Title: Expected Stock Returns, Common Idiosyncratic Volatility, and Average Idiosyncratic Correlation

Author names and affiliations:

Xuanming Ni[1]

Long Qian[2]

Huimin Zhao[3*]

Jia Liu[4]

1. School of Software and Microelectronics, Peking University, Beijing 100871, China;
2. School of Economics and Management, Tsinghua University, Beijing 100084, China;
3. School of Business, Sun Yat-sen University, Guangzhou 510275, China;
4. Business School, University of Portsmouth, Portsmouth, PO1 3DE, United Kingdom

JEL Classification Code: E44, G12

[*] **Corresponding author**, E-mail: zhaohuim@sysu.edu.cn (Huimin Zhao)

E-mail Address: nixm@ss.pku.edu.cn (Xuanming Ni), qianl20@mails.tsinghua.edu.cn (Long Qian), jia.liu@port.ac.uk (Jia Liu).

Acknowledgement: We acknowledge helpful comments from Prof. Youwei Li and 2019 International Conference on Internet Finance in Hangzhou.

Foundations: The research of this paper has been partially supported by grants from the National Nature Science Foundation of China (Project No.71991474) and the Fundamental Research Funds for the Central Universities (Project No.31620527).

Expected Stock Returns, Common Idiosyncratic Volatility, and Average Idiosyncratic Correlation

Abstract

Motivated by Herskovic et al. (2016), we examine the role of the average idiosyncratic correlation (ICOR) in two types of markets: an emerging market and a developed market. Examining daily stock data from the Chinese stock market for the period 1926 to 2019 and from the US for the period 1995 to 2020, we adopt high-dimensional principal component analysis (PCA) and thresholding methods to re-estimate ICOR. We find that ICOR plays an important role in explaining the expected stock returns, as the common idiosyncratic volatility (CIV) does in Herskovic et al. (2016). ICOR has been neglected in the literature due to large estimation error in the idiosyncratic covariance matrix and our analysis provides evidence that ICOR is nonnegligible in both markets when we control for several common market factors. We show that the average idiosyncratic covariance, which is the numerator of ICOR, presents the same patterns as CIV. Furthermore, our regression analyses of expected stock returns in response to ICOR change in both markets show that, in contrast to the negative result for CIV, the stocks' high risk exposure to ICOR change comes with a higher risk premium, perhaps because of the synchronized but disproportionate changea in the monthly idiosyncratic covariance and idiosyncratic volatility.

Keywords: idiosyncratic volatility, PCA, idiosyncratic correlation, high-dimensional covariance estimator.

2010 MSC: 00-01, 99-00

1. Introduction

Ang et al. (2006) find that stocks with high idiosyncratic risk predict lower average return. This empirical finding, called "idiosyncratic puzzle", challenges traditional asset-pricing theory in that a portfolio's idiosyncratic risk can be diversified away, and thus, it is uncorrelated with the expected return under the capital asset pricing model (CAPM) and arbitrage pricing theory (APT) frameworks; otherwise it should be related to a higher expected return (Merton, 1987). Even though the information in the idiosyncratic covariance matrix plays a very important role in asset pricing, thus far stduies have largely focused on the diagonal elements, that is, the idiosyncratic volatility (IVOL) (Duarte et al., 2014; Herskovic et al., 2016; Chen and Strebulaev, 2018). Whether and how the off-diagonal elements are related to the pricing of the expected return remain unexplored. In this study, we adopt a new method to re-estimate the idiosyncratic covariance matrix and determine if and how the off-diagonal elements matter in pricing stock returns.

A number of studies investigate the idiosyncratic puzzle considering various aspects, such as higher moments and firm characteristics (Boyer et al., 2009; Avramov et al., 2013; Hasan and Habib, 2017). Duarte et al. (2014) price the IVOL using principal component analysis (PCA) and examine its commonality from a cross-sectional perspective. Herskovic et al. (2016) calculate the common IVOL (CIV) by averaging the IVOL cross-sectionally to measure the comovement of IVOL between stocks. Their results suggest that the increasing CIV in the course of US market recessions is synchronized with the rise of firms' cash flow risk. They show that IVOL displays the same comovement even after 10 principal components (PCs) have been removed. Chen and Strebulaev (2018) demonstrate synchronized idiosyncratic risk by dynamically modelling risk shifting over the business cycle. However, these studies mainly focus on the IVOL, which is estimated by the square root of the diagonal elements of the idiosyncratic covariance matrix. In this paper, we take a new angle by examining the average idiosyncratic correlation (ICOR), which is the average of the off-diagonal elements of the idiosyncratic correlation matrix. Although the average covariance between idiosyncratic returns is smaller than the idiosyncratic volatiltiy, we find that ICORs are not close to zero in either emerging and well-develpoed markets, which are represented in this study by the Chinese and US stock markets, respectively, until 100 or more PCs have been removed from the stock returns. Herskovic et al. (2016) and Chen and Strebulaev (2018) find that the comovement of CIV is caused by firm-level idiosyncratic macroeconomic shocks. However, there is insufficient evidence that these macroeconomic shocks do not affect the covariance of idiosyncratic returns. We discover that the market-wide high-dimensional idiosyncratic covariance matrix contains a large number of estimation errors with most of the elements being close to zero. After removing these small errors, the average idiosyncratic covariance shows the same patterns as CIV, with a correlation of 0.63 in the Chinese stock market and 0.81 in the US stock market. During market recessions, ICOR remains steady in the Chinese stock market but declines in the US stock market. The cause of this difference is straightforward. The steadiness in the Chinese stock market is due to the proportionate surge in the numerator (idiosyncratic covariance) and denominator (idiosyncratic volatilities) if ICOR, and the decline in the US stock market is due to much larger estimation errors in the idiosyncratic covariance and higher market efficiency, which decreases the speed at which the numerator increases.

We adopt two high-dimensional methods to estimate ICOR. The first is the principal orthogonal complement thresholding (POET) method proposed by Fan et al. (2013). The second is asymptotic principal component analysis (APCA), proposed by Connor and Korajczyk (1986, 1988), which is also used by Duarte et al. (2014) to estimate IVOL. We conduct portfolio analysis and Fama and MacBeth (1973) regression analysis to examine the relation between the expected return and ICOR change/innovation. Our results show that stocks with more positive ICOR beta can earn higher expected returns in both markets, which is opposite to the result of Herskovic et al. (2016). However, our results are consistent with Herskovic et al. (2016) in that the large estimation error contained in the monthly estimators lowers the change rate of idiosyncratic covariance further than that of idiosyncratic volatility. Thus, the sign of ICOR change is opposite to that of CIV changes and hence, we obtain the opposite outcome.

