Three Essays in Factor Analysis of Asset Pricing

Author: Wenzhi Wang

Boston College Electronic Thesis or Dissertation, 2018

Finance Department

The Carroll Graduate School of Management

Boston College

Three Essays in Factor Analysis of Asset Pricing

a dissertation

by

Wenzhi Wang

submitted in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy in Finance

April 2018

Three Essays in Factor Analysis of Asset Pricing

Wenzhi Wang

Abstract

My dissertation is comprised of three chapters. The first chapter is motivated by many low-frequency sources of systemic risk in the economy. We propose a two-stage learning procedure to construct a high-frequency (i.e., daily) systemic risk factor from a cross-section of low-frequency (i.e., monthly) risk sources. In the first stage, we use a Kalman-Filter approach to synthesize the information about systemic risk contained in 19 different proxies for systemic risk. The low frequency (i.e., monthly) Bayesian factor can predict the cross-section of stock returns out of sample. In particular, a strategy that goes long the quintile portfolio with the highest exposure to the Bayesian factor and short the quintile portfolio with the lowest exposure to the Bayesian factor yields a Fama–French–Carhart alpha of 1.7% per month (20.4% annualized). The second stage is to convert this low frequency Bayesian factor into a high-frequency factor. We use textual analysis Word2Vec that reads the headlines and abstracts of all daily articles from the business section of the New York Times from 1980 to 2016 to collect distributional information on a per word basis and store it in high-dimensional vectors. These vectors are then used in a LASSO model to predict the Bayesian factor. The result is a series of coefficients that can then be used to produce a high-frequency estimate of the Bayesian factor of systemic risk. This high-frequency indicator is validated in several ways including by showing how well it captures the 2008 crisis. We also find that the high frequency factor is priced in the cross-section of stock returns and able to predict large swings in the VIX using a quantile regression approach, which sheds some light on the puzzling relation between the macro-economy and stock market volatility.

The second chapter of my dissertation provides a basic quantitative description of a compendium of macro economic variables based on their ability to predict bond returns and stock returns . We use three methods(asymptotic PCA, LASSO and Support Vector Machine) to construct factors out of 133 monthly time series of economic activity spanning a period from 1996:1 to 2015:12 and classify these factors into two groups: bond demand factors and bond supply factors. In PCA regression, we find both demand factors and supply factors are unspanned by bond yields and have stronger predictability power for future bond excess returns than CP factors. This predictability finding is confirmed and enhanced by machine learning technique LASSO and Support Vector Machine. More interestingly, LASSO can be used to identify 15 most important economic variables and give direct economic explanations of predictors for bond returns. Regarding to stock predictability, we find both demand and supply PC factors are priced by the cross-section of stock returns. In particular, portfolios with highest exposure to aggregate supply factor outperform portfolios with

lowest exposure to aggregate supply factor 1.8% per month while portfolios with lowest exposure to aggregate demand factor outperform portfolios with highest exposure to aggregate demand factor 2.1% per month. The finding is consistent with "fly to safety" explanation. Furthermore, variance decomposition from VAR shows that demand factors are much more important than supply factors in explaining asset returns. Finally, we incorporate demand factors and supply factors into macro-finance affine term structure (MTSMs) to estimate market price of risk of factors and find that demand factors affect level risk and supply factors affect slope risk. Moreover, MTSMs enable us to decompose bond yields into expectation component and yield risk premium component and we find MTSMs without macro factors under-estimate yield risk premium.

The third chapter,coauthored with Dmitriy Muravyev and Aurelio Vasquez, is motived from the fact that a typical stock has hundreds of listed options. We use principal component analysis (PCA) to preserve their rich information content while reducing dimensionality. Applying PCA to implied volatility surfaces across all US stocks, we find that the first five components capture most of the variation. The aggregate PC factor that combines only the first three components predicts future stock returns up to six months with a monthly alpha of about 1%; results are similar out-of-sample. In joint regressions, the aggregate PC factor drives out all of the popular option-based predictors of stock returns. Perhaps, the aggregate factor better aggregates option price information. However, shorting costs in the underlying drive out the aggregate factor's predictive ability. This result is consistent with the hypothesis that option prices predict future stock returns primarily because they reflect short sale constraints.

"Thus, when Heaven is about to confer a great office on any man, it first exercises his mind with suffering, and his sinews and bones with toil. It exposes his body to hunger, and subjects him to extreme poverty. It confounds his undertakings. By all these methods it stimulates his minds, hardens his nature, and supplies his incompentencies.(Translated by James Legge in The Works of Mencius)"

Mencius (Mengzi. Gaozi. Part II)

Acknowledgments to my advisors Robert Taggart, Rui Albuquerque, Dmitriy Muravyev, Thomas J. Chemmanur, Alan Marcus, Natesh Pillai, Ronnie Sadka, Jeffrey Pontiff, and my coauthors and colleagues Aurelio Vasquez, Ding Ding, Hugo Benedetti, Yu Wang, Siyi Shen, Xuechuan Ni, Jun Yang, Yao Shen, and seminar participants at Boston College Brown Bag Seminar for useful comments.

Contents

I Chapter 1: Two-Stage Learning of a Systemic Risk Factor **8**

A Data and Methodology . . . 13

B The Bayesian Systemic Risk Factor . . . 17

C High Frequency Textual Systemic Risk Factor . . . 23

D Asset Pricing Implication of Bayesian Factor . . . 32

E Property of High Frequency Daily Textual Factor . . . 43

F Conclusion . . . 46

II Chapter 2: Combining Bond Demand and Supply Factors **47**

A Related literature . . . 50

B PCA Analysis of Economic Variables . . . 51

C Machine Learning of Important Economic Variables . . . 62

D Stock Portfolio Sort Results based on Exposure to Demand and Supply Factors . . . 67

E VAR and Impulse Response Function . . . 71

F Market Price of Risk on Bond Factors . . . 76

G Conclusion . . . 87

III Chapter 3: Making Better Use of Options to Predict Stock Returns **89**

A Data and Methodology . . . 93

B Principal Components and Future Stock Returns . . . 100

C Robustness . . . 108

D Conclusion . . . 110

A Appendix: Two-Stage Learning of a Systemic Risk Factor **119**

B Appendix: Combining Bond Demand and Supply Factors **130**

C Appendix: Making Better Use of Option Prices to Predict Stock Returns **130**

I. Chapter 1: Two-Stage Learning of a Systemic Risk Factor

The ability of systemic risk to trigger sharp financial downturns has made this risk a focal point of research. In the aftermath of the 2007-2009 financial crisis, many systemic risk measures have been proposed at the low-frequency (monthly) level. In this paper, we have two complementary objectives for establishing an understanding of systemic risk measures.

Our first goal is to condense a cross-section of low-frequency systemic risk measures into one low-frequency systemic risk factor. We examine 19 previously proposed monthly systemic risk measures and use Particle-MCMC to combine them to one latent monthly systemic risk factor, which we have named the Bayesian factor. We find this monthly Bayesian factor can be projected onto the monthly textual information from the New York Times based on LASSO. A simple Granger Causality test shows that the monthly textual information from the New York Times can Granger cause the monthly Bayesian factor while the monthly Bayesian factor can not Granger cause the monthly textual information. Based on this causal relationship, our second goal is to feed the LASSO regression with the high-frequency (daily) textual information from the New York Times and use the fitted value for constructing a high-frequency (daily) textual factor. We find that both the low-frequency Bayesian factor and the high-frequency textual factor are significantly priced in a cross-section of stock returns. In particular, a strategy that goes long the quintile portfolio with the highest exposure to monthly Bayesian factor and short the quintile portfolio with the lowest exposure to monthly Bayesian factor yields a Fama–French–Carhart alpha of 1.7% per month (20.4% annualized). Additionally, We find that our monthly Bayesian factor is linked to distress related firm characteristics and more distressed firms are more sensitive to our Bayesian factor. This sheds some light on why distress risk could be a severe risk for investors. Incorporating our monthly Bayesian factor can help explain 36% of the distress anomaly. Lastly, our daily textual systemic risk factor is very informative in predicting future macro financial conditions, for example, the time-varying VIX index shocks and the TED-spread shocks.

The estimation of the price of risk on both factors consistently identifies a negative market price of systemic risk. As financial crash risk is accumulating, intermediary investors marginal value of wealth is going up. Assets that pay off poorly when crash risk peaks are therefore risky and must offer high returns. Equivalently, financial systemic risk is countercyclical and the cross-sectional price of systemic risk should be negative. This significantly negative systemic risk premium is also consistent with the intertemporal capital asset pricing model (ICAPM) ofMerton (1973). An increase in systemic risk and economic downside risk shock reduces future investment and consumption opportunities. To hedge against such an unfavorable shift, investors prefer to hold stocks whose returns increase in times of high systemic risk. When systemic risk rises, investors suffer through a reduction in optimal consumption and future investment opportunities. They are able to compensate for this loss by holding stocks that positively correlate with systemic risk. This intertemporal hedging demand argument implies that investors are willing to hold stocks with higher covariance with systemic risk, and they pay higher prices and accept lower returns for stocks with higher systemic risk betas.

Our study contributes to several strands of the asset pricing literature. First, it is related to how to aggregate individual measures for systemic risk. Giglio, Kelly, and Pruitt (2016) (henceforth GKP) extract a latent factor out of the same cross section of systemic risk measures and find that their factor can predict lower quantiles of future macroeconomic shocks instead of the central tendency of those shocks. There are two key differences between GKP and our paper. First, the two papers' motivations and focus are very different. They focus their analysis on the interactions between systemic risk and the macro-economy to highlight which measures are valuable as an input to regulatory or policy choices. Therefore, it is not surprising to find out that their measure can not explain the cross-section of equity excess returns. Instead, our focus lies on constructing a measure for the pricing kernel and examining the asset pricing implications of this measure. Empirical results in this paper suggest that our factors are better proxies for the pricing kernel than their Partial Quantile Regression (henceforth PQR) Factor. Second, the two papers employ very different methodologies. Motivated by exploring the relationship between systemic risk and the distribution of macro-economic shocks, GKP uses the quantile regression to find that systemic risk can predict a downside quantile of industry production innovations. Moreover, their method of dimension reduction, Partial Quantile Regression, condenses the cross section of predictors according to each predictors quantile covariation with the forecast target,choosing a linear combination of predictors that is a consistent quantile forecast. So, their methods rely highly on quantile regression and have already utilized the information of the forecast variable which consists of industry production shocks. In contrast, our Bayesian method is casting the dynamics of cross-section of predictors into a State-Space Model and it uses Particle-MCMC to extract a Bayesian factor. Our high-frequency textual Bayesian factor is constructed by projecting the Bayesian factor onto textual information and using the fitted value as our second financial systemic risk factor. So, we condense the cross section according to covariance within the predictors, disregarding how closely each predictor relates to the target, which are asset prices in our case.

Secondly, this paper advances the empirical ICAPM literature. Over 40 years of macro-finance research, there are still very few of empirical studies successfully linking macroeconomic risk exposure to the cross-section of stock returns. The most recent paper focusing on this issue is by Bali, Brown, Peng, and Tang (2016). They use Ludvigson and Ng's macro uncertainty measure to proxy for the ICAPM state variable and find this uncertainty measure is priced in a cross-section of stocks over the sample of 1977 and 2014. Our paper is very different from their paper because we are focusing on a priced crash risk or a downside tail risk while they are paying attention to a priced uncertainty measure. More importantly, their uncertainty measure is not significantly priced over the sample from 1995 to 2014, but our financial crash risk is significantly priced during this period. It is important to use the sample after 1995 to test the asset pricing implications because several important priced factors are published before 1995 and their predictability power for stock returns shrinks a lot after their publication,McLean and Pontiff (2016). Of course, this paper is also related to the intermediary asset pricing literature. Brunnermeier and Pedersen (2009) show how intermediary funding liquidity determines the pricing kernel when investors face funding constraints. Specifically, funding liquidity, ϕ_1, is also the Lagrange multiplier on the time-one funding constraint. Risk-neutral investors subject to a funding constraint maximize $E_0[\phi_1 W_1]$, where W_1 is investors' wealth at time 1. Equivalently, SDF is simply $\frac{\phi_1}{E_0[\phi_1]}$. There are several funding illiquidity measures constructed along this line, for example, Fontaine, Garcia, and Gungor (2016), Chen and Lu (2016), and Lee (2013). Unfortunately, all of these funding illiquidity risk factors are not significantly priced in our sample period. In contrast, our financial crash risk is significantly priced, which gives direct empirical support for our strategy of combining both macro information and financial intermediary information.

Thirdly, the construction of our high-frequency textual risk factor is motivated by an explosion of empirical economics research using text as data. An ever increasing share of human interaction, communication, and culture is recorded as digital text and the information encoded in text is a rich complement to the more structured kinds of data traditionally used in research. For example, major newspapers usually contain a large collection of high quality articles analyzing, summarizing, and even predicting people's thinking and belief on current economic issues. Manela and Moreira (2017) take a text regression approach to construct an index of news-implied market volatility based on text from the Wall Street Journal from 1890-2009. They apply support vector regression, which uses a penalized least squares objective to identify a small subset of words whose frequencies are most useful for matching patterns of turbulence in financial markets. They find that high levels of news-implied volatility forecast high future stock market returns. However, the training subsample used to estimate the dependency between news data and implied volatility is from 1996 to 2009 while the test subsample used for out-of-sample tests of model fit is from 1986 to 1995. A potential concern is that because the training sample period is chronologically after the predict subsample, they are using new information, unavailable to those who lived during the predict subsample, to predict future returns. Instead, our training sample is from 1990 to 2003 and the test sample is from 2004 to 2011, which is a regular sample splitting design. Furthermore, we adopt a cutting-edge Natural Language Processing (NLP) technology called Word Embedding to construct our textual systemic risk factor out of Business Section articles in New York Times from 1980 to 2016. This method of utilizing textual data brings about a high-frequency textual factor which shows more asset pricing power in a cross-section of portfolio. To my best knowledge, we are the first to construct such a high-frequency textual systemic risk measure. Most the of well-studied measures are constructed at the monthly level either due to lack of high frequency data, such as book leverage of big financial institutions, or due to a long-term rolling-window based estimation procedure, such as Marginal Expected Shortfall or DCI. It is well-known that low frequency data would give rise to a noisy estimation of risk factor exposure β in the first stage of a Fama-Macbeth regression, and therefore constructing high frequency data could possibly mitigate this measurement error concern. Lastly, we are interested in explaining latent factors such as a Bayesian latent factor and textual information can help us to achieve this goal to some degree.

Finally, we contribute to the literature on measuring the state of the economy in a time-series setting, commonly referred to as "now-casting." One of the important papers in this field is Beber, Brandt, and Luisi (2015). They use PCA to extract daily principal components from economic news release associated with a specific information type. They think of each news series as a continuously evolving time series observed only once per month and simply forward-fill the lat observed release until the next announcement. Put differently, there is no update of information between two announcements and this is the limitation of using macro news announcement data to now-cast high-frequency macro fundamentals. Instead of using numerical news data, we utilize the real high-frequency data, the textual data, to fill in the gap between two announcements and solve the missing variable problem. Our method would give rise to a high-frequency time-varying series instead of a high-frequency step-function series.

The remainder of the paper is organized as follows. Section I describes the data and empirical strategy used to collect individual financial systemic risk measures. Section II describes the first-stage Bayesian aggregation of individual systemic risk variables. Section III constructs the second-stage high-frequency textual risk factor out of monthly numeric systemic risk measures in three steps. Section IV conducts a number of asset pricing tests in the cross-section of stock returns and shows the predictive power of the systemic risk factor. Section V investigates the properties of our high-frequency textual systemic risk factor. Section VI concludes.

A. *Data and Methodology*

A.1. Data

This section outlines our construction of financial systemic risk measures and provides a brief summary of comovement among the measures. Individual measures are based on data for financial institutions identified by 2-digit SIC codes 60 through 67 (finance, insurance and real estate). Equity returns for US financial institutions are obtained from CRSP and book data are obtained from Compustat. We are interested in capturing primary intermediary systemic risk, and thus construct our measures using data for the 20 largest financial institutions in each period, including the likes of Goldman Sachs, JP Morgan, and Deutsche Bank. Whenever the financial systemic risk measure is aggregated from institution-level measures, we compute the measure for each of the 20 largest institutions in each period and take an equal weighted average. We closely follow the Giglio, Kelly, and Pruitt (2016) procedure to categorize and collect 19 financial systemic risk measures, which also covers the major part of 30 risk measures discussed by Bisias, Flood, Lo, and Valavanis (2012). Below we provide a brief overview of the measures that we build grouped by their defining features. A more detailed list of definitions can be found in Table 1.

