%
	(0.0007)	(0.0086)	(0.0025)
Panel B: UK	16.50%	69.57%	13.92%

	(0.0012)	(0.0052)	(0.0010)
Panel C: Euro Area	14.46%	78.23%	7.31%
	(0.0015)	(0.0082)	(0.0008)

In order to look at how the risks vary over time, we estimate the SFM using a rolling window of 60 months. The results are presented in Figure 3 and show the spatial coefficient over time along with the variance decomposition over time. In all three markets, idiosyncratic risk plays the most significant role in the majority of the time. In the US and the euro area, idiosyncratic risk takes up the largest proportion in every period. In the UK, this is also the case apart from the period between 2004 and 2007 in which market risk dominates. In the US, spillover risk is larger than the market risk in every period. Spillover risks was almost zero before 2007 and strongly increased in 2007 and remained high until 2012 when in reversed back to close to zero. Looking at the spatial coefficient, we also see a sharp increase in 2007 up to 0.8 that remains high up until 2012 mirroring the dynamics in the spillover risk. In the UK, the spillover risk has increased gradually between 2003 and 2007 but it takes up only a small proportion of total asset risk. The spatial coefficient has a similar dynamic, sharply rising to 0.45 in 2006 and then gradually falling to 0.15 in 2013. In the euro area, up until 2009 the spillover risk is close to zero and then slightly increases in 2010-2011. The spatial coefficient sharply rises starting in 2006 up to 0.6 in 2011. The euro area is dominated by the idiosyncratic risk which peaks during the GFC and then falls sharply in 2011. This risk dynamics reflects the uncertainties associated with the underling direct real estate market. The observation that the idiosyncratic risk has gradually fallen in the US and UK prior to the crisis may be associated with increased transparency and liquidity on the real estate market and the increased sophistication of the listed real estate market.

Overall, the market risk and the spillover risk becomes noticeable only during more turbulent periods as is the case during the GFC. The increase in the spillover risks mostly during more volatile periods suggests that it can be seen as evidence of contagion effects. The three regions have very different pattern in terms of when the spillover risk emerges and how long it takes to decrease. US is clearly the country which is most exposed to spillover risk implying that investors should look for companies outside of the US in order to better diversify their portfolios. The euro area provides the best destination when it comes to reducing the effects of spillover risks. What is also interesting is that the market risk in the UK is much higher than in any of the other countries during volatile periods meaning that UK investors should also consider getting exposure to other developed markets.