Our study makes original contributions to the asset pricing literature. First, our study provides empirical evidence that ICOR contains important pricing information, with the empirical results showing that the innovations of ICOR and CIV are both priced factors. Second, we show that our proposed method can be applied to both developed and emerging markets, such as the US and Chinese stock markets, respectively.

The remainder of this paper is organized as follows. Section 2 discusses the definitions of CIV and ICOR and the methodologies. Section 3 discusses the data construction process and parameter settings. Section 4 reports and discusses the empirical results. Section 5 concludes the paper.

2. Common Idiosyncratic Volatility and Idiosyncratic Correlation

In this section, we first define CIV and ICOR, and then discuss how to construct CIV and ICOR using high-dimensional PCA methods.

2.1. CIV and ICOR

Idiosyncratic volatility is defined as the standard error of idiosyncratic returns, which captures the variation unexplained by some common factors (Duarte et al., 2014; Herskovic et al., 2016; Chen and Strebulaev, 2018). IVOL is constructed as follows. Consider a Fama and French (1992) three-factor model:

$$R_{i,t} = \alpha_i + \beta_i^{MKT} MKT_t + \beta_i^{\text{SMB}} SMB_t + \beta_i^{HML} HML_t + \varepsilon_{i,t}, \tag{1}$$

where IVOL can be easily estimated by $IVOL_{i,t} = \sqrt{\text{Var}(\varepsilon_{i,t})}$. Herskovic et al. (2016) calculate CIV as the average IVOL of all stocks, which is

$$CIV_t = \frac{1}{n} \sum_{i=1}^{n} IVOL_{i,t}, \tag{2}$$

where CIV represents the significant volatility comovement even after removing all common variation in returns. The correlation matrix of idiosyncratic return is

$$\mathbf{R}_t = \begin{pmatrix} 1 & \mathrm{Corr}(\varepsilon_{1,t}, \varepsilon_{2,t}) & \cdots & \mathrm{Corr}(\varepsilon_{1,t}, \varepsilon_{n,t}) \\ \mathrm{Corr}(\varepsilon_{2,t}, \varepsilon_{1,t}) & 1 & \cdots & \mathrm{Corr}(\varepsilon_{2,t}, \varepsilon_{n,t}) \\ \cdots & \cdots & \cdots & \cdots \\ \mathrm{Corr}(\varepsilon_{n,t}, \varepsilon_{1,t}) & \cdots & \cdots & 1 \end{pmatrix} \quad (3)$$

where $\mathrm{Corr}(\varepsilon_{i,t}, \varepsilon_{j,t}) = \mathrm{Cov}(\varepsilon_{i,t}, \varepsilon_{j,t})/(IVOL_{i,t}IVOL_{j,t})$.

The average idiosyncratic correlation, ICOR, is the average of all off-diagonal elements of $\mathbf{R}_t$, whereas CIV focuses on the variances, which are the diagonal elements of the idiosyncratic covariance matrix $\mathbf{\Sigma}_{\varepsilon_t} = \mathrm{Cov}(\varepsilon_t)$. It is common for $\mathbf{\Sigma}_{\varepsilon_t}$ to be high-dimensional. For example, if we use daily Chinese stock market data to estimate ICOR or CIV, the sample size is normally around 240 days for one year, but the number of available stocks in the whole market is far higher than sample size, usually more than 1000. It is therefore more much difficult to accurately estimate off-diagonal elements than diagonal ones because of the relatively small sample size. The estimation error will be huge if we use the sample covariance estimator of ε_t. We first adopt the Fama and French (1992) three-factor model (FF3), market factor model (Rm-Rf), APCA, and POET to estimate the high-dimensional idiosyncratic correlation matrix. To lower the estimation error, we adopt the thresholding method to threshold the sample covariance estimator of ε_t. This method requires a predefined thresholding parameter and then sets off-diagonal elements to zero if their absolute values are less than the thresholding parameter. This method dramatically reduces the number of estimations, mitigating the shortage of samples when the number of stocks is far larger than the sample size.

2.2. POET and APCA Method

We adopt the adaptive thresholding[1] method proposed by Fan et al. (2013) to resolve the high-dimensional problems in the estimation of the idiosyncratic

[1]Thresholding means to discard some off-diagonal elements in order to reduce the overall estimation error; then, the covariance matrix will become a sparse matrix.

covariance matrix. $\widehat{\mathbf{\Sigma}}_{\varepsilon}^{\mathcal{T}}$ is the thresholding estimator of the idiosyncratic covariance matrix $\mathbf{\Sigma}_{\varepsilon}$:

$$\widehat{\mathbf{\Sigma}}_{\varepsilon}^{\mathcal{T}} = \left(\hat{v}_{ij}^{\mathcal{T}}\right)_{n\times n}, \quad \hat{v}_{ij}^{\mathcal{T}} = \hat{v}_{ij} I\left(|\hat{v}_{ij}| \geq \tau_{ij}\right) \tag{4}$$

where $\hat{v}_{ij}$ represents the elements of $\widehat{\mathbf{\Sigma}}_{\varepsilon}$. I is an indicator that equals one if $(|\hat{v}_{ij}| \geq \tau_{ij})$ is true; otherwise, it equals zero. τ_{ij} is an entry-dependent threshold parameter, which is defined as $\tau_{ij} = \tau\left(\hat{v}_{ii}\hat{v}_{jj}\right)^{1/2}$, where $\tau > 0$ is a parameter that determines the intensity of thresholding. We next discuss how we attach different values of τ to calculate ICOR.