The first set of measures are institution-specific measures and include CoVaR and ΔCoVaR from Adrian and Brunnermeier (2016), marginal expected shortfall (MES) from Acharya, Pedersen, Philippon, and Richardson (2017), and MES-BE, a version of marginal expected shortfall proposed by Brownlees and Engle (2012). The second set of measures is intended to quantify co-movement and contagion among intermediary equity returns. Kritzman, Li, Page, and Rigobon (2011) construct the Absorption Ratio to capture the fraction of the financial system variance explained by the first K=3 PCAs. The Dynamic Causality Index (DCI) from Billio, Getmansky, Lo, and Pelizzon (2012) counts the number of significant Granger-causal relationships among financial institutions equity returns, and the International Spillover Index from Diebold and Yilmaz (2009) measures co-movement in macroeconomic variables across different countries. The third set of measures is intended to capture the primary intermediary volatility. One simple measure is just the average equity volatility of the primary institutions. In addition, a turbulence variable following Kritzman and Li (2010) considers returns recent covariance relative to a longer-term covariance estimate. Another VaR measure from Allen, Bali, and Tang (2012) is derived by looking at the cross section of financial firms at any one point in time. Size concentration in the financial industry (the market capitalization Herfindal index) is also included to capture a potential instability in the sector. The fourth set of measures is designed to proxy for Liquidity and credit conditions in financial markets. These are Amihud (2002)'s illiquidity measure (AIM) aggregated across financial institutions, the TED spread (LIBOR minus the T-bill rate),the default spread (BAA bond yield minus AAA bond yield), the Gilchrist-Zakrajsek measure of the credit spread (GZ, proposed in Gilchrist and Zakrajšek (2012)) and the term spread (the slope of the Treasury yield curve). Of course, we include both book leverage and market leverage for the largest 20 financial institutions in our sample to serve as counterparts to the AEM leverage factor and the HKM capital ratio factor.

In Figure 8, We inspect the correlation among individual systemic risk measures. Most correlations are quite low. Only two groups of measures comove strongly. First, Catfin, turbulence, volatility, and the TED spread are relatively highly correlated. Secondly, CoVaR, $\Delta CoVaR$, MES, GZ, size concentration, and Absorption tend to comove. The other measures display low or even negative correlations with each other, suggesting that many measures capture different aspects of financial system distress or are subject to substantial noise. Only using one or two leverage factors to proxy for pricing kernel might not be good enough due to a complicated relationship among different distress proxies and potential noise within those proxies. So, we need dimension reduction techniques to help detect the relationship between the large collection of financial distress measures and asset prices, above and beyond the information in these potentially noise-ridden individual measures.

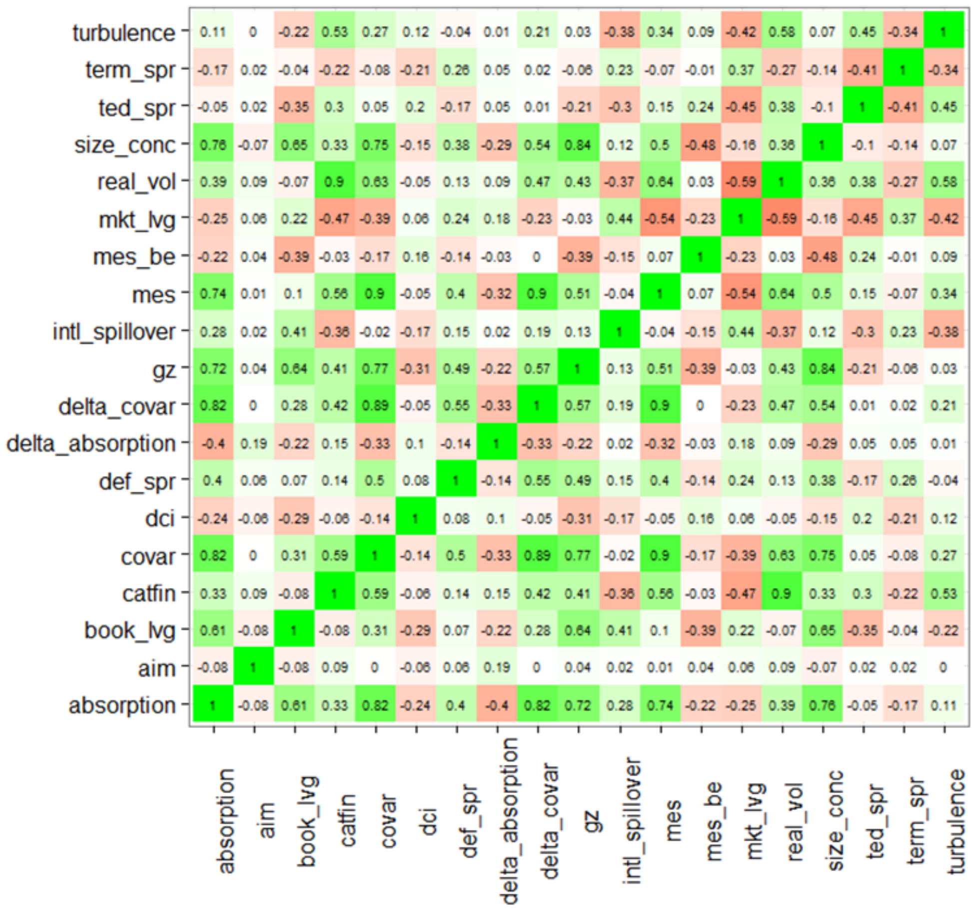


Figure 1. This figure plots the correlation matrix of all the individual financial systemic risk measures.

B. The Bayesian Systemic Risk Factor

Our Bayesian Factor model is similar to the dynamic factor framework used in Stock and Watson (1989) to estimate the coincident economic indicator. I put the dynamic factor model into a state space framework. For $i = 1, \cdots, N$, denote $x_{i,t}$ to be observation for unit i at time t. The measurement equation for the state space model is:

$$(1 - \psi_i L)X_{i,t} = (1 - \psi_i L)(\beta_{i,0} + \beta_{i,1}L + \beta_{i,2}L^2)g_t + \epsilon_{x,i,t} \tag{1}$$

or more compactly,

$$X^{\star}_{i,t} = \beta^{\star}_i g_t + \epsilon_{X,i,t}. \tag{2}$$

where g_t is the latent Bayesian Factor at time t, L is a lag operator, and $\epsilon_{X,i,t} \sim N(0, \sigma_{X,i,t})$ The transition equation is :

$$g_t = \psi_g g_{t-1} + \epsilon_{g,t}. \tag{3}$$

where $\epsilon_{g,t} \sim N(0, \sigma_{g,t})$.

We adopt a general method Particle-MCMC similar to Lindsten, Jordan, and Schön (2014) and Andrieu, Doucet, and Holenstein (2010) to estimate the latent factor g_T and the model parameters. The only difference here is that the dynamic factor model is not Markovian. Existing literatures suggest two ways to solve these non-Markov models. One way is to construct auxiliary variables and transform the non-Markov model into a Markovian State Space Model, such as Herbst and Schorfheide (2015). The other way to attack this problem is to derive the joint smoothing distribution directly and approximate them. In this model, because the observation equation only relies on four lags of the latent factor, I can derive the joint smoothing distribution analytically without any approximation and I can also derive the importance weights in sampling as follows:

$$\text{Smoothing Distribution} \equiv \frac{P_\theta\left((g^i_{1:\,t-1}, g^{'}_{t:\,T}), x_{1:\,T}\right)}{P_\theta\left(g^i_{1:\,t-1}, x_{1:\,t-1}\right)} = P_\theta\left(g^{'}_{t:\,T}, x_{t:\,T}|g^i_{1:\,t-1}, x_{1:\,t-1}\right) \tag{4}$$

$$= P_\theta\left(g^{'}_t, x_t, x_{t+1}, x_{t+2}|g^{'}_{t+1:\,T}, x_{t+3:\,T}, g^i_{1:\,t-1}, x_{1:\,t-1}\right) \times P_\theta\left(g^{'}_{t+1:\,T}, x_{t+3:\,T}|g^i_{1:\,t-1}, x_{1:\,t-1}\right) \tag{5}$$

$$\propto P_\theta(x_{t+2}|g^i_{t-1}, g^{'}_{t:\,t+2}) \times P_\theta(x_{t+1}|g^i_{t-2:\,t-1}, g^{'}_{t:\,t+1}) \times P_\theta(x_t|g^i_{t-3:\,t-1}, g^{'}_t) \times f_\theta(g^{'}_t|g^i_{t-1}) \tag{6}$$

$$\text{Importance Weights} \equiv w^i_t = \frac{P(g^i_{1:\,t}|x_{1:\,t})}{r_{\theta,t}(g^i_{1:\,t}|x_{1:\,t})} \tag{7}$$

$$\propto \frac{G(x_t|g^i_{t-3:\,t})f_\theta(g^i_t|g^i_{t-1})P(g^i_{1:\,t-1}|x_{1:\,t-1})}{r_{\theta,t}(g^i_t|g^i_{t-1}, x_t)r_{\theta,t-1}(g^i_{1:\,t-1}|x_{1:\,t-1})} \tag{8}$$

$$= G(x_t|g^i_{t-3:\,t}) \tag{9}$$

where $r_{\theta,t}(g^i_t|g^i_{t-1}, x_t) = f_\theta(g^i_t|g^i_{t-1})$

Here I summarize the procedure in Algorithm 1 and 2.

Figure 9 plots the time series of the Bayesian Factor and compares our Bayesian Factor(red dots) to the Partial Quantile Regression Factor (green dots) from Giglio, Kelly, and Pruitt (2016). The black line plots the S&P500 risk premia. Surprisingly, the shape of our Bayesian Factor coincides with the shape of PQR spanning the sample period from 1990 to 2011. This is very interesting because these two factors are estimated based on two different models: one is Bayesian estimation of a dynamic factor model and the other is Frequentist estimation of a partial quantile regression

Algorithm 1 PGAS non-Markov kernel for the smoothing distribution $p_\theta(g_{1:T}|x_{1:T})$

Require: Reference trajectory $g'_{1:T}$ and parameter $\theta \in \Theta$.
Ensure: Sample $g^\star_{1:T} \sim P^N_\theta(g'_{1:T}, \cdot)$ from the PGAS Markov kernel.

Draw $\{g^i_1, g^i_2, g^i_3, g^i_4\} \sim r_{\theta,1}(g_1, g_2, g_3, g_4|x_1, x_2, x_3, x_4)$ for $i = 1 \cdots N-1$.
Set $\{g^N_1, g^N_2, g^N_3, g^N_4\} = \{g'_1, g'_2, g'_3, g'_4\}$.
Set $\{w^i_1, w^i_2, w^i_3, w^i_4\} = G_\theta(x_1, x_2, x_3, x_4|g^i_1, g^i_2, g^i_3, g^i_4)/\star$ By observation equation $\star/$
for $t = 5$ to T **do**
  Generate $\{\tilde{g}^i_{1:t-1}\}^{N-1}_{i=1}$ by sampling $N-1$ times with replacement from $\{g^i_{1:t-1}\}^N_{i=1}/\star$ with probability proportional to the importance weights $\{w^i_{t-1}\}^N_{i=1}$
  $/\star$ Multinomial Resampling Scheme $\star/$
  Draw J with
  $P(J=i) = \frac{w^i_{t-1} G_\theta(x_{t+2}|g^i_{t-1}, g'_{t:t+2}) G_\theta(x_{t+1}|g^i_{t-2:t-1}, g'_{t:t+1}) G_\theta(x_t|g^i_{t-3:t-1}, g'_t) f_\theta(g'_t|g^i_{t-1})}{\Sigma_l w^l_{t-1} G_\theta(x_{t+2}|g^l_{t-1}, g'_{t:t+2}) G_\theta(x_{t+1}|g^l_{t-2:t-1}, g'_{t:t+1}) G_\theta(x_t|g^l_{t-3:t-1}, g'_t) f_\theta(g'_t|g^l_{t-1})}$, for $i = 1 \cdots N$
  and set $\tilde{g}^N_{1:t-1} = g^J_{1:t-1}$.
  $/\star$ Particle propagation $\star/$
  Simulate $g^i_t \sim f_\theta(g_t|\tilde{g}^i_{t-1})$ for $i = 1, \cdots, N-1/\star$ By transition equation $\star/$
  Set $g^N_t = g'_t$.
  Set $g^i_{1:t} = (\tilde{g}^i_{1:t-1}, g^i_t)$ for $i = 1, \cdots, N$
  $/\star$ Importance Weights $\star/$
  Set $w^i_t = G_\theta(x_t|g^i_t)$
end for
Draw k with $P(k=i) \propto w^i_T$.

Algorithm 2 PGAS for State Space Model

Set $\theta[0]$ and $g_{1:T}[0]$ arbitrarily.
for $n \geq 1$ **do**
  Draw $g_{1:T}[n] \sim P^N_{\theta[n-1]}(g_{1:T}[n-1], \cdot)./\star$ By running Algorithm 1 $\star/$
  Draw $\theta[n] \sim p(\theta|g_{1:T}[n], x_{1:T})$
end for

model. This coincidence implies that both factors could be good candidates to proxy for systemic risk. However, our Bayesian Factor is better in tracking the movement of S&P500 risk premia while PQR is almost always below the risk premia line. Furthermore, PQR can not explain the cross-section of individual stock excess returns based on a two-stage Fama-Macbeth regression, which we will discuss more in the asset pricing implication section. More interestingly, Figure 18 shows our Bayesian factor is really correlated with new measures capturing the spillover effect of distress risk or financial crash risk, such as Co_Var, DCI,and MES, instead of old fashion volatility measures or liquidity measures. In contrast, PQR can not achieve this goal. In all, we can say our Bayesian Factor is a better measure of financial systemic risk for asset pricing purposes.

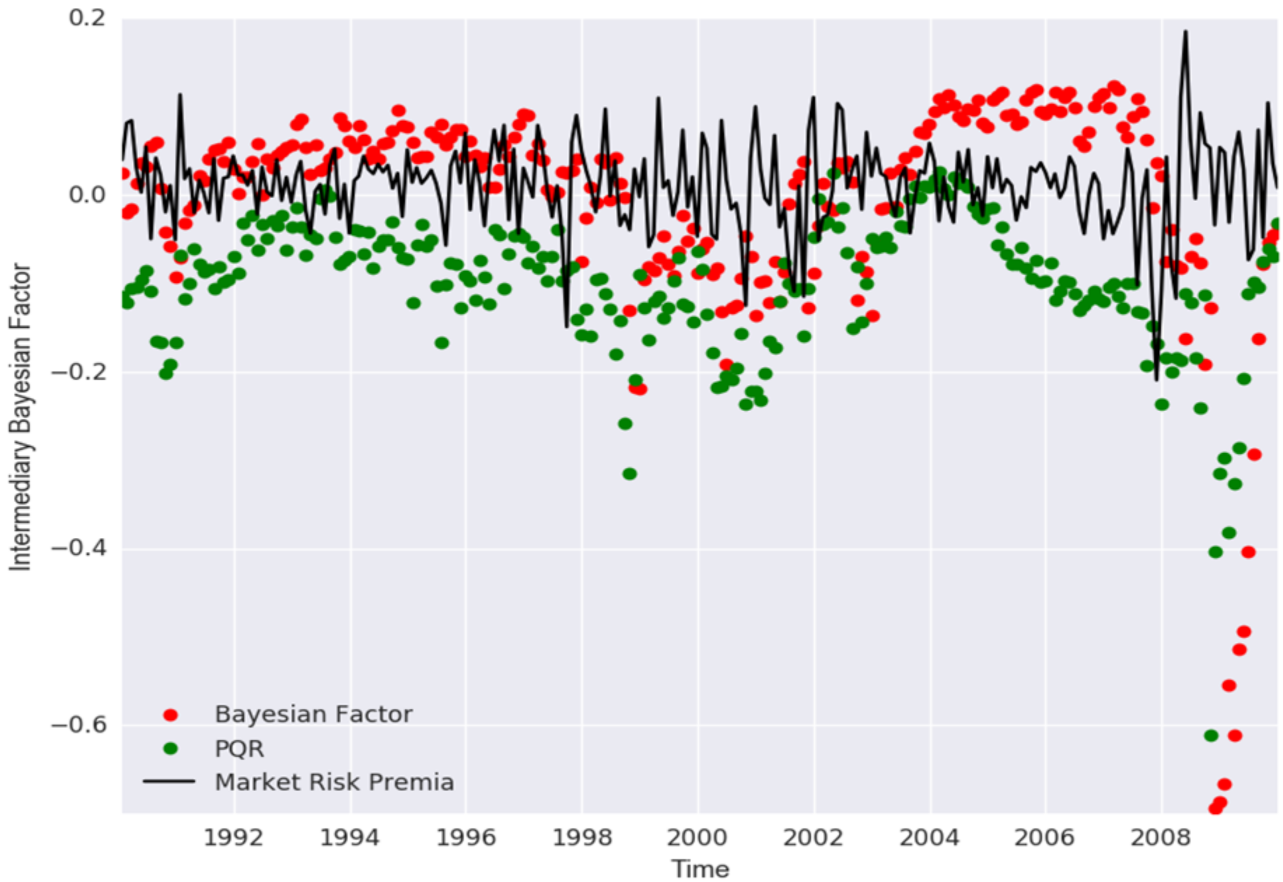


Figure 2. This figure compares Bayesian factor to partial quantile regression factor (PQR) by Gigli, Kelly, and Pruitt (2016).