Figure 3: Variance decomposition based on a Spatial Factor Model

Figure 3.1 a: Estimated rho for US

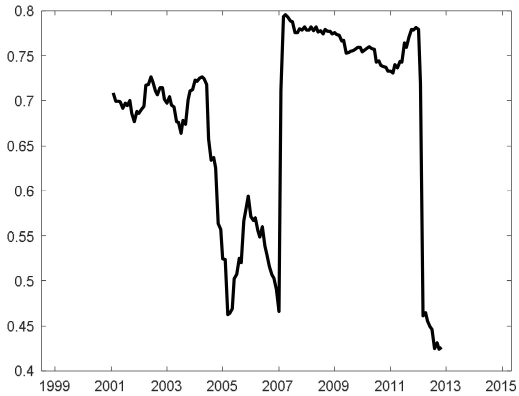


Figure 3.1.b: US Spillover risk, market risk and idiosyncratic risk

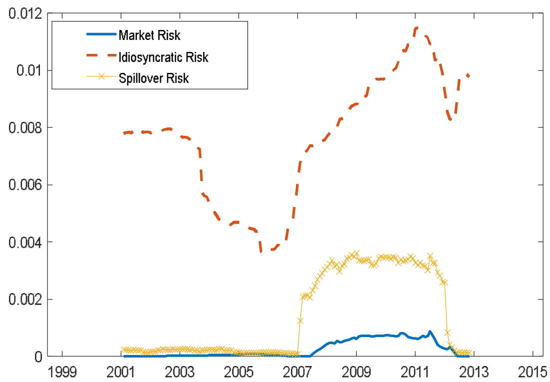


Figure 3.2 a: Estimated rho for UK

Figure 3.2.b: UK Spillover risk, market risk and idiosyncratic risk

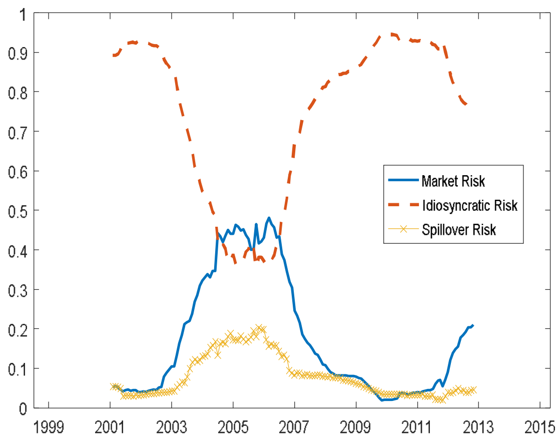
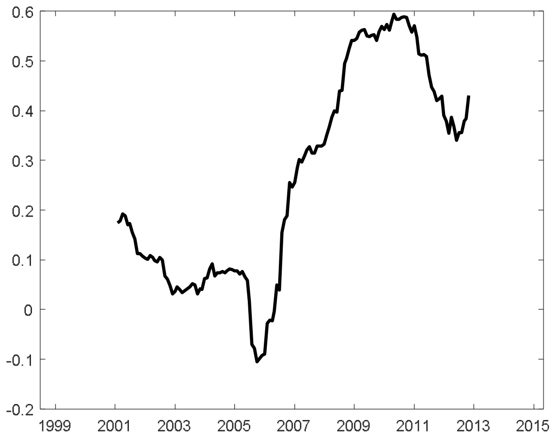
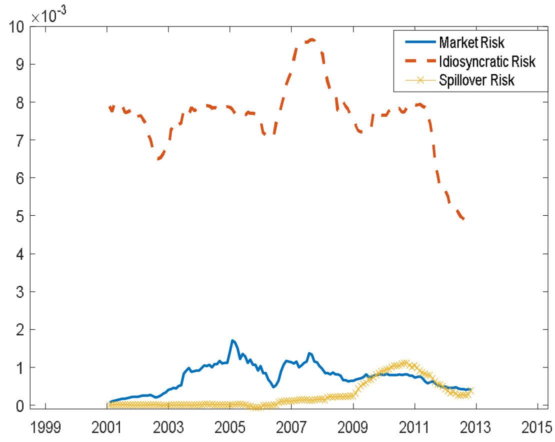


Figure 3.3 a: Estimated rho for Euro Area

Figure 3.3.b: Euro Area Spillover risk, market

risk and idiosyncratic risk





VIII Conclusion

We extend the four-factor model in Fama and French (2012) to incorporate spatial linkages across companies' underlying assets using firm-level data of listed real estate companies. We find that the spatial parameters in the spatial factor model are significant throughout the sample period and the model performs better than the factor model, considerably improving the model fit. Proximity across property holdings of pairs of firms can be used to model returns in addition to size, style, momentum and sector factors. The SFM can be used to disentangle spillover risks from market and idiosyncratic risks. We find that the spillover risk varies considerably across regions and across time with it rising sharply during the global financial crisis and being most dominant in the US. While market risks can be low implying good diversification potential, spillover risks also need to be accounted for as they can be high and mask the diversification benefits. Our results imply that investors looking for diversification should consider exposure to euro area firms which have low market risk and low spillover risk.

References

Anselin, L. “Spatial econometrics.” In *Palgrave Handbook of Econometrics, Econometric Theory,* Vol. 1, T. C. Mills and K. Patterson, eds. Basingstoke: Palgrave Macmillan (2006) 901–969.

Asgharian, H.; W. Hess; and L. Liu. “A Spatial Analysis of International Stock Market Linkages.” *Journal of Banking and Finance,* 37 (2013), 4738–4754.

Baltagi, B. H. and Liu, L. "An Improved Generalized Moments Estimator for a Spatial Moving Average Error Model," Economics Letters, Elsevier, vol. 113(3) (2011b), 282-284.

Becker, B.; H. Cronqvist; and R. Fahlenbrach. “Estimating the Effects of Large Shareholders Using a Geographic Instrument.” *Journal of Financial and Quantitative Analysis,* 46 (2011), 907–942.

Bekaert, G.; R. J. Hodrick; and X. Zhang. “International Stock Return Comovements.” *Journal of Finance*, 64 (2009), 2591–2626.

Bernile, G.; A. Kumar; and J. Sulaeman. “Home Away from Home: Geography of Information and Local Investors”. *Review of Financial Studies,* doi: 10.1093/rfs/hhv004 (2015).

Eder, A. and S. Keiler. “CDS Spreads and Contagion among Systemically Important Financial Institutions -- A Spatial Econometric Approach”. International Journal of Finance & Economics. (2015).

Fama, E., and K. French. “Size, Value, and Momentum in International Stock Returns.” *Journal of Financial Economics*, 105 (2012), 457–472.

Fazio, G. “Extreme Interdependence and Extreme Contagion between Emerging Markets.” *Journal of International Money and Finance*, 26 (2007), 1261–1291.

Fernandez, V. “Spatial Linkages in International Financial Markets”. *Quantitative*

Finance, 11 (2) (2011), 237–245.

Fingleton, B. "Equilibrium and Economic Growth Spatial Econometric Models and Simulations." *Journal of Regional Science*, 41 (2001), 117–141.