PCA is often used to calculate IVOL (Duarte et al., 2014; Herskovic et al., 2016). We use a high-dimensional PCA method POET (Fan et al., 2013), to deal with the high-dimensional situation. In addition, we adopt APCA method (Connor and Korajczyk, 1986) for comparison. The estimation process is as follows. Suppose that $\widehat{\lambda}_1 \geq \hat{\lambda}_2 \geq \cdots \geq \widehat{\lambda}_n$ are the eigenvalues of the sample covariance estimator $\widehat{\mathbf{\Sigma}}_R$ of the $T \times n$ return matrix $\mathbf{R}$, and $\left\{\widehat{\boldsymbol{\xi}}_i\right\}_{i=1}^{n}$ are corresponding eigenvectors. We decompose the $n \times n$ matrix $\widehat{\mathbf{\Sigma}}_R$ into a low-rank matrix plus a sparse full-rank matrix:

$$\widehat{\mathbf{\Sigma}}_R = \sum_{i=1}^{\widehat{r}} \widehat{\lambda}_{\xi} \widehat{\boldsymbol{\xi}}_i \widehat{\boldsymbol{\xi}}_i' + \widehat{\mathbf{\Sigma}}_{\varepsilon} = \widehat{\mathbf{\Sigma}}_c + \widehat{\mathbf{\Sigma}}_{\varepsilon}, \tag{5}$$

where $\widehat{\mathbf{\Sigma}}_{\varepsilon} = \sum_{i=\widehat{r}+1}^{n} \widehat{\lambda}_i \widehat{\boldsymbol{\xi}}_i \widehat{\boldsymbol{\xi}}_i' = \left(\hat{v}_{ij}\right)_{n\times n}$ is the principal orthogonal complement estimator or idiosyncratic covariance estimator. $\widehat{\mathbf{\Sigma}}_c = (\widehat{\sigma}_{ij})_{n\times n}$ is the common covariance matrix that captures the variation of return caused by $\hat{r}$ leading common factors. Here we let $\hat{r} = 3$. The thresholding estimator $\widehat{\mathbf{\Sigma}}_{\varepsilon}^{\mathcal{T}}$ of $\widehat{\mathbf{\Sigma}}_{\varepsilon}$ is the POET residual covariance estimator. Here, ICOR is the average of all off-diagonal elements of $\widehat{\mathbf{\Sigma}}_{\varepsilon}^{\mathcal{T}}$ and CIV is the average of all diagonal elements of $\widehat{\mathbf{\Sigma}}_{\varepsilon}^{\mathcal{T}}$.

Instead of calculating the $n \times n$ sample covariance matrix $\widehat{\mathbf{\Sigma}}_R$, which is not full-rank, APCA first calculates a $T \times T$ covariance matrix:

$$\widehat{\mathbf{\Omega}}_T = \frac{1}{k}\left(\boldsymbol{R} - \mathbf{1}_T\overline{\boldsymbol{r}}'\right)\left(\boldsymbol{R} - \mathbf{1}_T\overline{\boldsymbol{r}}'\right)', \tag{6}$$

where $\mathbf{1}_T$ is the $T \times 1$ vector of one and $\bar{\boldsymbol{r}} = (\bar{r}_1, \ldots, \bar{r}_n)'$, with $\bar{r}_i$ being the mean of the ith column of the return matrix $\mathbf{R}$. Then, APCA estimates the refined statistical factor estimators through the normal PCA process and then estimates the idiosyncratic returns using the OLS regression (Connor and Korajczyk, 1988; Tsay, 2005).

3. Data, Variable Construction, and Parameter Settings

We use daily return data and Fama-French three-factor data for 4,018 stocks[2] from the Chinese A Share market, ranging from 1995-11-27 to 2020-09-30, and 26883 stocks from US stock market, ranging from 1926-01-02 to 2019-12-31. The Chinese stock data are acquired from the RESSET database, and the US data are collected from the CRSP and Kenneth French's data library.

For the Chinese stock markets, we exclude stocks recently listed on the Science and Technology Innovation Board (STAR) to avoid the issue of insufficient data. Therefore, the common ordinary stocks under examination are from the Main Board, Growth Enterprise Market (GEM) Board, Small and Medium Enterprise (SME) Board, and New Third Board. For the US stock market, we examine ordinary common stocks and keep CRSP securities from the NYSE, AME, and Nasdaq with share codes 10, 11, and 12.

In estimating yearly and monthly ICOR and CIV, we keep stocks that have missing return data for no more than 10 days within a year and no more than two days within a month to ensure sufficient data for estimation. To examine the ICOR pattern at the industry level, we estimate the market ICOR and industry ICOR according to the Industry Classification Standard of the China Securities Regulatory Commission (CSRC) for Chinese stocks. We group US stocks using the Standard Industrial Class (SIC) for the period of 1926 to 1997 and the North American Industry Classification System (NAICS)[3] for the period of 1998 to 2019.

[2]The number 4,018 is the number of tradeable stocks in the A Share market on 2020-09-30.
[3]NAICS replaced SIC in 1997.

The parameters in estimation models Equation 4 to Equation 6 are set as follows. We adopt yearly stock data To estimate industry level ICOR. We show how ICOR varies under different values of the thresholding parameter τ in Subsection 4.1, and we set it to 0.3 as the default. We adopt monthly data to estimate ICOR to conduct empirical asset pricing. However, the typical sample size to estimate monthly ICOR is only about 20 days. If we estimate the same number of off-diagonal elements in the monthly covariance matrices as in the yearly covariance matrices, the estimation error will be enormous. Thus, to lower the estimation error, τ is set to 0.7, to threshold more off-diagonal elements and to generate much more sparse idiosyncratic correlation matrices than in the case of yearly ICOR estimation. APCA and POET follow the same process of first eliminating PCs and then thresholding the remaining idiosyncratic covariance matrix. With respect to the choice of the number of PCs, we show that ICORs are not very sensitive to the number of PCs. To be consistent with the normal method of calculating IVOL, which is to use the Fama and French (1992) three-factor model, the number of PCs is set to three by default.

4. Empirical Results and Discussions

In this section, we first examine the patterns of ICOR and CIV in the Chinese and US stock markets. We then examine the relationship between ICOR, CIV, and the expected returns by exploring the information innovation in them.

4.1. Patterns of ICOR and CIV

Following Herskovic et al. (2016), Figure 1 plots the trends of the yearly ICOR of the Chinese stock market, estimated by four different methods: POET, APCA-Threshold, Rm-Rf-Threshold, and FF3-Threshold. The thresholding parameter τ basically determines how many off-diagonal elements are to be dropped. According to the definition of ICOR, if τ is sufficiently high, all off-diagonal elements are set to zero, and ICOR equals zero. We list the ICOR results under five different τ in Figure 1. Surprisingly, when τ gets higher, ICOR grows rapidly, reaching the level of 0.3, and then remains relatively stable during the 25-year period. The estimation error of the high-dimensional correlation matrix (normally, the sample size equals 240, and the number of stocks is over 2000) is much lower, because in this case, when $\tau = 0.3$, most of the off-diagonal correlation elements have been deleted [4], and the estimation error can be greatly reduced.