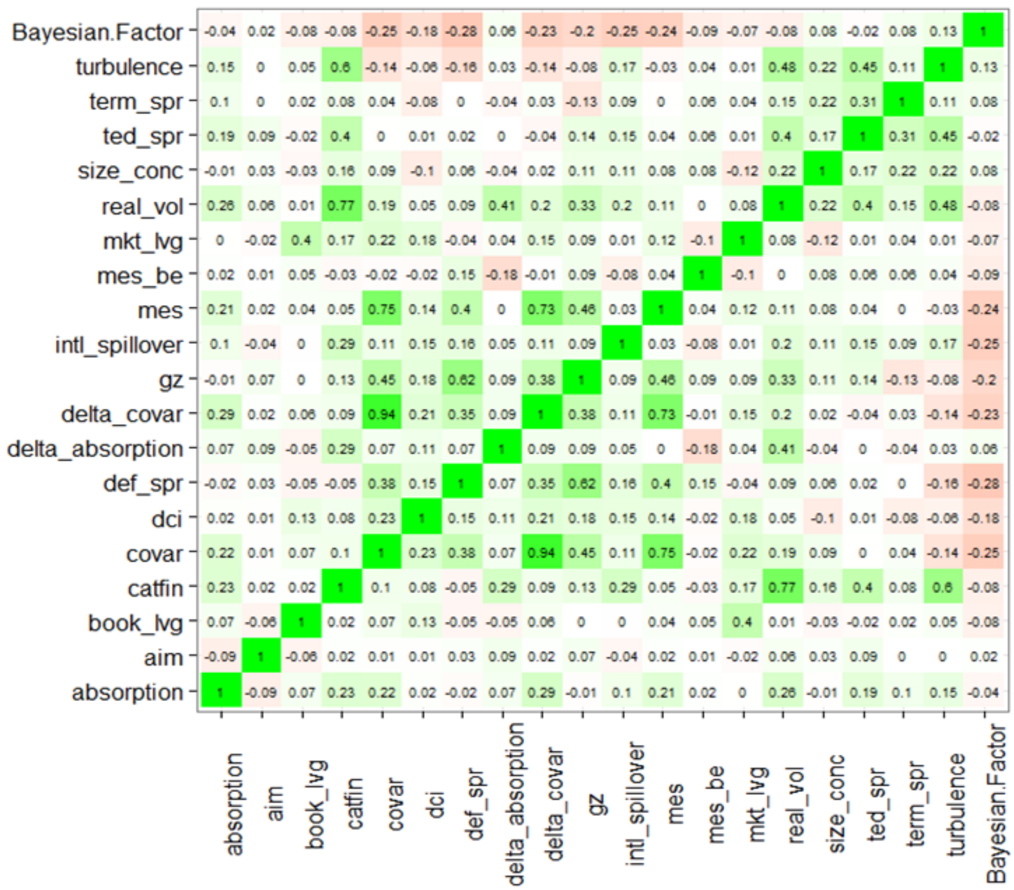


Figure 3. This figure plots the correlation matrix of shocks to all the individual financial systemic risk measures plus the Bayesian Factor.

C. High Frequency Textual Systemic Risk Factor

This section describes how our high-frequency textual factor is constructed in three steps. The first step is sift through news articles from the New York Times business section and to construct a textual vector based on the semantic neural network language model Word Embedding. Then, the second step is using LASSO to project the monthly Bayesian factor onto the monthly textual vector and using the fitted value as our monthly textual Bayesian factor. Thirdly, we can feed the high-frequency textual data to this trained LASSO model and take the fitted value as our high-frequency textual risk factor. We also project principal components of cross-section of risk measures onto word vectors as a robustness check. More details about PCA can be found in the Appendix.

C.1. Neural Network Language Model

The textual information in this paper consists of the headlines and abstracts of all the articles from the business section of New York Times from 1980 to 2016. We collect the daily newspaper information from New York Times API website using Python from January 1, 1980 to December 31, 2016. We adopt a new technology called Word Embedding to collect distributional information on a per word basis and store it in high-dimensional vectors. The vector can then be used as a representational framework to characterize how any given word is semantically related to other words in the corpus. Simply speaking, one word can be represented by one high-dimension vector. This step is done using neural networks as in Mikolov, Sutskever, Chen, Corrado, and Dean (2013)

.

Word embedding/Word2Vec tries to understand words meaning based on those appearing in their contexts. Our goal is to parameterize this model using word vectors, a simple feed-forward neural network model. In this model, Word embedding/Word2Vec is just an unsupervised training of a shallow two-layer feed forward neural network model of words. Usually, the neural network includes three layers: the input layer, the hidden layer, and the output layer. The input layer consists of one-hot vector representation of word w_t. So the input layer has $|V|$ nodes, which is the size of vocabulary. The hidden layer consists of $d = 100$ hidden units, which is determined via state-of-art technique. The values of these components are computed using the embedding matrix W of size $|V| \times d$. The output layer has each component to represent the probability of each of context words given input word w_t. Similar to the input word w_t, the context words are also represented as one-hot vectors. Assuming the number of context word is 4, the number of output component in this case is just $4|V|$. The output layer is computed via multiplying the hidden layer and another output matrix $W^{'}$ of size $d \times |V|$. Then, we can extract the ith row of the embedding matrix W as a vector for word i: $v(w_i) = W_i^T$. We can now write down easily the probability of a context word w_{t+j} given input word w_t in terms of the word vectors, or just the likelihood of this neural network model. Specifically,

$$P(W_{t+j}|W_t) = \frac{e^{v(w_t)\times v^{'}(w_{t+1})}}{\Sigma_{w\in V} e^{v(w_t)\times v^{'}(w)}} \tag{10}$$

where the same distribution is used for $j = \pm 1, \pm 2$ (because the number of context word is 4, so the window size k = 2 in our example). To train this model our goal will be to find $v(w)$ and $v'(w)$, $w \in V$, so as to maximize

$$\max \Sigma_t \Sigma_{j=-k}^{k} \log P(W_{t+j}|W_t) \tag{11}$$

The optimization is done via stochastic gradient ascent algorithm. Again, $v(w_i)$ is just what we call a word vector and it is used to representing each word w_i.

Table XIII shows the machine understanding of the word's meaning based on Word2Vec. Panel

A shows the detection of the least-matching word in one word groups. For example, within the group of "man, woman, child, and kitchen", Word2Vec can detect that the most different word in this group is kitchen. Another example is that Word2Vec can pick Austria out of a group of words including " Paris, Berlin, London, and Austria". Panel B shows the most similar words to one specific word selected by Word2Vec. For example, the top ten most similar words to "Stock" are: Shares, Capital, Nasdaq, Bond, Securities, Eurobond, Value, Bull, Volume, and Mondays. The interesting thing is that Word2Vec can detect a Monday effect in the stock market. Additionally, because each word is represented as a vector, we can implement simple algebraic operations directly on words. A very famous example is " Women + King = Queen ".

C.2. LASSO Trained Textual Factor of Systemic Risk Measures

Now that we have one Bayesian factor, two PCs , and a 100 dimensional word vector representation of textual information, we can project both the Bayesian factor and the PCs onto the 100 dimensional vector using LASSO and take the fitted values as our textual risk measures. The reason to use LASSO here is because LASSO is a dimension reduction technique and we can use LASSO to reduce the over-fitting problem due to the 100 dimensionality of our independent variables. We discuss a little bit more about LASSO as follows. Consider the penalized estimation problem:

$$\hat{\beta}_{penalized} = \arg\min(\Sigma_{t=1}^{T}(y_t - \Sigma_i \beta_i x_{i,t})^2 + \lambda \Sigma_{i=1}^{K} |\beta_i|^q) \tag{12}$$

More specifically, the lasso (short for Least Absolute Shrinkage and Selection Operator), introduced in the seminal work of Tibshirani (1996), solves the OLS plus L1-penalized problem :

Table I: Machine Understanding of Words' Meaning based on Word2Vec

Reported is the mismatch and similarity among words based on Trained Model.

Panel A: Non-Matching Words

Group of Words				Non-Matching Word
Man	Woman	Child	Kitchen	Kitchen
France	England	Germany	Berlin	Berlin
Paris	Berlin	London	Austria	Austria

Panel B: Most Similar Words

Man		Stock		Government		Negative		Positive	
Hero	0.78	Shares	0.71	Authorities	0.73	Gloomy	0.71	Upbeat	0.75
Salesman	0.73	Capital	0.56	Admin	0.70	Conflicting	0.69	Rosy	0.73
Friend	0.73	Nasdaq	0.55	Ministry	0.68	pessimism	0.69	Solid	0.72
Teenager	0.72	Bond	0.54	Parliament	0.68	troubling	0.68	Strong	0.72
Woman	0.70	Securities	0.53	Congress	0.68	persistent	0.66	Strength	0.71
Guy	0.69	Eurobond	0.52	IMF	0.67	Dire	0.64	Resilience	0.69
Boy	0.68	Value	0.49	Regulators	0.64	Lingering	0.63	Reassuring	0.67
Actor	0.68	Bull	0.45	Bailout	0.63	Worsening	0.63	Firmer	0.65
Himself	0.67	Volume	0.44	Reserve	0.62	Warnings	0.62	Robust	0.65
Father	0.66	Mondays	0.43	Reform	0.60	Adverse	0.62	Subdued	0.63

$$\hat{\beta}_{LASSO} = \arg\min(\Sigma_{t=1}^{T}(y_t - \Sigma_i \beta_i x_{i,t})^2 + \lambda \Sigma_{i=1}^{K}|\beta_i|) \tag{13}$$

So Lasso shrinks and selects. It uses the smallest q for which the minimization problem is convex, which is valuable computationally.

Table XVI shows both the in-sample and out-of-sample fit of projecting factors onto word vectors based on LASSO. The results are generally very good. In the training period from January 1990 to December 2006, the in-sample R^2 of projecting PC1 onto the news vector is 90% and the in-sample R^2 of projecting PC2 onto the news vector is 73%, which means news information can be well-trained to capture financial systemic risks. The in-sample R^2 of projecting Bayesian factor onto the news vector is 95%, much higher than the projection R^2 of PCs. We also examine the out-of-sample performance of training, and the R^2 for PC1, PC2, and Bayesian factor are respectively 25% , 17%, and 30%. This reduces a lot of over-fitting concerns about our LASSO model. We also show that news information can be trained to explain individual systemic risk measures very well, such as the Ted-spread and turbulence. The goodness of fit can also be seen from Figure 13, Figure 16, and Figure 14, in which textual factors closely track the original numerical factors.

Another natural question to ask here is whether textual information can ex-ante predict numeric information or whether it just reflects numeric information ex-post. To shed some light on this issue, we further explore the Granger Causality relationship between the textual systemic measures and the numeric risk measures in Table XVII. The results are in favor of textual information granger causing numeric information. Textual PC1 and Textual PC2 both Granger-cause numerical PC1 and PC2 at 1% significance level while numeric PC1 does not cause Textual PC1. Additionally, Textual Bayesian factor can Granger cause the Bayesian Factor while the reverse relationship does

Table II: In-sample and Out-of-sample Distress Risk prediction using word vector based on LASSO

Reported are in-sample and out-of-sample LASSO fit statistics using both the train subsample and test subsample. LASSO is specified as $DistressRisk_t = \beta_0 + \sum_{i=1...100} \beta_{1,i} X_{i,t} + \epsilon_{t+1}$, where $X_{1,t}, \ldots, X_{100,t}$ are 100 vectors unsupervised trained from all business section articles of New York Times in each month from 1980 to 2016. Panel A reports variance of the predicted value Textual measures as a fraction of actual Distress Risk measures variance using in-sample data from 1990.01 to 2006.12. Panel B reports out-of-sample variance of the predicted value Textual factors as a fraction of original distress measures' variance using out-of-sample data from 2007.01 to 2009.12.

Lasso Regression:	$DistressRisk_t = \beta_0 + \sum_{i=1...100} \beta_{1,i} X_{i,t} + \epsilon_{t+1}$
Distress Risk:	In sample R^2
Ted Spread	0.75
Turbulence	0.70
Bayesian factor	0.95
PC1	0.90
PC2	0.73
Distress Risk:	Out of sample R^2
Ted Spread	0.21
Turbulence	0.12
Bayesian factor	0.30
PC1	0.25
PC2	0.17

not hold. This supports the conjecture that news information is ex-ante predicting the numeric information instead of only ex-post summarizing or explaining the numeric information. However, PC2 indeed Granger-causes textual PC2.

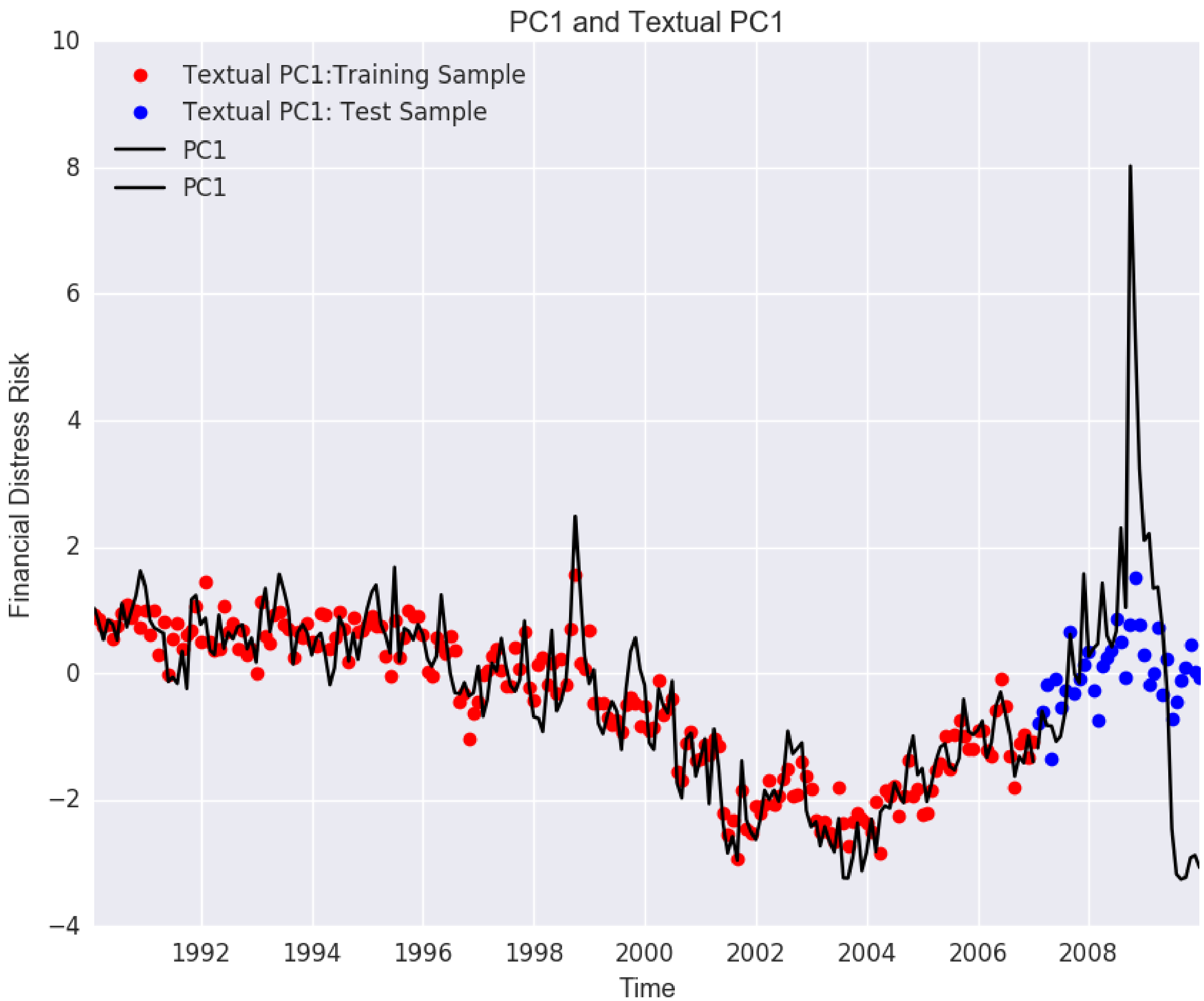


Figure 4. Textual PC1 from 1990 to 2009. Solid line is end-of-month PC1 extracted out of 19 risk measures . Dots are Textual PC1,fitted values from the regression $PC_{1,t} = \beta_0 + \sum_{i=1...100} \beta_{1,i} X_{i,t} + \epsilon_{t+1}$, where $x_{i,t}$ are vector representation of end-of-month text and β is estimated with LASSO. The training subsample, 1990 to 2006, is used to estimate the dependency between news data and financial systemic risk. The test subsample, 2007-2009, is used for out-of-sample tests of model fit.