Fingleton, B. "A Generalized Method of Moments Estimator for a Spatial Model with Moving Average Errors, with Application to Real Estate Prices." *Empirical Economics*, 34 (2008), 35–57.

Flavin, T. J.; M. J. Hurley; and F. Rousseau. "Explaining Stock Market Correlation: A Gravity Model Approach." *The Manchester School Supplement*, (2002), 87–106.

Forni, M. and M. Lippi. "The Generalized Dynamic Factor Model: Representation Theory." Econometric Theory, 17(6) (2001), 1113–1141.

French, K. (2016) "Kenneth R. French" http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/ (2016)

Gibbons, M. R., S. A. Ross, and J. Shanken. "A Test of the Efficiency of a Given Portfolio." *Econometrica*, 57 (1989), 1121–1152.

Griffin, J. M. "Are the Fama and French Factors Global or Country Specific?" *Review of Financial Studies*, 15 (2002), 783–803.

Haining, R. *Spatial Data Analysis: Theory and Practice*. Cambridge University Press: Cambridge (2003).

Hong, H.; J. D. Kubik; and J. C. Stein. "The Only Game in Town: Stock-price Consequences of Local Bias." *Journal of Financial Economics*, 90 (2008), 20–37.

Hou, K.; G. A. Karolyi; and B. C. Kho. "What Factors Drive Global Stock Returns?" *The Review of Financial Studies*, 24 (2011), 2527–2574.

Jourdain, B. and M. Sbai. "Coupling Index and Stocks", Quantitative Finance, 12(5) (2012), 805 – 818.

Karolyi, G. A., and R. M. Stulz. “Are Assets Priced Locally or Globally?” In *Handbook of the Economics of Finance*, G. Constantinides, M. Harris, and R. M. Stulz eds. North Holland (2003).

Kou, S., X. Peng and H. Zhong. “Asset Pricing with Spatial Interaction”, *Management Science*, forthcoming (2017).

Laghi, E. and M. Di Marcantonio. “Beyond CAPM: Estimating the Cost of Equity Considering Idiosyncratic Risks”, 16(8) (2016), 1273 – 1296.

LeSage, J. “Bayesian Estimation of Spatial Autoregressive Models” *International Regional Science Review*, 20 (1-2) (1997), 113–129.

LeSage, J., and R. K. Pace. “Introduction to Spatial Econometrics.” Chapman & Hall/ CRC, Boca Raton (2009).

Ling, D. C, A. Naranjo and B. Scheick “Asset Location, Timing Ability, and the Cross-Section of Commercial Real Estate Returns.” *Real Estate Economics*, forthcoming (2017).

Lioui, A. and P. Maio. “Interest Rate Risk and the Cross Section of Stock Returns”. *Journal o Financial and Quantitative Analysis*. 49. (2014), 483-511.

Lopez, J. A. “Empirical Analysis of the Average Asset Correlation for Real Estate Investment Trusts”, *Quantitative Finance*. 9(2) (2009), 217 – 229.

Miao, H.; S. Ramchander; and M. C. Simpson. “Return and Volatility Transmission in U.S. Housing Markets.” *Real Estate Economics*, 39 (2011), 701–741.

Milcheva, S., and B. Zhu. “Bank Integration and Co-movements across Housing Markets.” *Journal of Banking & Finance*, 72S (2016), S148–S171.

Orlov, A. G. “A Cospectral Analysis of Exchange Rate Comovements during Asian Financial Crisis.” *Journal of International Financial Markets, Institutions & Money*, 19 (2009), 742–758.

Ortalo-Magne, F., and F. Prat. “Spatial Asset Pricing: A First Step.” CERP Discussion paper 7842 (2009).

Pesaran, M. H., and E. Tosetti. “Large Panels with Common Factors and Spatial Correlation.” *Journal of Econometrics*, 161 (2011), 182–202.

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

Wieand, K. “The Urban Homeowner's Residential Location Decision in an Asset-pricing Context.” *Real Estate Economics*, 27 (1999), 649–667.

Appendix: Bayesian estimation of the spatial factor model

Following LeSage (1997 and 2009), we use a Bayesian estimation to estimate the spatial factor model.

$$\tilde{r}_t = \alpha + \rho W \tilde{r}_t + \mathrm{B} f_t + e_t, \qquad \text{(A1)}$$

with t=1,2,..T. $\tilde{r}_t$ is an N by 1 vector of dependent variables. W is an N by N matrix. f_t is a K by 1 vector of the common factors and e_t is a N by 1 vector of error terms. α is an N by 1 vector of intercept $\alpha = [\alpha_1, \alpha_2, ..., \alpha_N]'$. and B is an N by K matrix of coefficient and

$$\mathrm{B} = \begin{bmatrix} \beta_{11} & \beta_{12} & \cdots & \beta_{1K} \\ \beta_{21} & \beta_{22} & \cdots & \beta_{2K} \\ \vdots & & & \vdots \\ \beta_{N1} & \beta_{N2} & \cdots & \beta_{NK} \end{bmatrix}.$$