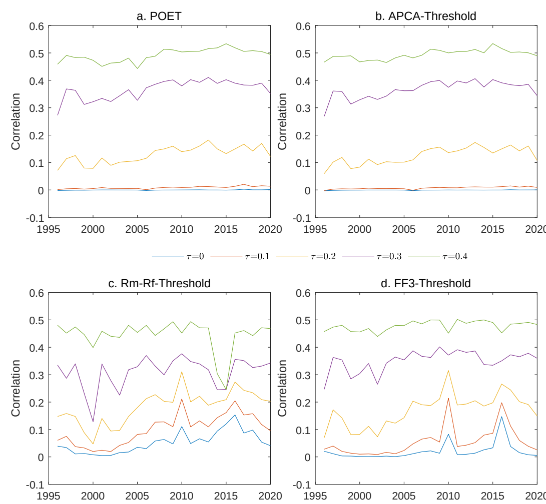

Figure 1: ICOR in the Chinese stock market under different thresholding parameters

Furthermore, the correlations between the idiosyncratic returns that ICOR represents are very difficult to eliminate or diversify. Figure 2 reports the process of eliminating PCs and the corresponding ICOR levels. It is clear that in most of the cases, ICORs do not drop to zero until over 100 PCs have been removed. This finding indicates that the existence of ICOR cannot be explained by any dominating statistical factors; otherwise, ICOR would quickly drop to zero after the first few PCs are removed.

We calculate CIV by averaging the yearly IVOL cross-sectionally, as in Her-

[4] The sparsity of residual correlation matrix drops from 100% to 30.9%, 5% and 1%, if τ increases from 0 to 0.1, 0.2 and 0.3, respectively.

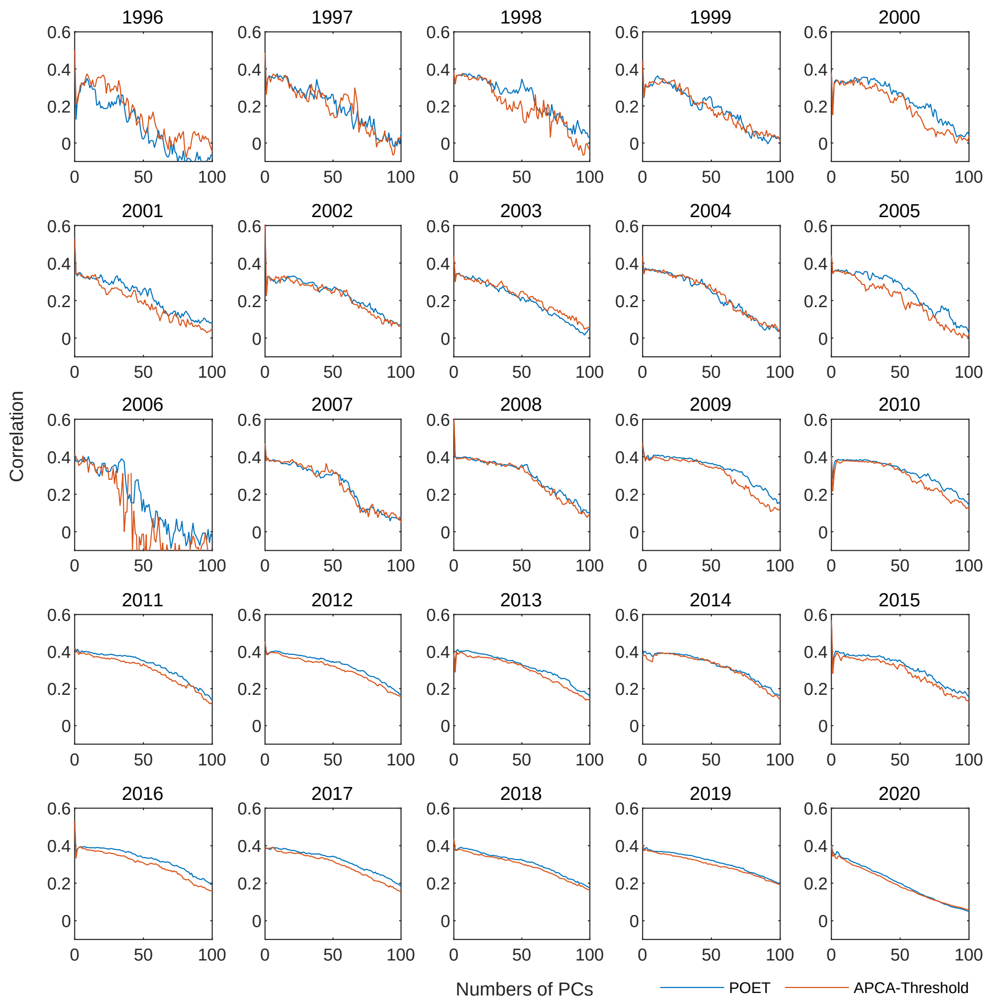

Figure 2: The process of eliminating principal components

skovic et al. (2016). Figure 3 shows the CIV patterns estimated by four different factor models. They are very similar in terms of their trends and levels. CIV becomes very high in 2007 and 2015, when huge stock market crashes occurred. Recall that ICOR is independent of the market crashes in the Chinese stock market in Figure 1, so it is completely different from the CIV trends. To show that our methodologies are consistent with Herskovic et al. (2016), we also include the CIV trends in the US stock market. We use the same factor models and thresholding methods to estimate CIV and ICOR in the US market, and we report the CIV results in Figure 4. As the figure shows, the CIV patterns calculated by all four methods are virtually identical, with CIV peaks during the recessions in 1932, 1992, 2000, and 2008. The trends of ICOR in the US stock markets under different thresholding parameters are shown in Figure 5. Theses ICOR trends are more uneven than those in China, and they have no apparent peaks during the recessions. In summary, ICOR exhibits completely different patterns from CIV. We discuss this difference later.