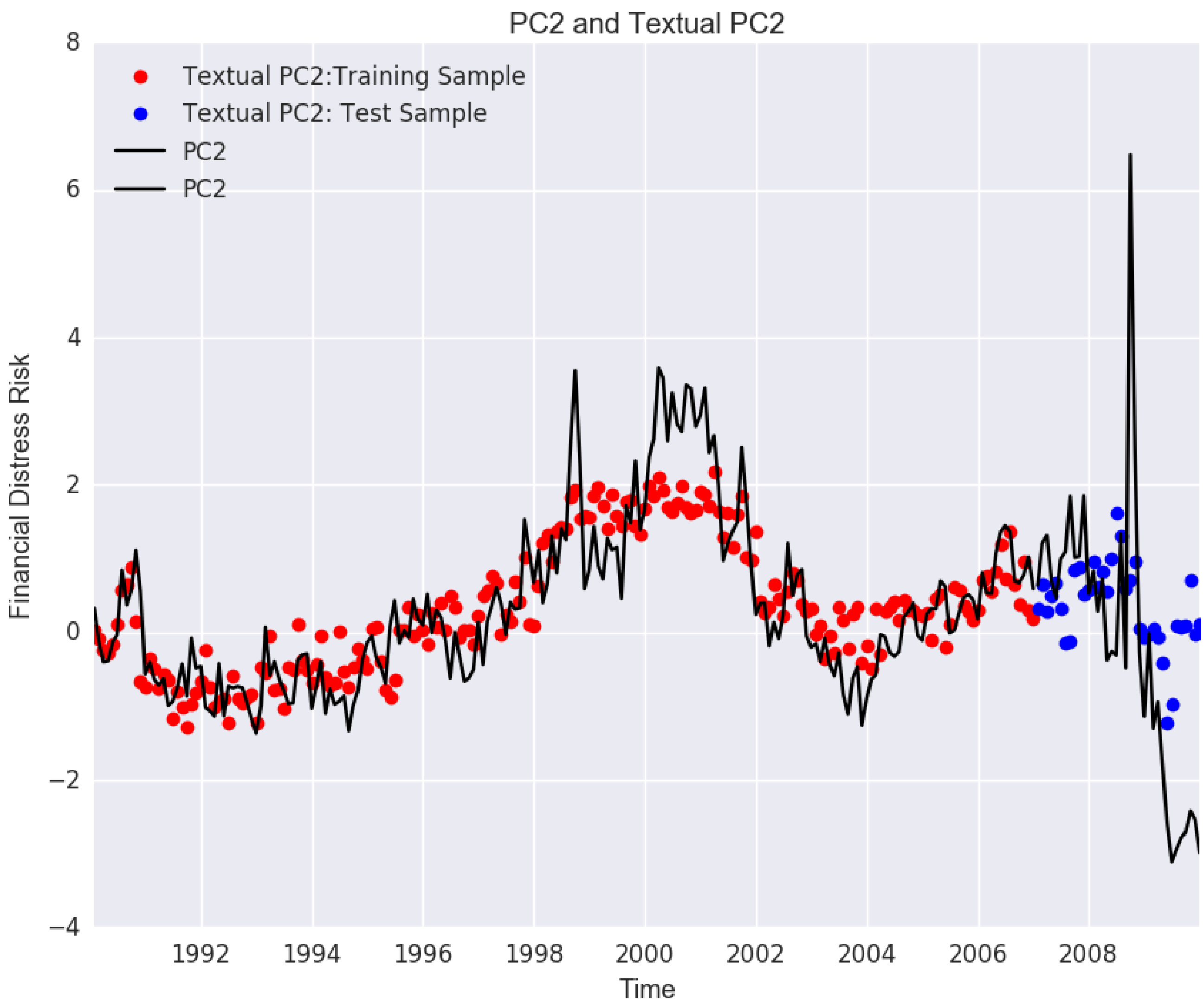


Figure 5. Textual PC2 from 1990 to 2009. Solid line is end-of-month PC2 extracted out of 19 risk measures . Dots are news implied PC2, fitted values from the regression $PC_{2,t} = \beta_0 + \sum_{i=1...100} \beta_{1,i} X_{i,t} + \epsilon_{t+1}$, where $x_{i,t}$ are vector representation of end-of-month text and β is estimated with LASSO. The training subsample, 1990 to 2006, is used to estimate the dependency between news data and financial systemic risk. The test subsample, 2007-2009, is used for out-of-sample tests of model fit.

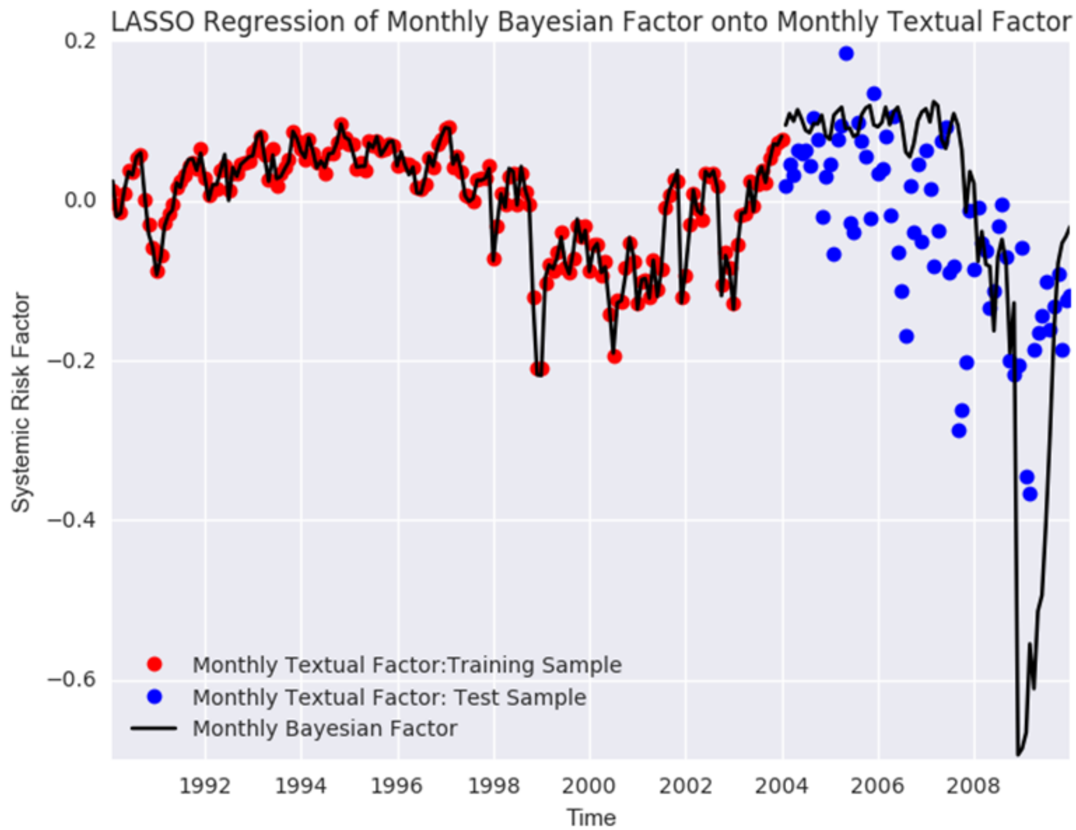


Figure 6. Textual Bayesian Factor from 1990 to 2009. Solid line is end-of-month Bayesian factor extracted out of 19 risk measures . Dots are Textual Bayesian Factors, fitted values from the regression $BayesianFactor_{2,t} = \beta_0 + \sum_{i=1...100} \beta_{1,i} X_{i,t} + \epsilon_{t+1}$, where $x_{i,t}$ are vector representation of end-of-month text and β is estimated with LASSO. The training subsample, 1990 to 2003, is used to estimate the dependency between news data and intermediary systemic risk. The test subsample, 2004-2009, is used for out-of-sample tests of model fit.

Table III: Granger causality tests between news information and systemic risk

Reported are Granger causality tests between textual information and numeric information based on training sample. Panel A reports the test of causality from numeric information to textual information. Panel B reports the test of causality from textual information to numeric information.$*$, $**$, and $***$ indicate 10%, 5%, and 1% significance levels, respectively.

Numerical Information causes News Information:	p-value
PQR	0.026
Bayesian Factor	0.31
PC1	0.18
PC2	0.002***
News Information causes Numerical Information:	**p-value**
PQR	0.001***
Bayesian Factor	0.001***
PC1	0.000***
PC2	0.000***

D. Asset Pricing Implication of Bayesian Factor

D.1. Portfolio Sort Results based on exposure to Bayesian Factor

Table V reports out-of-sample factor regression results across the Bayesian Factor Exposure $\beta_{i,t}^{Bayes}$, where $\beta_{i,t}^{Bayes}$ is firm i's exposure to the Bayesian systemic risk factor in month t. Decile portfolios are formed at the conclusion of each month, ranging from 1 to 10 with the highest (lowest) values located in the 10th (1st) decile. Value-Weighted portfolio returns are measured in the next month and regressed on contemporaneous risk factors: the three Fama-French Factors and momentum factors(UMD).The intercept in this regression(Intercept) is the portfolio alpha. Our main result is that intercepts from these regressions increase with β^{Bayes}, indicating high β^{Bayes} outperform low β^{Bayes} firms. The row "1-10" contains a statistical test for the difference between low and high β^{Bayes} decile portfolios and shows that the 1.1% difference in the four-factor alphas are statistically significant (t-statistics = 2.77). The final row "(1+2) - (9+10)" at the bottom of the table contains another statistic test for the difference of low and high β^{Bayes} quintile portfolios. A strategy that goes long the quintile portfolio with the highest β^{Bayes} and short the quintile portfolio with the lowest β^{Bayes} yields a Fama–French–Carhart alpha of 1.7% per month (t-statistics = 2.77). The sharpe ratio for this strategy is almost 1. Similar results hold for out-of-sample equal-weighted portfolio sort based on exposure to the Bayesian factor in table IV.

Table IV: Out-of-Sample Portfolio Sort Based on Bayesian Factor Exposure

Reported are out-of-sample factor regression results across Bayesian Factor Exposure $\beta_{i,t}^{Bayes}$, where $\beta_{i,t}^{Bayes}$ is firm i's exposure to Bayesian Factor in month t. Decile portfolios are formed at the conclusion of each month, ranging from 1 to 10 with the highest (lowest) values located in the 10th (1st) decile. The sample consists of 814,997 firm-months spanning 1995 through 2013. Equal-Weighted portfolio returns are measured in next month and regressed on contemporaneous risk factors: the three Fama-French Factors and momentum factors(UMD).The intercept in this regression(Intercept) is the portfolio alpha. t-statistics are shown in the underlying row. $*$, $**$, and $***$ indicate 10%, 5%, and 1% significance levels, respectively.

Portfolio Decile	Intercept	mktrf	SMB	HML	UMD
1	0.01***	0.961***	1.06***	-0.016	-0.184***
	5.88	24.37	21.29	-0.3	-5.61
2	0.006***	0.866***	0.695***	0.23***	-0.094***
	4.42	28.36	18.21	5.51	-3.84
3	0.005***	0.8***	0.564***	0.342***	-0.1***
	4.18	32.32	18.21	10.1	-5.01
4	0.005***	0.786***	0.496***	0.374***	-0.091***
	5.06	36.76	18.57	12.81	-5.31
5	0.006***	0.755***	0.491***	0.361***	-0.115***
	5.91	33.98	17.66	11.9	-6.4
6	0.006***	0.788***	0.502***	0.38***	-0.103***
	6.9	37.25	18.98	13.14	-6.03
7	0.006***	0.829***	0.57***	0.386***	-0.119***
	6.17	38.1	20.95	12.97	-6.81
8	0.008***	0.831***	0.631***	0.33***	-0.173***
	6.37	30.52	18.53	8.86	-7.87
9	0.01***	0.933***	0.819***	0.282***	-0.203***
	7.21	30.19	21.2	6.68	-8.15
10	0.017***	0.999***	1.107***	0.111	-0.335***
	7.48	19.65	17.42	1.6	-8.19
1-10	-0.009***	-0.022	-0.016	-0.181**	0.134***
	-3.35	-0.36	-0.21	-2.14	2.69
(1+2)-(9+10)	-0.013***	-0.089	-0.14	-0.233*	0.242***
	-3.22	-0.97	-1.22	-1.86	3.28

Table V: Out-of-Sample Portfolio Sort Based on Bayesian Factor Exposure

Reported are out-of-sample factor regression results across Bayesian Factor Exposure $\beta_{i,t}^{Bayes}$, where $\beta_{i,t}^{Bayes}$ is firm i's exposure to Bayesian Factor in month t. Decile portfolios are formed at the conclusion of each month, ranging from 1 to 10 with the highest (lowest) values located in the 10th (1st) decile. The sample consists of 814,997 firm-months spanning 1995 through 2013. Value-Weighted portfolio returns are measured in next month and regressed on contemporaneous risk factors: the three Fama-French Factors and momentum factors(UMD).The intercept in this regression(Intercept) is the portfolio alpha. t-statistics are shown in the underlying row. $*$, $**$, and $***$ indicate 10%, 5%, and 1% significance levels, respectively.