Equation (A1) can be written as:

$$y = \rho (W \otimes I_T) y + X\gamma + e. \qquad \text{(A2)}$$

where y is an (NT x 1) vector of returns stacked by time, and $y = [\tilde{r}_{11}, \tilde{r}_{21}, ..., \tilde{r}_{N1,} \tilde{r}_{12}, ..., \tilde{r}_{N2}, ..., \tilde{r}_{NT}]'$. e is an (NT x 1) vector of normally distributed random variables with non-constant variance. $X = [1_T \quad f_1 \quad f_2 \quad \cdots f_K] \otimes I_N$, and γ is a $N(K$+1) by 1 vector of the coefficient and $\gamma = [\alpha_1, \alpha_2, ..., \alpha_3, \beta_{11}, \beta_{21}, ..., \beta_{N1}, \beta_{12,} ..., \beta_{N2}, ..., \beta_{NK}]'$.

The information prior for the heteroscedastic linear regression model can be written as:

$$
\begin{aligned}
&e \sim N(0, \sigma^2 V) \\
&V = diag(v_1, v_2, ..., v_N) \otimes I_T \\
&\rho \sim N(c, L) \\
&\sigma^2 \sim (1/\sigma) \\
&q / v_i \sim ID\chi^2(q) / q \\
&\sigma \sim \Gamma(\upsilon_0, d_0)
\end{aligned}
$$
.

where V is the relative variance terms and is assumed to be fixed but unknown parameters that need to be estimated. The prior distribution for v_i take the form of an independent distribution $\chi^2(q)/q$.

We allow for an informative prior on the spatial autoregressive parameter ρ, the heteroscedastic control parameter q and the disturbance variance σ. The diffuse prior for ρ is implemented by setting the prior mean c to zero and prior mean c to zero and prior variance of L, which is set as $1e+12$. The diffuse prior for σ is set to $\upsilon_0 = 0$, $d_0 = 0$.

The parameter q allows the v_i estimates to deviate from their prior means of unity. Small values for q allow for non-constant variance and are associated with a prior belief that outliers or non-constant variance exist. Large values for q would produce v_i estimates that are all close to unity, forcing the model to take on a homoscedastic character. Following LeSage (1997 and 2009), q is set to be 4.

The posterior distribution is based on the likelihood function:

$$L(\rho, \gamma, \sigma^2, v; y, W) = \sigma^{-NT} \left| I_{NT} - \rho(W \otimes I_T) \right| \prod_{i=1}^{NT} v_i^{-1/2} \exp\left[-\sum_{i=1}^{NT} (e_i^2 / 2\sigma^2 v_i) \right]. \qquad \text{(A3)}$$

The posterior density kernel function is

$$p(\rho, \gamma, \sigma, V) = \ln \left| I_{NT} - \rho(W \otimes I_T) \right| \prod_{i=1}^{NT} v_i^{-(q+3)/2} \exp(-q / 2v_i) \cdot \sigma^{-(n+1)} \exp\left[-\sum_{i=1}^{NT} (\sigma^{-2} e_i^2 + q) / 2v_i \right]$$

.

(A4)

The conditional distribution for parameter σ assuming that we know the parameters *γ*, *ρ* and *V* is:

$$\left[\sum_{i=1}^{TN}(e_i^2/v_i)/\sigma^2\right]\Big|(\gamma,\rho,V)\sim\chi^2(N)\,. \qquad \text{(A5)}$$

The conditional distribution for Β assuming that we know *σ*, *ρ* and *V* would be:

$$\begin{aligned}&\gamma\,|\,(\sigma,V)\sim N[H(\dot{X}'V^{-1}\dot{y}+\sigma^2Q'L^{-1}c),\sigma^2H]\\&H=\left(\dot{X}'V^{-1}\dot{X}+Q'L^{-1}Q\right)^{-1}\\&\dot{X}=[I_{NT}-\rho(W\otimes I_T)]X\\&\dot{y}=[I_{NT}-\rho(W\otimes I_T)]y\end{aligned} \qquad \text{(A6)}$$

The posterior distribution of v_i is set as:

$$(\sigma^{-2}e_i^2+q)/v_i\sim\chi^2(q+1)\,. \qquad \text{(A7)}$$

Conditioning on *σ*, *B*, and v_i, we have:

$$p(\rho\,|\,\gamma,\sigma,V)\propto\left|I_{NT}-\rho(W\otimes I_T)\right|\exp\left\{-(1/2\sigma^2)(e'V^{-1}e)\right\}. \qquad \text{(A8)}$$

The model is run with 2000 iterations.