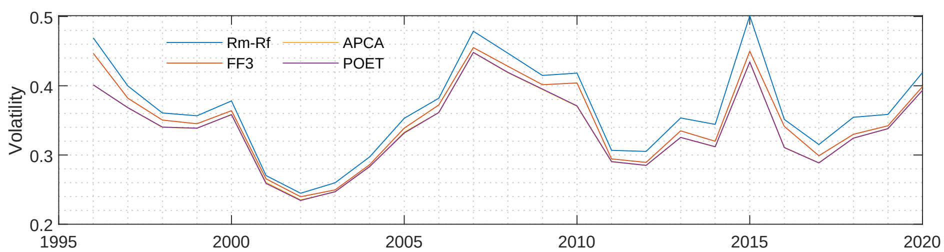

Figure 3: CIV estimated by four different methods in the Chinese stock market

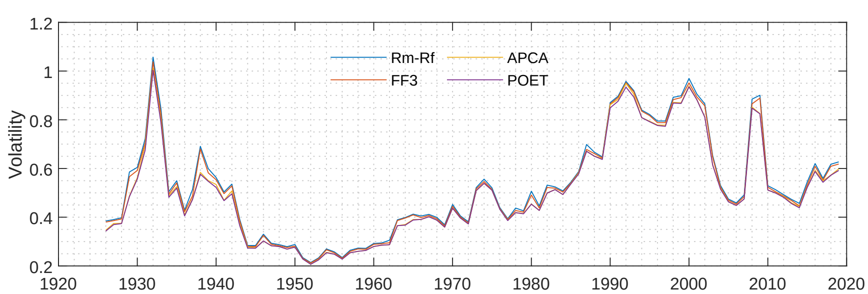

Figure 4: CIV estimated by four different methods in the US stock market

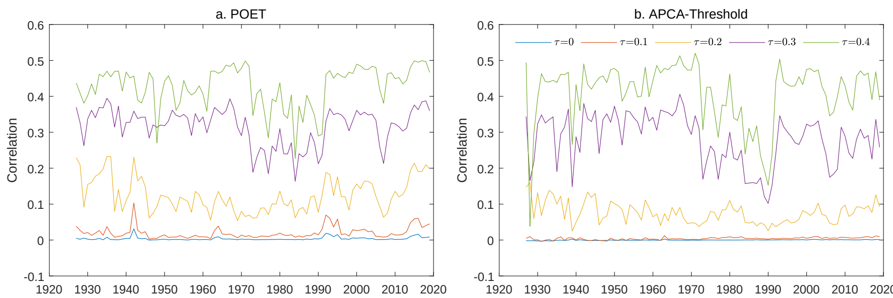

Figure 5: ICOR in the US stock market under different thresholding parameters

Herskovic et al. (2016) calculate the average pairwise idiosyncratic correlations of the US stock market, which corresponds to the case of $\tau = 0$ in Figure 5, and find that it is negligible. Our result aligns with theirs in that ICOR is close to zero when $\tau = 0$. However, when τ increases, ICOR increases. This is largely because the histogram of idiosyncratic correlations is positively skewed, and our thresholding process is equal to directly thresholding the correlation matrix (Fan et al., 2013)[5], where a correlation is removed if its absolute value is below τ. The histogram of the idiosyncratic correlation elements in the Chinese stock market is shown in Figure 6. The skewness is 0.496, and it can be seen from the log counts that the number of large positive correlations is greater than that of large negative correlations. Thus, ICOR increases as τ increases. To further understand why ICOR can be "falsely" close to zero, we must consider the idiosyncratic covariance.

Recall that the idiosyncratic correlation between stock i and stock j is $\text{Corr}(\varepsilon_{i,t}, \varepsilon_{j,t}) = \text{Cov}(\varepsilon_{i,t}, \varepsilon_{j,t})/(IVOL_{i,t} IVOL_{j,t})$, where $\text{Cov}(\varepsilon_{i,t}, \varepsilon_{j,t})$ is the idiosyncratic covariance. Adopting the thresholding method, we remove the elements in the idiosyncratic covariance matrix with absolute correlations lower than $\tau = 0.3$. We thus obtain the thresholded average idiosyncratic covariance (ICOV). The thresholding process plays an important role here. In Figure 7, we

[5]To see this, recall that in Equation 4, the indicator is $I(|\hat{v}_{ij}| \geq \tau_{ij})$, where $\tau_{ij} = \tau(\hat{v}_{ii}\hat{v}_{jj})^{1/2}$ and $\hat{v}_{ii}^{1/2}$ is the stock i's idiosyncratic standard deviation/volatility. Divided by $(\hat{v}_{ii}\hat{v}_{jj})^{1/2}$, the indicator becomes $I(|\text{Corr}(\varepsilon_{i,t}, \varepsilon_{j,t})| \geq \tau)$.

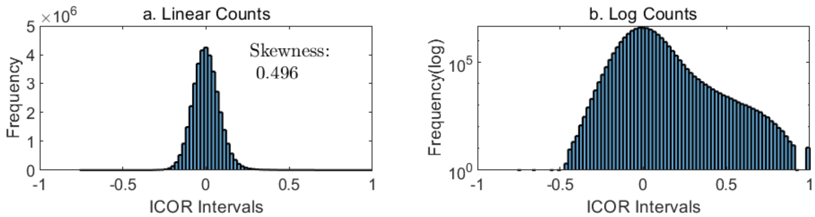

Figure 6: The histogram of ICOR in Chinese stock market

compare the ICOV and CIV of the Chinese and US stock markets. The ICOV is highly correlated with CIV, with correlations of 0.6324 for the Chinese stock market and 0.8082 for the US stock market. The average thresholded ICOV from 1996 to 2020 is 0.000055581, which is at least 100 times larger than the average value of ICOV without thresholding. Further, the correlation with CIV will drop to -0.4620 if ICOV is not thresholded. According to Herskovic et al. (2016) and Chen and Strebulaev (2018), the surge in CIV is caused by synchronized macroeconomic shocks. Logically, the macroeconomic shocks should also cause ICOV to change, which only occurs after we remove most of the small elements in the covariance matrix. As a consequence, ICOV becomes highly and positively correlated with CIV. The overall results lead us to conclude that these small covariances, which are most likely caused by estimation error, contaminate the idiosyncratic covariance and result in wrong statistical inference regarding ICOV and ICOR.