Portfolio Decile	Intercept	mktrf	SMB	HML	UMD
1	-0.003*	1.171***	0.353***	-0.334***	-0.055
	-1.72	26.57	6.34	-5.47	-1.51
2	-0.001	1.091***	0.082	-0.127**	0.03
	-0.28	26.71	1.6	-2.27	0.9
3	-0.002	1.066***	-0.135***	-0.092*	0.06**
	-1.45	30.29	3.07	-1.91	2.12
4	-0.001	0.924***	-0.099**	-0.017	0.016
	-0.56	29.88	-2.56	-0.4	0.65
5	-0.0	0.862***	-0.103***	0.108***	-0.018
	-0.22	32.36	-3.08	2.97	-0.83
6	0.0	0.947***	-0.178***	0.152***	0.018
	0.34	36.5	-5.48	4.28	0.87
7	0.002	0.945***	-0.083**	0.172***	-0.035
	1.56	34.43	-2.42	4.59	-1.6
8	0.003*	0.946***	0.104**	0.188***	-0.098***
	1.91	28.51	2.51	4.15	-3.68
9	0.005***	1.009***	0.119***	0.038	-0.165***
	3.02	28.67	2.71	0.79	-5.84
10	0.004	1.474***	0.575***	0.191**	-0.18***
	1.4	21.06	6.57	1.99	-3.2
1-10	-0.011***	-0.272***	-0.165	-0.579***	0.11
	-2.77	-2.93	-1.42	-4.56	1.46
(1+2)-(9+10)	-0.017***	-0.19	-0.203	-0.744***	0.304***
	-2.77	-1.41	-1.2	-4.02	2.8

D.2. FamaMacbeth Regression of Bayesian Factor Exposure

We run Fama and MacBeth (1973a) regressions on the individual firm level in the period from 1995 to 2011. Table VI presents the regression results of monthly future excess returns on systemic risk exposure in the first two columns and the results of existing funding illiquidity/constraint exposures in other columns. The first two columns show a significant positive relationship between systemic risk exposure and future stock returns. Our systemic risk measure is low when financial systemic risk is high and so a positive coefficient here implies a negative market price of risk on systemic risk. More interesting, column two shows that after controlling for firm characteristics, the statistical significance of our risk exposure is even increasing. The regression results also give out the economic significance of risk exposure and a one standard deviation increase in systemic risk exposure would result in a 20 basis point increase in stock excess return per month. In contrast, existing measures for funding liquidity or financial constraint are not working very well. β_{FGG} is the exposure to funding liquidity measure from Fontaine, Garcia, and Gungor (2016), β_{CL} is similar funding liquidity measure from Chen and Lu (2014), and β_{Lee} is another measure from Lee (2013). None of three risk exposures are significantly priced in our sample period. This finding is consistent with our initial conjecture that combining both intermediary information and macro economic information is better in constructing the pricing kernel than only using either of them. This is also empirical evidence to motivate a model combining intermediary asset pricing model , disaster risk model, and the ICAPM. Column seven shows another significant risk factor which is the intermediary capital ratio factor. However, the HKM factor is not robust enough because Table VII column four shows that including our risk exposure into the regression would make the HKM factor insignificant. The last column in Table VII shows regression results including all of the risk exposures and our risk exposure is the only significant variable in this regression.

Table VI: Fama-MacBeth Regression of Different Factor Exposures

This table presents Fama-MacBeth regression results from regressing next month returns on my Bayesian factor exposure and on other risk exposures. The sample consists of 566,186 firm-months spanning 1995 through 2011. Each column corresponds to a different regression with different risk factor exposure.β_{FGG} is the funding liquidity measure from Fontaine, Garcia, and Gungor (2016), β_{CL} is a similar funding liquidity measure from Chen and Lu (2014), and β_{Lee} is another measure from Lee (2013).β_{HKM} is the intermediary capital ratio factor exposure. All regressions have the same controls, including StkVol, Ret(0), Ret(11M), Size, B/M, ROA, and ΔROA. Ret(0) is past month firm excess returns. StkVol equals the firm's total stock trading volume in the observation month. Ret(11M) equals the cumulative market-adjusted returns measured over the prior eleven months, and it capture the momentum effect. Size is the log of market capitalization of the firm and B/M is the firm's book-to-market ratio measured at the firm's last yearly announcement date. Amihud is the Amihud illiquidity ratio of the firm in the observation month. All independent variables are standardized in cross section before Fama-Macbeth regression. Standard errors are computed across weekly coefficient estimates,following Fama and MacBeth (1973). The resulting t-statistics are shown in parentheses. $*$, $**$, and $***$ indicate 10%, 5%, and 1% significance levels, respectively.

	M1	M2	M3	M4	M5	M6	M7	M8
Intercept	0.012**	0.028	0.023	0.029	0.024	0.027	0.022	0.028
	(2.43)	(1.48)	(1.39)	(1.53)	(1.33)	(1.45)	(1.24)	(1.47)
β_Bayes	0.029***	0.015***	–	–	–	–	–	–
	(2.72)	(3.12)	–	–	–	–	–	–
β_Mkt	–	–	0.003	–	–	–	–	–
	–	–	(1.21)	–	–	–	–	–
β_FGG	–	–	–	0.024	–	–	–	–
	–	–	–	(0.68)	–	–	–	–
β_CL	–	–	–	–	-0.001	–	–	–
	–	–	–	–	(-0.34)	–	–	–
β_Lee	–	–	–	–	–	-0.011	–	–
	–	–	–	–	–	(-1.16)	–	–
β_HKM	–	–	–	–	–	–	0.007**	–
	–	–	–	–	–	–	(2.60)	–
β_PQR	–	–	–	–	–	–	–	-0.000
	–	–	–	–	–	–	–	(-1.18)
Ret(1M)	–	-0.055***	-0.057***	-0.054***	-0.056***	-0.054***	-0.056***	-0.054***
	–	(-6.83)	(-7.43)	(-7.06)	(-7.07)	(-7.26)	(-7.28)	(-6.70)
Ret(11M)	–	-0.002	-0.001	-0.001	-0.002	-0.002	-0.002	-0.002
	–	(-0.48)	(-0.39)	(-0.32)	(-0.61)	(-0.66)	(-0.50)	(-0.50)
Size	–	-0.004***	-0.004***	-0.004***	-0.004***	-0.004***	-0.004***	-0.004***
	–	(-4.26)	(-4.21)	(-4.29)	(-4.09)	(-4.25)	(-4.17)	(-4.19)
StkVolume	–	0.000***	0.000***	0.000***	0.000**	0.000**	0.000**	0.000***
	–	(2.68)	(2.80)	(2.66)	(2.59)	(2.57)	(2.56)	(2.62)
Amihud	–	0.002**	0.002***	0.002**	0.001**	0.001**	0.002***	0.002**
	–	(2.35)	(2.94)	(2.32)	(2.31)	(2.39)	(2.71)	(2.33)
Profitability	–	0.016*	0.017*	0.017*	0.016*	0.017*	0.016*	0.016*
	–	(1.77)	(1.89)	(1.93)	(1.87)	(1.90)	(1.79)	(1.85)
Distress	–	-0.003**	-0.003**	-0.003**	-0.003**	-0.003**	-0.003**	-0.003**
	–	(-2.01)	(-2.28)	(-2.00)	(-2.22)	(-2.09)	(-2.48)	(-1.99)
BE/ME	–	0.005**	0.005**	0.005**	0.005**	0.005**	0.005**	0.005**
	–	(2.33)	(2.59)	(2.30)	(2.39)	(2.34)	(2.41)	(2.46)
Adj_R^2	0.003	0.048	0.054	0.049	0.049	0.05	0.05	0.048
No. Obs	566185	566185	566185	566185	566185	566185	566185	566185

Table VII: Comparison among Different Factor Exposures

This table compares Fama-MacBeth regression results from regressing the next month returns on my Bayesian factor exposure and on other risk exposures. The sample consists of 566,186 firm-months spanning 1995 through 2011. Each column corresponds to a different regression with different risk factor exposure after controlling for our risk exposure. All regressions have the same controls, including StkVol, Ret(0), Ret(11M), Size, B/M, ROA, and ΔROA. Ret(0) is past month firm excess returns. StkVol equals the firm's total stock trading volume in the observation month. Ret(11M) equals the cumulative market-adjusted returns measured over the prior eleven months, and it captures momentum effect. Size is the log of market capitalization of the firm and B/M is the firm's book-to-market ratio measured at the firm's last yearly announcement date. Amihud is the Amihud illiquidity ratio of firm in observation month. All independent variables are standardized in cross section before Fama-Macbeth regression. Standard errors are computed across weekly coefficient estimates,following Fama and MacBeth (1973). The resulting t-statistics are shown in parentheses. $*$, $**$, and $***$ indicate 10%, 5%, and 1% significance levels, respectively.

	M1	M2	M3	M4	M5	M6
Intercept	0.025	0.021	0.024	0.023	0.023	0.022
	(1.49)	(1.30)	(1.44)	(1.40)	(1.38)	(1.36)
β_Bayes	0.013***	0.013***	0.014***	0.012**	0.016***	0.009**
	(3.03)	(3.09)	(3.26)	(2.47)	(3.49)	(2.04)
β_FGG	0.019	–	–	–	–	0.031
	(0.72)	–	–	–	–	(1.24)
β_CL	–	-0.002	–	–	–	-0.004
	–	(-0.80)	–	–	–	(-1.29)
β_Lee	–	–	-0.008	–	–	-0.004
	–	–	(-1.20)	–	–	(-0.61)
β_HKM	–	–	–	0.004	–	0.004
	–	–	–	(1.16)	–	(1.26)
β_PQR	–	–	–	–	-0.000	-0.000
	–	–	–	–	(-0.67)	(-0.83)
β_Mkt	0.002	0.002	0.002	0.001	0.002	-0.000
	(1.10)	(0.99)	(0.97)	(0.34)	(0.91)	(-0.06)
Ret(1M)	-0.061***	-0.060***	-0.059***	-0.060***	-0.060***	-0.063***
	(-7.44)	(-7.21)	(-7.37)	(-7.31)	(-6.99)	(-7.09)
Ret(11M)	-0.000	-0.001	-0.000	-0.001	-0.001	-0.000
	(-0.14)	(-0.37)	(-0.11)	(-0.37)	(-0.28)	(-0.08)
Size	-0.004***	-0.004***	-0.004***	-0.004***	-0.004***	-0.004***
	(-4.34)	(-4.18)	(-4.32)	(-4.34)	(-4.23)	(-4.43)
StkVolume	0.000***	0.000***	0.000***	0.000***	0.000***	0.000***
	(2.90)	(2.82)	(2.81)	(2.86)	(2.83)	(2.82)
Amihud	0.002***	0.002***	0.002***	0.002***	0.002***	0.002***
	(2.95)	(2.99)	(2.99)	(2.97)	(2.99)	(3.03)
Profitability	0.017*	0.016*	0.016*	0.015*	0.016*	0.016*
	(1.89)	(1.86)	(1.87)	(1.77)	(1.85)	(1.85)
BE/ME	0.005**	0.005**	0.005**	0.005**	0.005**	0.004**
	(2.40)	(2.51)	(2.46)	(2.43)	(2.54)	(2.35)
Adj_R^2	0.055***	0.055***	0.055***	0.056***	0.055***	0.059***
No. Obs	566185	566185	566185	566185	566185	566185

D.3. Relationship between Systemic Risk Factor and Firm Distress Characteristics

As our systemic risk factor is intending to measure macro financial crash risk, a firm's exposure to this factor should be capturing firm's distress probability: the more sensitive a firm is to the systemic risk factor, the higher is the firms distress probability. So, in this section, we explore the link between systemic risk exposure and firm characteristics related to distress risk. More specifically, we try to see if a firm's exposure to the systemic risk factor can affect the ability of the firm's distress-related characteristics to predict stock returns. We choose three firm characteristics related to distress risk: distress probability from Campbell, Hilscher, and Szilagyi (2008) Chava and Jarrow (2004) Chava (2014) Alanis, Chava, and Kumar (2016), gross profitability from Novy-Marx (2013), and idiosyncratic volatility from Ang, Hodrick, Xing, and Zhang (2006). Table VIII shows the Fama-MacBeth regression results from regressing next month returns on those three characteristics with or without controlling for our systemic risk exposure. The main result is that systemic risk exposure indeed shrinks the predictability of distress related characteristic for stock returns. Column four in table VIII shows that distress probability is negatively associated with stock returns with a t-statistics of -2.48. Inclusion of systemic risk exposure in column five shrinks in absolute value the distress probability t-statistic to -1.77, a barely significant level. Similarly, the idiosyncratic volatility t-statistic has been reduced n absolute value from -2.07 in column six to -1.33 in column 7 after including systemic risk exposure. To gain more insight in this direction, we construct a five-factor alpha based on a new five-factor model that includes systemic risk factor. The systemic risk factor here consists of a mimicking portfolio returns constructed by long the low systemic risk exposure quintile and short the high systemic risk exposure quintile. This table shows a big difference between the four-factor alpha and the five-factor alpha. For example, t he four-factor distress alpha is -0.638 per month while the five-factor distress alpha is -0.406 per month, a 37% decrease in alpha. The t-statistics also shrinks in absolute value from -2.07 to -1.21, making the distress anomaly not significant anymore. People used to struggle to explain this negative distress anomaly. Here, as distress anomaly loads positively on the systemic risk mimicking factor, a possible way to explain the distress anomaly is that more distressed firms have bigger exposure to the systemic risk and then distress risk is a severe risk to investors. With respect to the profitability anomaly and the idiosycratic volatility anomaly, both of them are explained by the systemic risk factor about 22%. Table X shows the results of a bivariate portfolio sort conditioned on systemic risk exposure first and then each of three firm characteristics. We find that after neutralizing systemic risk exposure, those three anomalies are not significant anymore.

Table VIII: Bayesian Factor Exposure and Firm Characteristics

This table compares Fama-MacBeth regression results from regressing next month returns on distress-related firm characteristics with or without controlling for systemic risk exposure. The sample consists of 566,186 firm-months spanning 1995 through 2011. Each column corresponds to each regression with or without including the Bayesian systemic factor exposure. GP is gross profitability (Robert Novy-Marx,2013), Distress is default probability(Campbell, Hilscher and Szilagyi, 2008), and IdVOL is idiosyncratic volatility. All regressions have the same controls, including StkVol, Ret(0), Ret(11M), Size, B/M, ROA, and ΔROA. Ret(0) is past month firm excess returns. StkVol equals the firm's total stock trading volume in the observation month. Ret(11M) equals the cumulative market-adjusted returns measured over the prior eleven months , and it captures the momentum effect. Size is the log of the firm's market capitalization and B/M is the firm's book-to-market ratio measured at the firm's last yearly announcement date. Amihud is the Amihud illiquidity ratio of firm in observation month. All independent variables are standardized in cross section before the Fama-Macbeth regression. Standard errors are computed across weekly coefficient estimates,following Fama and MacBeth (1973). The resulting t-statistics are shown in parentheses. $*$, $**$, and $***$ indicate 10%, 5%, and 1% significance levels, respectively.

	M1	M2	M3	M4	M5	M6	M7
Intercept	0.026	0.022	0.024	0.025	0.023	0.022	0.023
	(1.53)	(1.32)	(1.41)	(1.41)	(1.37)	(1.33)	(1.37)
GP	–	0.0025***	0.002**	–	–	–	–
	–	(3.3)	(2.36)	–	–	–	–
Distress	–	–	–	-0.0041**	-0.0033*	–	–
	–	–	–	(-2.48)	(-1.77)	–	–
IdVOL	–	–	–	–	–	-0.0029**	-0.002
	–	–	–	–	–	(-2.07)	(-1.33)
β_{Bayes}	0.0015***	–	0.0013**	–	0.0011**	–	0.0011**
	(2.58)	–	(2.35)	–	(2.04)	–	(2.04)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adj_R^2	0.051	0.051	0.051	0.051	0.051	0.051	0.051
No. Obs	566185	566185	566185	566185	566185	566185	566185

Table IX: Anomaly strategy average returns and abnormal performance

This table reports the average excess returns and alphas for strategies formed by sorting on profitability, distress and idiosyncratic volatility. Strategies are longshort extreme quintiles from a sort on the corresponding variable,employing NYSE breaks,and returns are value-weighted. FF4 alpha is Fama-French four factor alpha, and alpha is constructed from five factor model including systemic risk mimicking portfolio. t-statistics are shown in the underlying row. $*$, $**$, and $***$ indicate 10%, 5%, and 1% significance levels, respectively. The sample covers January 1995 through December2011.

Sorting Variables			alpha and factor loadings					
	H-L	FF4 alpha	alpha	Mkt	SMB	HML	UMD	MimicFactor
Profitability	0.412**	0.485**	0.378	-0.136***	-0.121**	-0.165**	0.092**	-0.035*
	2.02	2.00	1.51	-2.7	-2.01	-2.43	2.31	-1.66
Distress	-0.761	-0.638**	-0.406	0.509***	0.430***	0.017	-0.910***	0.122***
	-1.14	-2.07	-1.21	8.25	5.82	0.21	-18.67	2.91
IdVOL	-0.926	-0.948**	-0.774	0.994***	1.194***	-0.798***	-0.436***	0.096*
	-1.14	-2.10	-1.61	10.17	10.75	-6.39	-5.95	1.67

Table X: Value-Weighted Portfolio Sort based on Bayesian Factor Exposure and Firm Characteristics

This table reports average monthly returns of portfolios sorted on stock characteristics and Bayesian factor exposure over our sample period from January 1995 to December 2011. Each month we sort stocks in ascending order into quintile portfolios on the basis of one of our Bayesian factor exposure measures. Within each quintile, we sort stocks into five additional portfolios (Low, 2, 3, 4, and High) based on each firm characteristics, and then average each of the portfolios across the five quintiles that resulted from the first sort. We use value-weigh stocks in each portfolio.We compute the difference between portfolio High and portfolio Low (H-L) and four-factor alpha. t-statistics are shown in the underlying row. $*$, $**$, and $***$ indicate 10%, 5%, and 1% significance levels, respectively.