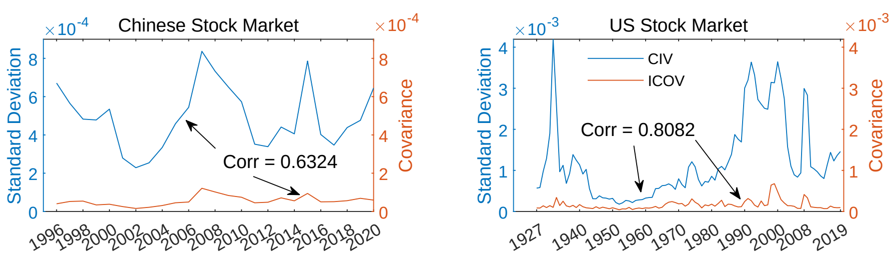

Figure 7: Comparison of ICOV and CIV in the Chinese and US market

However, as shown in Figure 1, ICOR does not increase with CIV and remains steady during market recessions in the Chinese market; however, ICOR decreases during market recessions in the US stock market, as Figure 5 shows. Regarding the Chinese market, when the denominator ($IVOL_{i,t}IVOL_{j,t}$) of the idiosyncratic correlation increases and the idiosyncratic correlation remains unchanged, the numerator $\text{Cov}(\varepsilon_{i,t}, \varepsilon_{j,t})$, which is the idiosyncratic covariance, will increase proportionally. The decline of ICOR in the US scenario is caused by the change in the denominator, which is greater than that of the numerator. The reasons for the difference in the two markets are twofold. First, the number of stocks in the Chinese stock market is much smaller than that of the US. In the datasets we adopted in this paper, the number of stocks is on average 1,481 for the Chinese stock market and 2,205 for the US stock market. Typically, to estimate ICOR for the US stock market, we need to estimate an extra 1 million or more idiosyncratic correlations than for the Chinese case. Hence, the estimation error may cause the difference in the ICOR between the two markets. Second, the idiosyncratic characteristics in the Chinese stock market are more persistent, and the variation of the idiosyncratic covariance cannot be easily absorbed by common factors and subsequently removed by the PCA process, thereby causing ICOR to remain steady. In the US, the stock market is more efficient. The variation of stock returns can be more effectively absorbed by fewer PCs or common factors. When the market is in recessions, the variation of the idiosyncratic covariances captured by PCA process is greater than that of idiosyncratic volatilities, which caused ICOR to decline. Overall, our results indicate that the variation of ICOR contains information from the idiosyncratic covariance and volatility. Next, we test its ability to price the expected stock returns.

4.2. Expected Stock Returns and ICOR

We construct ICOR beta sorted portfolios and investiagte whether there is a significant difference in the expected returns between different portfolios. First, we calculate the monthly ICOR changes and estimate the ICOR beta. To enable comparison with Herskovic et al. (2016), we also estimate market variance (MV) shocks as a control variable when estimating the ICOR beta.

The rolling window for beta estimation is 60 months, which is consistent with Herskovic et al. (2016). We construct 5 portfolios sorted by ICOR beta[6] or 25 portfolios double-sorted by ICOR beta and MV beta. Then, we calculate the equally weighted returns for each portfolio. Table 1 reports the portfolio performance sorted by ICOR beta and MV beta.

Panel A reports the portfolios sorted by ICOR beta. Column 5-1 shows that the long-short portfolio, which longs the portfolios with the highest ICOR beta and shorts the portfolios with the lowest ICOR beta, has a positive return of 6.17% at the 5% significance level. The long-short portfolio also has a positive alpha return when we control for Fama-French factors. Panel B reports the one-way sorted portfolios on the ICOR beta controlling for the MV beta, and the results are similar to those in Panel A. The returns of the portfolios bi-sorted by the ICOR beta and MV beta in Panel C show that with the exception of the smallest MV beta group, long-short portfolios of ICOR beta have significant positive returns 5.6%, 6.36%, 11.14% and 8.69%. However, the long-short portfolio returns of the MV beta do not show a clear significant pattern. The first portfolio return equals -2.87% and the second, third and fourth returns do not have sufficiently large t-values.

[6] This means that we first sort stocks by the numerical values of their ICOR beta and then divide these stocks into different groups to obtain ICOR beta sorted portfolios.

Table 1: **Portfolio performance sorted by ICOR and MV betas in the Chinese stock market**

		ICOR beta						
		1(small)	2	3	4	5(big)	5-1	t-value
Panel A: One way sorted portfolios on ICOR beta								
$E(R) - r_f$		9.5150	12.9109	13.6786	14.8148	15.6853	6.1703	2.2807***
α_{CAPM}		1.1078	4.4663	5.1369	6.3820	7.4285	6.3207	-
α_{FF3}		-1.0635	0.6659	0.5481	1.0761	1.5937	2.6572	-
Panel B: One way sorted portfolios on ICOR beta controlling for MV Beta								
$E(R) - r_f$		9.4250	12.6999	13.8382	15.1783	15.4634	6.0384	2.2444***
α_{CAPM}		0.9925	4.2116	5.2779	6.7659	7.2737	6.2812	-
α_{FF3}		-1.2899	0.4302	0.7255	1.3813	1.5731	2.8630	-
Panel C: Bi-sorted portfolios on ICOR beta and MV beta								
MV Beta	1(small)	10.0993	12.0399	11.7865	12.2675	10.5821	0.4828	0.1481
	2	9.2289	13.2296	13.8820	14.2768	14.8300	5.6011	1.9128***
	3	11.2564	12.1849	12.3445	17.4358	17.6136	6.3572	1.9541***
	4	9.5123	14.1272	16.2010	16.5028	20.6506	11.1383	3.0349***
	5(big)	7.2201	13.1785	14.1470	13.3441	15.9138	8.6937	2.3231***
	5-1	-2.8792	1.1386	2.3605	1.0766	5.3316	-	-
	T-value	-0.9041	0.3939	0.9381	0.3829	1.8076**	-	-

Notes: The portfolio returns are presented as annual percentages. α is the excess portfolio return under the CAPM and FF3 models. Column 5-1 reports the portfolio that shorts the smallest ICOR beta portfolio and longs the highest ICOR beta portfolio. *** indicates significane at 1% level.

We further report the portfolio pricing result of the CIV beta for the Chinese stock market[7] in Table 2. Panel A shows that the long-short portfolio has negative returns of -7.59% at the 1% significance level. The long-short portfolio also has a negative alpha return when we control for Fama-French factors. Panel B shows the same result as Panel A. Panel C shows that the long-short portfolios of the CIV beta have significant negative returns under all of the MV beta groups. Similar to Table 1, the results for the MV beta sorted portfolios have

[7]The portfolio result for the CIV beta for the US stock market is the same as in Herskovic et al. (2016). We omit it here.

no clear patterns. Clearly, our results in Table 2 align with Herskovic et al. (2016), as stocks with more negative sensitivity to positive CIV innovation will earn higher risk premiums.The reason is that, CIV surges always accompany with market recessions, and if the stock return drops at the same time, it is clearly riskier, and investors will demand higher expected return to hold onto them.

Table 2: **Portfolio performance sorted by CIV and MV betas in Chinese stock market**

		CIV Beta						
		1(small)	2	3	4	5(big)	5-1	t(5-1)
Panel A: One way sorted portfolio by CIV beta								
$E(R) - r_f$		15.8164	15.4014	13.9122	13.2545	8.2215	-7.5949	-2.8639
α_{CAPM}		7.4837	6.7672	5.3935	4.7613	0.1162	-7.3675	-
α_{FF3}		1.8512	1.2605	0.8853	1.0242	-2.1991	-4.0503	-
Panel B: One way sorted portfolios by CIV beta controlling for MV Beta								
$E(R) - r_f$		15.8505	15.3952	13.9621	13.0924	8.3043	-7.5462	-2.8243
α_{CAPM}		7.5897	6.8686	5.4211	4.4920	0.1497	-7.4400	-
α_{FF3}		1.9643	1.3640	0.7466	0.8823	-2.1378	-4.1021	-
Panel C: Bi-sorted portfolios by CIV beta and MV beta								
MV Beta	1(small)	17.0288	13.0804	12.0412	12.8713	7.5436	-9.4852	-1.9003
	2	15.2983	15.2076	14.7634	12.6230	7.3758	-7.9225	-2.5161
	3	15.4725	13.9529	14.4729	14.0133	11.2166	-4.2559	-1.3324
	4	18.1128	17.9110	15.0206	14.2177	10.1959	-7.9169	-2.4657
	5(big)	13.6104	16.6821	13.2250	11.7731	6.7483	-6.8621	-2.0424
	5-1	-3.4184	3.6017	1.1838	-1.0982	-0.7953	-	-
	t(5-1)	-1.1085	1.1742	0.4601	-0.4377	-0.1772	-	-

Notes: The portfolio returns are presented as annual percentages. α is the excess portfolio return under the CAPM and FF3 models. Column 5-1 reports the portfolio that shorts the smallest CIV beta portfolio and longs the highest CIV beta portfolio.

The portfolio results for the US stock market are reported in Table 3. As shown in column 5-1 of Panel A, the long-short portfolio has a positive return of 2.45%. It earns a positive alpha return when we control for CAPM and the Fama-French three factors. Panel B reports the portfolios sorted by ICOR beta when we control for MV beta. The long-short portfolio has a positive return of 1.95%. Panel C shows that the long-short portfolios of ICOR beta have significant positive returns of 3.32% and 3.10% on the MV beta, as rows 2 and 4 show. These results are not as strong as those for the Chinese stock market, partly because of the large estimation error contained in ICOR shock in US stock market, which may contribute to the inconsistent results between the two markets.

Table 3: **Portfolio performance sorted by ICOR and MV betas in the US stock market**

		CIV Beta						
		1(small)	2	3	4	5(big)	5-1	t(5-1)
Panel A: One way sorted portfolio by CIV beta								
$E(R) - r_f$		12.1373	12.4530	13.0673	13.9094	14.5854	2.4482	1.3860
α_{CAPM}		3.5053	3.3852	3.0648	2.5831	1.6105	-1.8948	-
α_{FF3}		2.2433	1.8808	1.0446	-0.1553	-1.8256	-4.0689	-
Panel B: One way sorted portfolios by CIV beta controlling for MV Beta								
$E(R) - r_f$		12.7413	12.2859	13.0854	13.3399	14.6949	1.9536	1.3116
α_{CAPM}		2.9425	3.0103	3.2813	2.7155	2.2008	-0.7417	-
α_{FF3}		1.4060	1.3985	1.3946	0.2717	-1.2793	-2.6853	-
Panel C: Bi-sorted portfolios by CIV beta and MV beta								
MV Beta	1(small)	12.4202	13.4750	12.7108	17.1308	13.2684	0.8482	0.0940
	2	12.8830	11.6200	13.0131	14.6560	16.2064	3.3234	1.4271
	3	14.2372	12.4422	13.1893	13.3950	14.6037	0.3665	0.1865
	4	12.4445	13.7283	13.5877	14.0557	15.5440	3.0995	1.8319
	5(big)	14.9847	13.9785	13.5840	13.0965	13.5933	-1.3914	-0.9895
	5-1	2.5645	0.5035	0.8732	-4.0343	0.3249	-	-
	t(5-1)	0.9101	0.2236	0.4511	-1.5585	0.2169	-	-

Notes: The portfolio returns are presented as annual percentages. α is the excess portfolio return under the CAPM and FF3 models. Column 5-1 reports the portfolio that shorts the smallest ICOR beta portfolio and longs the highest ICOR beta portfolio.

We also conduct Fama-Macbeth (Fama and MacBeth, 1973) regression analyses to price the portfolio returns in the Chinese stock market. The results are reported in Table 4. The risk premiums of the ICOR beta are positive and highly significant, with t-values of 5.6054 and 5.4502. The risk premiums of the CIV beta are negative and highly significant, with t-values of -4.1763 and -3.165, which are consistent with the result for the US stock market reported in Herskovic et al. (2016). The Fama-Macbeth results for the US stock market in Table 5 show that the risk premium of the ICOR beta and CIV beta are significantly positive and negative, respectively, indicating that ICOR is also priced in the US stock market.

Table 4: **Fama-Macbeth analysis results for the Chinese stock market**

	25 Portfolios Bi-sorted by ICOR/MV Betas			25 Portfolios Bi-sorted by CIV/MV Betas		
	(1)	(2)	(3)	(1)	(2)	(3)
Intercept	0.0028	0.0012	0.0026	0.0047	0.0018	0.0068
	(0.3619)	(0.2110)	(0.4143)	(0.5565)	(0.2413)	(0.7947)
$r_m - r_f$	0.0051	0.0037	0.0023	0.0032	0.0043	-0.0013
	(0.6861)	(0.6962)	(0.3833)	(0.3915)	(0.6328)	-(0.1541)
ICOR	-	**0.5837**	**0.5729**	-	-	-
	-	(5.6054)	(5.4502)	-	-	-
CIV	-	-	-	-	**-0.0797**	**-0.0709**
	-	-	-	-	(-4.1763)	(-3.1650)
MV	-	-	-0.6492	-	-	-2.5976
	-	-	(-0.4585)	-	-	-(1.1296)
r^2	0.0210	0.6031	0.6066	0.0065	0.4411	0.4905

Notes: The t-values are reported in parentheses. The Fama-Macbeth regression involves two stages. In the first stage, we regress the monthly excess return of the bi-sorted portfolios on factors to estimate their betas. In the second stage, we regress the average excess portfolio returns on their betas cross-sectionally to estimate the risk premium. ICOR, CIV, and MV are risk premium estimates. The 60-month rolling estimation period is from 2000/12 to 2020/09.