	Double Sort Conditioned on β_{Bayes}					
Variable	H-L	α	Mkt	SMB	HML	UMD
Profitability	0.192	0.305	-0.14**	-0.146**	0.092	0.063
	0.76	1.25	-2.47	-2.23	-1.33	1.57
Distress	-0.465	-0.448	0.408***	0.384***	0.054	-0.814***
	-0.74	-1.45	5.86	4.75	0.64	-16.55
IdVOL	0.018	-0.285	0.76***	1.056***	-0.607***	-0.397***
	0.02	-0.68	7.8	9.36	-5.14	-5.78

E. Property of High Frequency Daily Textual Factor

As we can use LASSO to project the monthly systemic risk factor onto monthly textual information with good fit both in sample and out of sample (section II.C), we can then multiply the LASSO weights by daily textual information to construct a high-frequency(daily) textual systemic risk factor. Actually, this an unsupervised learning of the high-frequency systemic risk factor due to the lack of the supervised objective, the true systemic risk factor. But, we are comfortable to say that our high-frequency measure captures what we want because we find that this daily textual systemic risk factor has a significant 12% correlation with the daily systemic-risk Mimicking-portfolio returns at 99% confidence level. On top of this, the daily textual systemic risk factor is better than the daily systemic-risk mimicking-portfolio returns in predicting future macro financial states, for example the VIX index. We focus on VIX here because Engle and Rangel (2008) refer to the relation between the macro economy and stock market volatility as the central unsolved problem of 25 years of volatility research. We try to shed some light on this issue and use our high-frequency daily systemic risk to predict this very informative measure VIX at the daily level. Table XI show estimates of predictive regressions of future VIX shocks onto high-frequency systemic risk shocks. The textual systemic risk shocks have significant positive association with future VIX shocks while the systemic risk mimicking portfolio does not show any predictive power for future VIX shocks.

We also formally use the Fama-Macbeth two-step procedure to estimate the market price of risk on this high-frequency daily textual systemic risk factor and Table XII show a consistent positive market price on this high-frequency risk factor. This is a pure out-of-sample test because the training period of the high-frequency risk factor is from 1990 to 2003 while the asset pricing test sample is from 2007 to 2015. Each row corresponds to an asset pricing test with different portfolios. The first row uses the Fama-French 25 portfolios based on size nad momentum, the second row utilizes the Fama-French 25 portfolio based on size and book-to-market ratio, and the last row uses

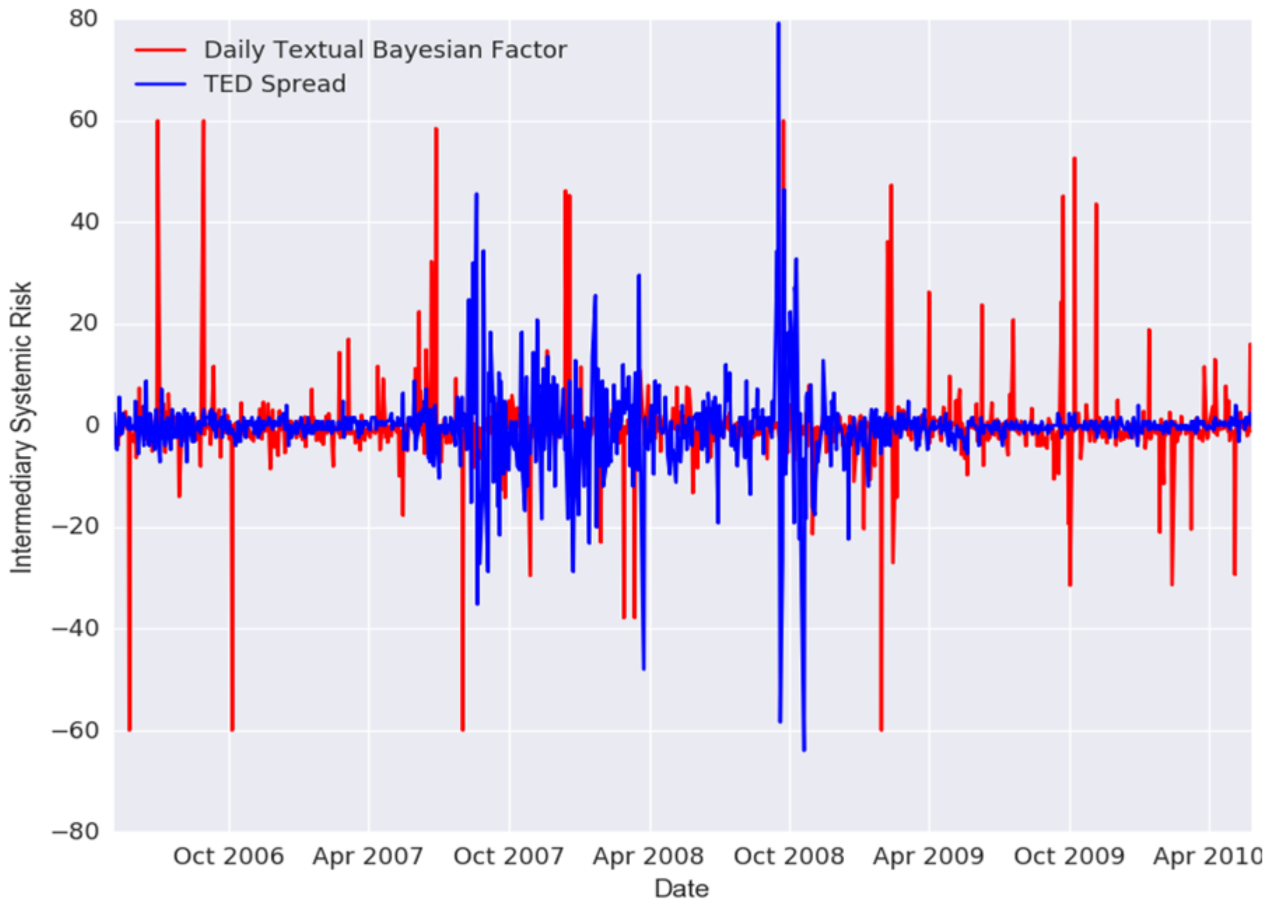


Figure 7. The Daily Textual Bayesian Factor and TED Spread from 2006 to 2010. Red line is the daily Textual Bayesian Factor and blue line is the daily TED Spread.

the Fama-French 25 portfolios based size and investment. All three tests show a positive market price on the daily systemic risk factor, consistent with monthly Fama-Macbeth regression in table VI. The estimated annualized risk premiums based on three sets of portfolios are respectively 8.7%, 6.44%, and 4.92%. More importantly, the J-statistics are not very high in this case, which means this four-factor model is probably a good asset pricing model.

Table XI: Predicting ΔVIX

This table shows estimates of the following predictive regression: $\Delta VIX_{t+1} = a + bSystemicShock_t + \epsilon_{t+1}$. The sample is daily observations from January 2006 through December 2014. Robust Newey-West t-statistics are reported in parentheses.

Predictive Regression:	$\Delta VIX_{t+1} = a + b * SystemicShock_t + \epsilon_{t+1}$			
Constant	0.0018	0.0022	0.0019*	0.0021**
	(1.14)	(1.377)	(1.860)	(2.056)
VIX_t		-0.119***		-0.0868***
		(-5.6)		(-5.396)
Textual Systemic Risk Shock	0.0129**	0.0115**		
	(2.245)	(2.024)		
Systemic Risk Mimicking Portfolio			-0.0546	0.0367
			(-0.562)	(0.373)

Table XII: Market Price of Daily Textual Factor Mimicking Portfolio

Reported are the market price risk for the Daily Textual Factor across different portfolios. Each row is estimated as $E[R^e] = \lambda_0 + \beta_{fac}\lambda_{fac}$. $E[R^e]$ are expected excess returns for each portfolios. VWM^e, SMB and HML are the Fama-French three factors. HKM denotes the intermediary capital factor from He, Kelly and Manela(2016). Bayesian Factor is the first difference of our Bayesian Factor. The first row estimates the market price of risk using monthly Fama-French 25 Portfolios Formed on Size and Book-to-Market from Jan 1990 to Dec 2009. The second row estimates market price of risk using monthly Fama-French 25 Portfolios Formed on Size and Momentum. The third row estimates market price of risk using both Fama-French 25 Portfolios Formed on Size and Book-to-Market and Fama-French 10 Portfolios Formed on Momentum from Jan 1990 to Dec 2009. The following rows report similar results for bond portfolios, option portfolios, commodity portfolios, and currency portfolios. The J statistic, which tests whether the pricing errors are jointly zero, is shown in the last column of each row. Fama-MacBeth t-statistics are reported in parenthesis and $*$, $**$, and $***$ indicate 10%, 5%, and 1% significance levels, respectively.

	VWM^e	SMB	HML	SystemicRisk		
25SizeMoM0715	10.974***	1.904***	-0.035	8.700 ***	J-stat:	46.2
	(22.081)	(8.198)	(-0.074)	(17.489)		(0.006)
25SizeBe0715	9.982 ***	0.954***	-1.735***	6.437 ***	J-stat:	54.0
	(20.105)	(4.303)	(-6.466)	(28.121)		(0.001)
25SizeInv0715	10.033***	-1.906***	8.117 ***	4.916 ***	J-stat:	45.0
	(20.221)	(-7.043)	(18.489)	(13.143)		(0.008)

F. Conclusion

We extract a commonality of 19 individual financial distress measures to construct an aggregate Bayesian systemic risk factor based on Particle-MCMC. The Bayesian factor is significantly and consistently priced by different portfolios and implies a negative market price of financial systemic risk. Furthermore, exposure to the Bayesian Factor can predict the future cross-section of stock excess returns and firms with high Bayesian Factor exposures out-perform those with low Bayesian Factor exposures by 1.7% per month. Additionally, We find our systemic risk factor is linked to distress related firm characteristics and more distressed firms are more sensitive to our systemic risk factor. This shed some light on why distress risk could be a severe risk for investors. Incorporating our systemic risk factor can help explain 36% of the distress anomaly.

More interestingly, we project our monthly Bayesian latent factor onto textual data to construct a high-frequency textual risk measure. Textual information indeed shows good explanatory power for systemic risk both in-sample and out-of-sample. Simple Granger Causality test shows that Textual Factors can Granger cause the Bayesian factor while the Bayesian factor can not Granger cause the news based factor. This sheds some light on an open question of whether textual information only reflects numeric information or could lead numeric information in the first place. Besides, we utilize the fitted values out of our LASSO projection to construct a high-frequency risk factor and we find this high-frequency factor is also consistently and significantly priced across different equity portfolios. Actually, it has a significant 12% correlation with daily systemic-risk mimicking-portfolio returns at a 99% confidence level. On top of this, the daily textual systemic risk factor is more informative than the daily systemic-risk mimicking-portfolio returns in predicting future macro financial conditions, for example, the time-varying VIX index.

II. Chapter 2: Combining Bond Demand and Supply Factors

Recent empirical study in asset pricing has uncovered significant predicable variation in the excess returns of U.S. government bonds, a violation of expectations hypothesis.(Cochrane and Piazzesi (2002),Ludvigson and Ng (2009)). Except for those literature mainly focusing on demand of bond, there is another stream of literate examining how supply of bond affect bond excess returns. (Greenwood, Hanson, and Stein (2015),Greenwood and Vayanos (2014),Krishnamurthy and Vissing-Jorgensen (2011)) While individual demand and supply measures are explored in separate papers, there has been little empirical analysis of them as a group.The first question this paper addresses is how to extract commonality of "factor zoos" related to both demand of bond and supply of bond. One way around "big data" problem is asymptotic principle component analysis for high-dimensional datasets. Asymptotic PCA allows me to extract a few estimated factors out of a large number of series , and eliminate the arbitrary reliance on a small number of indicators to proxy for macroeconomic fundamentals. To my best knowledge, this paper is the first one to combine both demand of bond and supply of bond to revisit the question of whether there exists important macro economic factors in bond risk premia . Another way to solve this kind of problem is using machine learning technique called LASSO. We use LASSO predictive regression to confirm the predictability of bond demand factors and supply factors and also to pick out the most important underlying macro economic variables for the purpose of prediction. So, this paper adopts two different methodology to estimates common latent factors from a monthly panel of 133 measures of economic activity and divide them into two groups: demand factor and supply factor. We begin with a comprehensive analysis of whether demand factor or supply factor predict excess bond returns, and then move on to investigate how related factors influence excess bond returns and which one would be more important. Our results indicate that excess bond returns are indeed forecastable by both demand and supply factor in the sample from 1996:1 to 2015:12. The predicative power of estimated factors is stronger for longer bonds and have their strongest predictive power for five-year bonds, explaining 24.58% of the one-year-ahead variation in their excess returns. More interestingly, CP factor is not significant in predicting excess bond returns anymore in this sample. This means that the estimated factors contain substantial information about future bond returns that is not contained in CP factor. We also try to put economic interpretation on estimated latent factors. The first demand factor, the most important demand factor out of all the estimated factors we study, is highly correlated with output and labor market but not highly correlated with prices or financial market. The second demand factor highly correlated with measures of the aggregate price-level and the fifth factor highly correlated with consumption and inventory orders also have significant predictive power for excess bond returns. On the other hand, the first supply factor highly correlated with public debt and the second supply factor highly correlated with money supply have substantial predictive power for excess bond returns. On top of this roughly economic characterization of bond factors, we use LASSO to identify 15 most important underlying economic variables out of original 133 variables and classify them into bond demand group and supply group. This procedure helps to give the direct economic meaning to bond demand variables and supply variables. We find labor market variables are the most important components of our demand factors and total amount of us government bonds outstanding is the most important supply variable.

Except for bond return predictability of demand factors and supply factors, we also examine the cross-sectional stock return predictability of those factors. This is motivated by the hypothesis that bond factors should be good state variables for pricing kernel and can price different classes of asset prices. Actually, this paper is the first study trying to sort stock portfolios based on stocks' exposure to bond factors. This idea is consistent with consumption/production capital asset pricing model or Merton (1973) ICAPM, the building blocks of modern asset pricing theory. Over the recently thirty yeas, there are very few research along this direction because people can not find empirical evidence for a priced consumption growth factor or production growth factor. This paper tries to utilize exhaustively many economic variables and macroeconomy-bond nexus to shed some light on this issue. Indeed, we find that both of aggregate demand factor and aggregate supply factor (constructed as CP factor) are priced by the cross-section of stock returns. In particular, portfolios with highest exposure to aggregate supply factor outperform portfolios with lowest exposure to supply factor 1.8% per month while portfolios with lowest exposure to aggregate demand factor outperform portfolios with highest exposure to demand factor 2.1% per month.

We further build a VAR including supply factor, demand factor, excess bond return, and CP factor to implement variance decomposition analysis. It is shown that most of the variation in excess bond return and yields are not due to contemporaneous demand factor and supply factor, but demand factor can explain more share of variation in excess bond returns than supply factor. The finding of VAR is also supportive of unspanned factor conjecture in existing literature including Duffee (2011),Joslin, Priebsch, and Singleton (2014). Unspanned factors are those factors which are unspanned by contemporaneous yield curve but have predictive power for excess bond returns in the future. This is puzzling because there is no consensus on why there exist unspanned factors and which kind of variables are unspanned factors for bond. Following this direction, the second question this paper tries to address is whether predictive demand factor or supply factor is unspanned and how these unspanned factor affect market price of risk of spanned factors. We adopt Adrian, Crump, and Moench(2013)'s three step linear regressions to test if demand and supply factors are unspanned and to estimate the effects of unspanned factors on the market price of spanned factors. The results show that both demand and supply factors are unspanned and they have significant effects on market price of risk. Moreover, MTSMs enable us to decompose bond yields into expectation component and yield risk premium component and we find MTSMs without unspanned factors under-estimate yield risk premium.

The rest of this paper is organized as follows. In the next section, we briefly review related literature discussed above. Section 2 lays out asymptotic principal components analysis to estimate demand and supply factors. We also present the results of one-year-ahead predictive regressions for excess bond returns and discuss the economic interpretation of significant demand and supply factors. Section 3 is alternative machine learning of bond demand factors and supply factors to identify the most important economic variables based on LASSO. Section 4 is portfolio sort based on stocks' exposure to demand and supply factors. Section 5 describe the VAR model with impulse response function and variance decomposition analysis for supply factor, demand factor, excess bond return, and bond yields. Section 6 follows Adrian, Crump, and Moench(2013)'s three step linear regressions to estimate market price of risk of affine term structure model and decompose bond yield into two different components. Section 7 concludes.

A. Related literature

This paper is related to Ludvigson and Ng (2009), which uses static and dynamic factor model to examine whether macroeconomic variables could predict bond excess returns. They construct a monthly panel of 132 measures of economic activity and extract eight latent common factors out of those 132 economic variables using two methods:Asymptotic PCA and Gibbs-MCMC. They find these latent factors could predict bond excess returns in the following twelve months beyond CP factors both in sample and out of sample. This paper differs from their paper in the sense that their macroeconomic variables are only reflecting the demand of bond while this paper constructs economic variables covering both demand and supply of bond and compares the predictability of bond excess returns using demand factors and/or supply factors. So, this paper is also related to existing literature talking about supply factors of bond excess returns. Greenwood, Hanson, and Stein (2015) finds that supply of T-bills affects "z-spread" of yield curve, which reflects a money-like premium on short-term T-bills, above and beyond the liquidity and safety premia embedded in longer term Treasury yields. Greenwood and Vayanos (2014) explores how the supply and maturity structure of U.S. debt affect U.S. Treasury bond yields and expected returns. They construct the duration of government debt as the maturity-weighted-debt-to-GDP ratio and find that this ratio is positively associated with bond yields and can predict future bond returns. Moreover, they show that these effects are stronger for longer-maturity bonds and following periods when arbitrageurs have lost money. Those literature on supply factors talk about one ore two ad hoc explanatory variables in each paper and never link factors to each other. In this paper, we combine all these supply factors to construct the latent common supply factor and examine its predictive power for bond excess returns.

B. PCA Analysis of Economic Variables

B.1. Implementation and Data

I study monthly data spanning the period 1996:12015:12, a sample newer than the one in Ludvigson and Ng (2008). Observations on one- through five-year zero-coupon U.S. Treasury bond prices are taken from the Fama-Bliss dataset available from the Center for Research in Securities Prices (CRSP). Excess bond returns, yields, and forward rates are constructed from this dataset. I estimate factors from a balanced panel of 133 macroeconomic time series, each spanning the period 1996:1 to 2015:12. Most of series are coming from large macroeconomic database called FRED-MD on Michael W. McCracken's website. The remaining series are from Global Insight. The series were selected to represent broad categories of macroeconomic time series including both demand and supply sides of economy: output and income, labor market, consumption and orders, order and inventories, international trade and foreign exchange measures, prices, stock market indicators, money and credit, treasury bond outstanding, fiscal deficit, UP public debt,and bond duration. The more detailed description of economic series is given in the Appendix, where the coding column indicates how the original data were transformed to stationary series. All of the raw data have been standardized prior to estimation.

The econometric framework in this paper involves estimating common factors from a big dataset of economic activity. This estimation is carried out via asymptotic principle components analysis, a procedure that has been used for forecasting bond excess returns in Ludvigson and Ng (2008). Principal component analysis (PCA) is based on the (NN) sample covariance matrix $\hat{\Omega}_N = \frac{1}{T}X^{'}X$ where X is $N \times T$ matrix of observed variables.Asymptotic principal component analysis is based on the (TT) covariance matrix $\underbrace{\hat{\Omega}_T}_{T\times T} = \frac{1}{N}XX^{'}$. My notation for bond returns and yields closely follows that in Ludvigson and Ng (2008) and Cochrane (2005). I refer the reader to those papers for a more detailed description of the econometric method. The implementation involves an estimation of the following equation:

$$x_{i,t} = \lambda_i f_t + e_{i,t} \tag{14}$$

$$rx^{(n)}_{t+1} = \alpha^{'} F_t + \beta^{'} Z_t + \epsilon_{t+1} \tag{15}$$

where $x_{i,t}, i = 1, \cdots N, t = 1, \cdots, T$ are $T \times N$ panel of macroeconomic data , where the cross-sectional dimension N is large, and possibly larger than the number of time periods, T. f_t is $r \times 1$ latent common factors , λ_i is a corresponding $r \times 1$ vector of factor loadings, and $e_{i,t}$ are idiosyncratic errors. $rx^{(n)}_{t+1}$ denotes the continuously compounded (log) excess return on an n-year discount bond in period t + 1. Excess returns are defined as the difference between the log holding period return from buying an n-year bond at time t and selling it as an n 1 year bond at time t + 1, and the log yield on the one-year bond. Finally, $F_t \subseteq f_t$, and Z_t are control variables such as CP factors.

I pick the number of latent common factors according to the panel information criteria developed in Bai and Ng (2002). As common factors are hidden factors and not observed , I estimate them via asymptotic PCA, which is optimally minimizing the sum of squared residuals $x_t - \Lambda f_t$ to estimate time t factors $\hat{f}_t$ .In this paper, hats denotes estimated values.

B.2. Asymptotic Principle Component Analysis

Table 1 Panel A shows the OLS predictive regression results for two year bond excess returns. In this newer sample, CP factor itself only explains 4.1% of the variation of bond excess returns. Back to Cochrane and Piazzesi (2005), CP factor could explain a much bigger portion of the variation of bond excess returns, which is around 31%. So the predictability of CP factor decreases a lot as times goes by and there might be a new economic regime after 1996. Adding seven demand factors into regression results in an increase in adjusted R-square of 7.73%, showing an important role of demand factor in predicting bond excess returns. Especially, the first demand factor has a significant negative effect on bond excess returns and 1 standard deviation increase in this factor would dampen bond excess return by 33 basis point. Another interesting observation is that CP factor is not statistically significant after adding the demand factors. So macro factors might be more useful in predicting bond risk premia than CP factors. In the last specification, supply factors are included and adjusted R-square increases further to 17.26%. Supply factors are at least as important as demand factors in predicting bond risk premia. I then follow Ludvigson and Ng (2009) to construct single demand factor and single supply factor as the fitted values from a regression of average (across four different maturities) excess returns on the set of demand and supply factors, respectively. I denote them as F_t^{Demand} and F_t^{Supply}. Panel A shows that both of them have economic and statistic significance in affecting excess bond returns. A 1 standard deviation increase in demand factor and/or supply factor would increase excess bond return by 36 basis point and 45 basis point respectively. Panel B, Panel C, and Panel D shows the predicative regression results for three year excess bond returns, four year excess bond returns, and five year excess bond returns respectively. For each excess bond return, the regressions show very similar qualitative implications. One important interesting observation here is that the more maturity bond is, the predictability of both demand and supply factors are. This might be because economic variables take effect in a medium run instead of in a short run. Another explanation might be that economic factors take time to affect excess bond returns and therefore a lagged effect exist here. Also, this fact is consistent with unspanned factor assumption: factors which are not spanned by contemporaneous bond yields could predict future bond returns. I will talk more about this later.

Table XIII Regression of monthly excess bond returns on lagged factors

The table reports estimates from OLS regressions of excess bond returns on the lagged variables named in column 1. The dependent variable $rx^n_{t+1}$is the excess log return on the n-year Treasury bond. $\hat{F}_t$ denotes a set of regressors. These denote factors estimated by the method of asymptotic principal components using a panel of data with 133 individual series over the period 1996:12014:12. F_{Supply} is the single factor constructed as a linear combination of the four estimated supply factors .F_{Demand} , is the single factor constructed as a linear combination of the seven estimated demand factors. CP_t is the Cochrane and Piazzesi (2005) factor that is a linear combination of five forward spreads. Newey and West (1987) corrected t-statistics have lag order 18 months and are reported in parentheses. Coefficients that are statistically significant at the 5% or better level are highlighted in bold. A constant is always included in the regression even though its estimate is not reported in the table. The sample spans the period 1996:12014:12.

	$\widehat{F}^D_{1t}$	$\widehat{F}^D_{2t}$	$\widehat{F}^D_{3t}$	$\widehat{F}^D_{4t}$	$\widehat{F}^D_{5t}$	$\widehat{F}^D_{6t}$	$\widehat{F}^D_{7t}$	$\widehat{F}^S_{1t}$	$\widehat{F}^S_{2t}$	$\widehat{F}^S_{3t}$	$\widehat{F}^S_{4t}$	CP_t	F^{Demand}_t	F^{Supply}_t	adj. R^2
							Panel A: $rx^{(2)}_{t+1} = \beta_0 + \beta_1'\widehat{F}_t + \beta_2 CP_t + \epsilon_{t+1}$								
a												0.18			4.1%
												(1.07)			
b	**−0.33**	0.09	-0.06	0.003	-0.08	0.008	-0.09					0.21			11.83%
	(-2.09)	(1.73)	(-0.33)	(0.03)	(-0.99)	(0.05)	(-0.60)					(1.08)			
c	**−0.47**	**0.16**	-0.15	0.01	**−0.16**	-0.04	-0.11	**0.29**	-0.003	0.12	**−0.17**	0.21			17.26%
	(-3.05)	(2.83)	(-0.97)	(0.09)	(-1.99)	(-0.31)	(-0.76)	(3.71)	(-0.07)	(1.62)	(-2.27)	(1.11)			
d												0.16	**0.23**		9.86%
												(0.93)	(2.29)		
e												0.13	**0.36**	**0.45**	14.37%
												(0.76)	(3.17)	(2.30)	
							Panel B: $rx^{(3)}_{t+1} = \beta_0 + \beta_1'\widehat{F}_t + \beta_2 CP_t + \epsilon_{t+1}$								
a												0.42			5.91%
												(1.24)			
b	**−0.57**	**0.19**	-0.25	0.05	-0.23	-0.08	-0.08					0.45			13.91%
	(-2.05)	(1.9284)	(-0.79)	(0.24)	(-1.56)	(-0.31)	(-0.3)					(1.17)			13.91%
c	**−0.81**	**0.30**	-0.40	0.06	**−0.36**	-0.16	-0.1	**0.48**	-0.007	**0.23**	**−0.28**	0.45			17.9%
	(-2.92)	(2.82)	(-1.44)	(0.29)	(-2.53)	(-0.67)	(-0.37)	(3.65)	(-0.09)	(1.78)	(-2.13)	(1.18)			
d												0.36	**0.53**		14.61%
												(1.08)	(2.76)		

	$\widehat{F}^D_{1t}$	$\widehat{F}^D_{2t}$	$\widehat{F}^D_{3t}$	$\widehat{F}^D_{4t}$	$\widehat{F}^D_{5t}$	$\widehat{F}^D_{6t}$	$\widehat{F}^D_{7t}$	$\widehat{F}^S_{1t}$	$\widehat{F}^S_{2t}$	$\widehat{F}^S_{3t}$	$\widehat{F}^S_{4t}$	CP_t	F_t^{Demand}	F_t^{Supply}	adj. R^2
Panel C: $rx^{(4)}_{t+1} = \beta_0 + \beta_1'\widehat{F}_t + \beta_2 CP_t + \epsilon_{t+1}$															
a												0.71			8.6%
												(1.6)			
b	-0.57	**0.3**	-0.66	0.04	-0.36	-0.26	-0.03					0.77			17.66%
	(-1.51)	(2.13)	(-1.59)	(0.16)	(-1.76)	(0.72)	(-0.06)					(1.41)			
c	**−0.87**	**0.44**	**−0.85**	0.06	**−0.51**	-0.36	0.01	**0.58**	0.01	**0.32**	**−0.39**	0.76			20.73%
	(-2.34)	(2.89)	(-2.28)	(0.19)	(-2.7)	(-1.05)	(0.02)	(3.38)	(0.08)	(1.95)	(-2.24)	(1.42)			
d												0.62	**0.79**		18.91%
												(1.43)	(2.73)		
e												0.55	**1.09**	**1.05**	22.32%
												(1.29)	(3.86)	(2.99)	
Panel D: $rx^{(5)}_{t+1} = \beta_0 + \beta_1'\widehat{F}_t + \beta_2 CP_t + \epsilon_{t+1}$															
a												0.91			8.94%
												(1.82)			
b	-0.59	**0.46**	**−1.11**	-0.01	**−0.54**	-0.45	0.20					1.02			22.6%
	(-1.44)	(2.63)	(-2.24)	(-0.03)	(-2.18)	(-1.03)	(0.44)					(1.57)			
c	**−0.91**	**0.60**	**−1.32**	0.01	**−0.70**	-0.56	-0.19	**0.64**	-0.01	**0.38**	**−0.38**	1			24.58%
	(-2.2)	(3.2)	(-2.9)	(0.03)	(-2.96)	(-1.34)	(0.42)	(3.25)	(-0.05)	(2.02)	(-1.98)	(1.56)			
d												0.78	**1.12**		21.76%
												(1.63)	(2.85)		

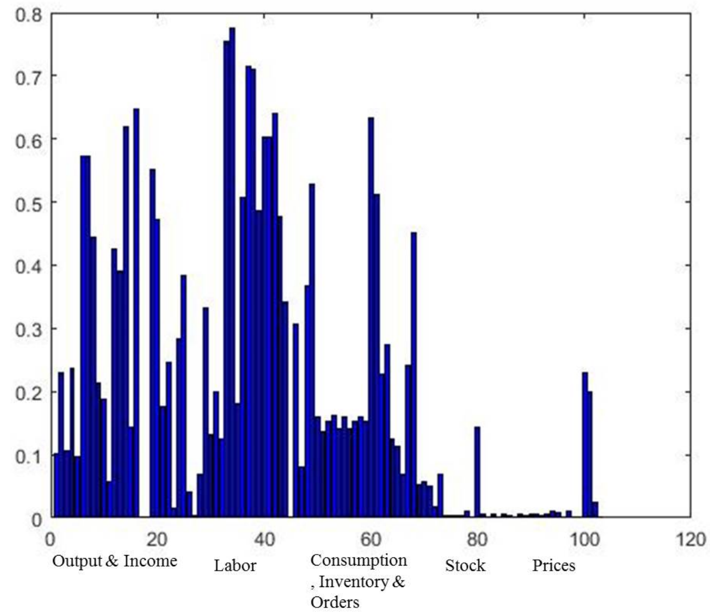


Figure 8. Marginal R-squares for $\widehat{F^D_{1t}}$.
Note: Chart shows the R-square from regressing the series number given on the x-axis onto $\widehat{F^D_{1t}}$. See the Appendix for a description of the numbered series. The factors are estimated using data from 1996:1 to 2015:12.

B.3. Economic interpretation of the factors

I also examine the economic interpretation of all estimated factors. The first demand factor, the most important demand factor out of all the estimated factors I study, is highly correlated with output and labor market but not highly correlated with prices or financial market. The second demand factor highly correlated with measures of the aggregate price-level and the fifth factor highly correlated with consumption and inventory orders also have significant predictive power for excess bond returns. On the other hand, the first supply factor highly correlated with public debt and the second supply factor highly correlated with money supply have substantial predictive power for excess bond returns. The details are shown in following graphs from figure 1 to figure 6.

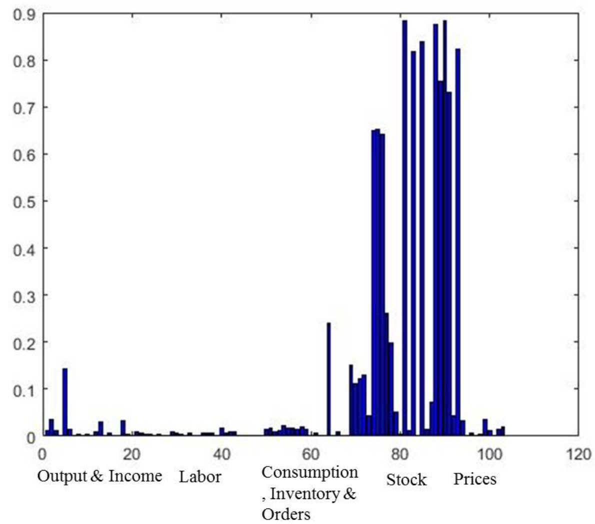


Figure 9. Marginal R-squares for $\widehat{F_{2t}^D}$.