As shown in Table 1, stocks that have more positive exposure to ICOR change earn a higher risk premia, which is opposite to the CIV change, as shown in Table 2. The reason for the contrasting result is that the change of idiosyncratic covariance (numerator) is smaller than that of idiosyncratic volatility (denominator), which causes ICOR to drop and the ICOR shock to become negative. However, from Figure 1, we know that the yearly ICOR in the Chinese stock market is very stable. Thus, the change of ICOR is unable to price the expected returns, because, to ensure sufficient samples, we adopt the monthly ICOR shock in asset pricing tests. The larger estimation error contained in the monthly ICOR shock will further lower the change of idiosyncratic covariance, which causes ICOR to drop and the ICOR shocks to become negative, resulting in the result opposite to that of CIV change. This explanation can be applicable to both markets.

Table 5: **Fama-Macbeth analysis results for the US stock market**

	25 Portfolios Bi-sorted by ICOR/MV Betas			25 Portfolios Bi-sorted by CIV/MV Betas		
	(1)	(2)	(3)	(1)	(2)	(3)
Intercept	0.0062	0.0086	0.0085	-0.0028	0.0002	-0.0020
	(10.1392)	(5.8029)	(5.1039)	(-1.4217)	(0.1005)	(-0.7164)
$r_m - r_f$	0.0038	0.0013	0.0013	0.0111	0.0078	0.0104
	(7.8576)	(0.8868)	(0.8338)	(6.6670)	(3.9045)	(3.7038)
ICOR	-	**0.0146**	**0.0139**	-	-	-
	-	(2.5278)	(1.9429)	-	-	-
CIV	-	-	-	-	**-0.0781**	**-0.0752**
	-	-	-	-	(-2.7805)	(-4.1561)
MV	-	-	0.0012	-	-	0.0131
	-	-	(0.5294)	-	-	(1.6922)
r^2	0.7931	0.8072	0.8084	0.7203	0.8062	0.8303

Notes: The t-values are reported in parentheses. The Fama-Macbeth regression involves two stages. In the first stage, we regress the monthly excess return of the bi-sorted portfolios on factors to estimate their betas. In the second stage, we regress the average excess portfolio returns on their betas cross-sectionally to estimate the risk premium. ICOR, CIV, and MV are risk premium estimates. The 60-month rolling estimation period is from 2000/12 to 2020/09.

5. Conclusions

We investigate the average idiosyncratic correlation (ICOR) and its pricing ability for an emerging market and a developed market: the Chinese stock market and the US stock market, respectively. Our analyses provide empirical evidence that ICOR plays an important role in both markets, as CIV does in the

US stock market (Herskovic et al., 2016). On the basis of daily stock returns, we adopt stocks in the Chinese A share market spanning from 1995 to 2020 and US from 1926 to 2019. To deal with the high-dimensional estimation problem, we adopt POET, a high-dimensional PCA method combined with the thresholding method, to better estimate ICOR. Our results demonstrate that the thresholding method may recover the trend of ICOV, with a high correlation to that of CIV. Furthermore, we find that the stocks' exposure to CIV shocks is negatively priced and their exposure to ICOR change is positively priced. These opposite results are due to large estimation error contained in the monthly ICOR. Our results are consistent with those ofHerskovic et al. (2016) for the Chinese and US stock markets, and shed new light not only on idiosyncratic volatilities, but also on idiosyncratic covariances and correlations, providing evidence that ICOR contains important pricing information that largely depends on the accurate estimation of ICOV.

References

Ang, A., Hodrick, R. J., Xing, Y., and Zhang, X. (2006). The cross-section of volatility and expected returns. *The Journal of Finance*, 61(1):259–299.

Avramov, D., Chordia, T., Jostova, G., and Philipov, A. (2013). Anomalies and financial distress. *Journal of Financial Economics*, 108(1):139–159.

Boyer, B., Mitton, T., and Vorkink, K. (2009). Expected idiosyncratic skewness. *The Review of Financial Studies*, 23(1):169–202.

Chen, Z. and Strebulaev, I. A. (2018). Macroeconomic risk and idiosyncratic risk-taking. *The Review of Financial Studies*, 32(3):1148–1187.

Connor, G. and Korajczyk, R. A. (1986). Performance measurement with the arbitrage pricing theory: A new framework for analysis. *Journal of Financial Economics*, 15(3):373–394.

Connor, G. and Korajczyk, R. A. (1988). Risk and return in an equilibrium apt: Application of a new test methodology. *Journal of Financial Economics*, 21(2):255–289.

Duarte, J., Kamara, A., Siegel, S., and Sun, C. (2014). The systematic risk of idiosyncratic volatility. *Working Paper*.

Fama, E. F. and French, K. R. (1992). The cross-section of expected stock returns. *The Journal of Finance*, 47(2):427–465.

Fama, E. F. and MacBeth, J. D. (1973). Risk, return, and equilibrium: Empirical tests. *Journal of Political Economy*, 81(3):607–636.

Fan, J., Liao, Y., and Mincheva, M. (2013). Large covariance estimation by thresholding principal orthogonal complements. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 75(4):603–680.

Hasan, M. M. and Habib, A. (2017). Firm life cycle and idiosyncratic volatility. *International Review of Financial Analysis*, 50:164–175.

Herskovic, B., Kelly, B., Lustig, H., and Van Nieuwerburgh, S. (2016). The common factor in idiosyncratic volatility: Quantitative asset pricing implications. *Journal of Financial Economics*, 119(2):249–283.

Merton, R. C. (1987). A simple model of capital market equilibrium with incomplete information. *The Journal of Finance*, 42(3):483–510.

Tsay, R. (2005). *Analysis of financial time series*. John Wiley: New York.