Note: Chart shows the R-square from regressing the series number given on the x-axis onto $\widehat{F_{2t}^D}$. See the Appendix for a description of the numbered series. The factors are estimated using data from 1996:1 to 2015:12.

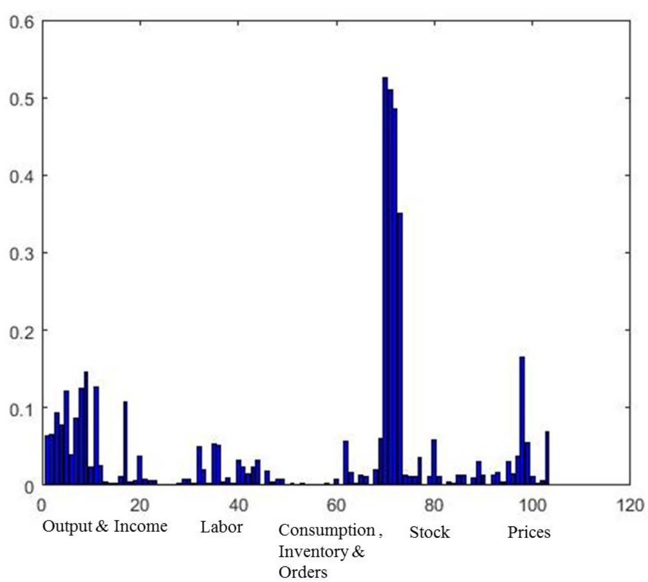


Figure 10. Marginal R-squares for $\widehat{F_{5t}^D}$.

Note: Chart shows the R-square from regressing the series number given on the x-axis onto $\widehat{F_{5t}^D}$. See the Appendix for a description of the numbered series. The factors are estimated using data from 1996:1 to 2015:12.

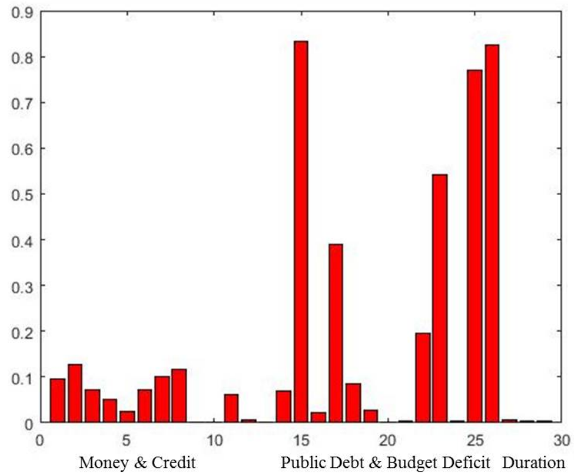


Figure 11. Marginal R-squares for $\widehat{F_{1t}^S}$.

Note: Chart shows the R-square from regressing the series number given on the x-axis onto $\widehat{F_{1t}^S}$. See the Appendix for a description of the numbered series. The factors are estimated using data from 1996:1 to 2015:12.

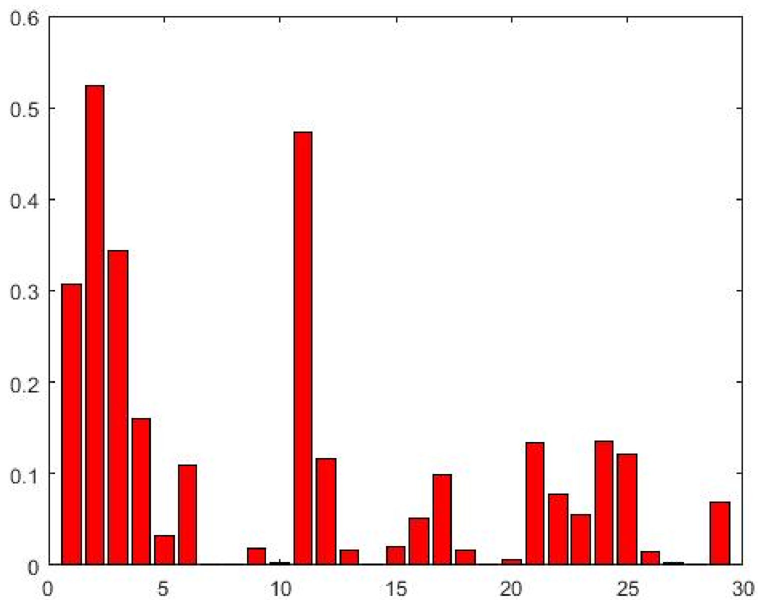


Figure 12. Marginal R-squares for $\widehat{F_{2t}^S}$.

Note: Chart shows the R-square from regressing the series number given on the x-axis onto $\widehat{F_{2t}^S}$. See the Appendix for a description of the numbered series. The factors are estimated using data from 1996:1 to 2015:12.

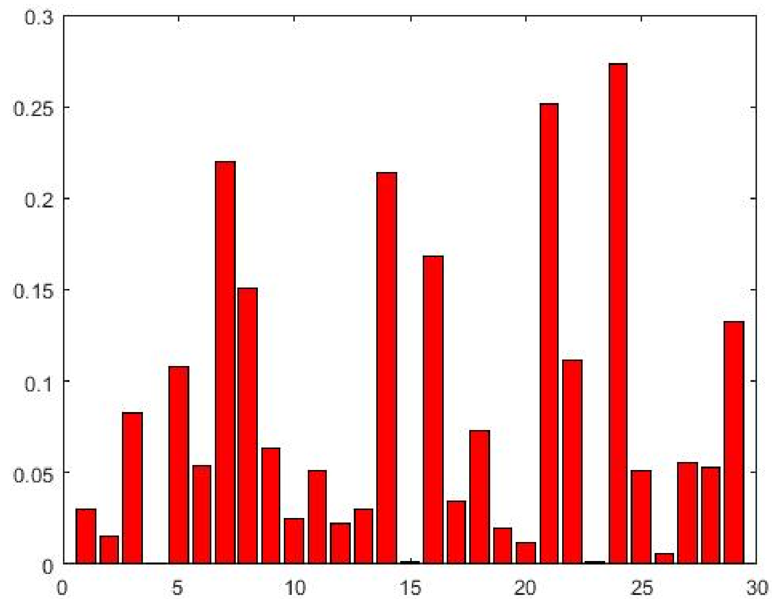


Figure 13. Marginal R-squares for $\widehat{F_{4t}^S}$.
Note: Chart shows the R-square from regressing the series number given on the x-axis onto $\widehat{F_{4t}^S}$. See the Appendix for a description of the numbered series. The factors are estimated using data from 1996:1 to 2015:12.

C. *Machine Learning of Important Economic Variables*

Although we briefly characterize the latent factors by relating them to the underlying variables in our panel dataset, the factors are identifiable only up to an r r matrix and a detailed interpretation of the individual factors would be inappropriate. Moreover, we caution that any labeling of the factors is imperfect, because each is influenced to some degree by all the variables in our large dataset and the orthogonalization means that no one of them will correspond exactly to a precise economic concept like output or unemployment, which are naturally correlated. To avoid this issue and give more economic interpretation about bond predictability, we unitize machine learning technique LASSO to identify the underlying meaningful economic variable for predicting bond risk permia. We first use LASSO to run similar predictive regression to see whether the same 133 macro economic variables can still generate bond predictability under new methodology. Simultaneously, LASSO will automatically pick the most important variables in predicting bond excess returns and we call them as key economic variables. Then, we examine the pricing power of those key economic variables for cross-section of stock returns.

C.1. LASSO Identification of Key Economic Variables

Table XIV shows both the in-sample and out-of-sample fit of projecting bond excess returns onto 133 macro economic variables based on LASSO. The results are generally very good and much better than PCA regression. In training period from March 1996 to March 2009, the in-sample R^2 of projecting the excess return on the 2-year treasury bond rx^2_{t+1} onto macro economic variables is 20% and is much better than the same R^2 in table XIII. The in-sample R^2 of regressing the excess return on the 5-year treasury bond rx^5_{t+1} is even reaching up to 40%. We also examine the out-of-sample performance of training, and the R^2 for rx^2_{t+1}, rx^3_{t+1}, rx^4_{t+1}, and rx^5_{t+1} are respectively 2%, 14%, 26%, and 34%, which would reduce a lot of over-fitting concerns about our

LASSO model. Similar strong predictability results hold for alternative machine learning technique Support Vector Machine, as shown in panel B of table XIV. Figures 14 and 15 show the fitness of LASSO and SVM graphically. More importantly, LASSO can automatically penalize unimportant variables and shrink coefficients on those unimportant variables to zero. Only important variables are left with nonzero coefficients, as shown in table XV. Now, instead of having to look at 133 macro economic variables, we only need to focus on those 15 important variables to predict bond excess returns. Those 15 economic variables can also be divided into two groups: bond demand group and bond supply group. In particular, Fiscal Deficit, US Total Public Debt Outstanding Notes, and US Total Public Debt Outstanding belong to supply group and the other variables belong to demand group. These 15 observed variables have their own economic meanings and can give us direct interpretations. For example , IPDMAT represents industry production of durable materials and IPNMAT is industry production of nondurable materials. Out of 13 industrial production index, only those two index are important for predicting bond returns and should be classified as one of demand variables. Another striking feature here is that labor market is the most important component of our demand factor. Seven out of ten demand variables are coming from labor market variables: UEMP27OV, CES1021000001, USCONS, SRVPRD, USWTRADE, USFIRE, and CES0600000007.

Table XIV: In-sample and Out-of-sample Bond Risk Premia prediction using macro-economic Variables based on LASSO and SVM

Reported are in-sample and out-of-sample LASSO fit statistics using both the train subsample and test subsample. LASSO is specified as $ret_t = \beta_0 + \sum_{i=1...131} \beta_{1,i} X_{i,t} + \epsilon_{t+1}$, where $X_{1,t}, \ldots, X_{131,t}$ are 131 macro-economic variables. Panel A reports variance of the predicted value as a fraction of actual bond risk premia's variance using in-sample data from 1996.03 to 2009.03. Panel B reports out-of-sample variance of the predicted value as a fraction of original risk premia' variance using out-of-sample data from 2009.04 to 2014.12.

Panel A Lasso:	$rx_{t+1} = \beta_0 + \sum_{i=1...131} \beta_{1,i} X_{i,t} + \epsilon_{t+1}$
Bond Risk Premia:	In sample R^2
$rx^{(2)}_{t+1}$	0.2
$rx^{(3)}_{t+1}$	0.31
$rx^{(4)}_{t+1}$	0.35
$rx^{(5)}_{t+1}$	0.40
Bond Risk Premia:	Out of sample R^2
$rx^{(2)}_{t+1}$	0.02
$rx^{(3)}_{t+1}$	0.14
$rx^{(4)}_{t+1}$	0.26
$rx^{(5)}_{t+1}$	0.34
Panel B SVM:	$rx_{t+1} = \beta_0 + \sum_{i=1...131} \beta_{1,i} X_{i,t} + \epsilon_{t+1}$
Bond Risk Premia:	In sample R^2
$rx^{(2)}_{t+1}$	0.16
$rx^{(3)}_{t+1}$	0.15
$rx^{(4)}_{t+1}$	0.27
$rx^{(5)}_{t+1}$	0.32
Bond Risk Premia:	Out of sample R^2
$rx^{(2)}_{t+1}$	0.01
$rx^{(3)}_{t+1}$	0.11
$rx^{(4)}_{t+1}$	0.11
$rx^{(5)}_{t+1}$	0.21

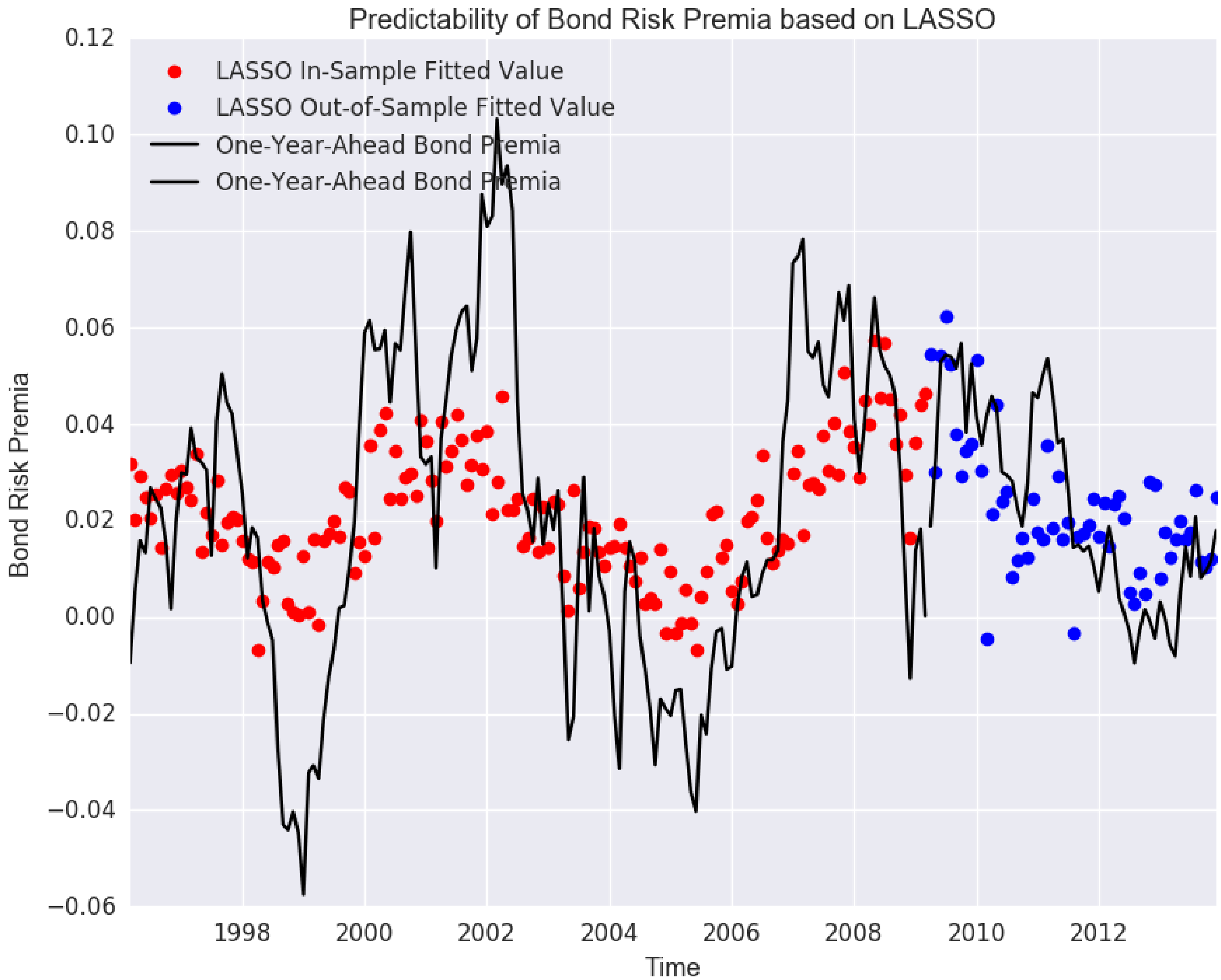


Figure 14. Predictability of Bond Risk Premia based on LASSO .
Note: Bond Risk Premia from 1960.3 to 2014.12. Solid line is one-year-ahead bond risk premia. Dots are LASSO Factor,fitted value from regression $ret_{1,t} = \beta_0 + \sum_{i=1...131} \beta_{1,i} X_{i,t} + \epsilon_{t+1}$, where $x_{i,t}$ are vector representation of end-of-month demand and supply macro variables and β is estimated with LASSO. The train subsample, 1996.3 to 2009.3, is used to estimate the dependency between bond risk premia and real macro-economic variables. The test subsample, 2009.4-2014.12, is used for out-of-sample tests of model fit.

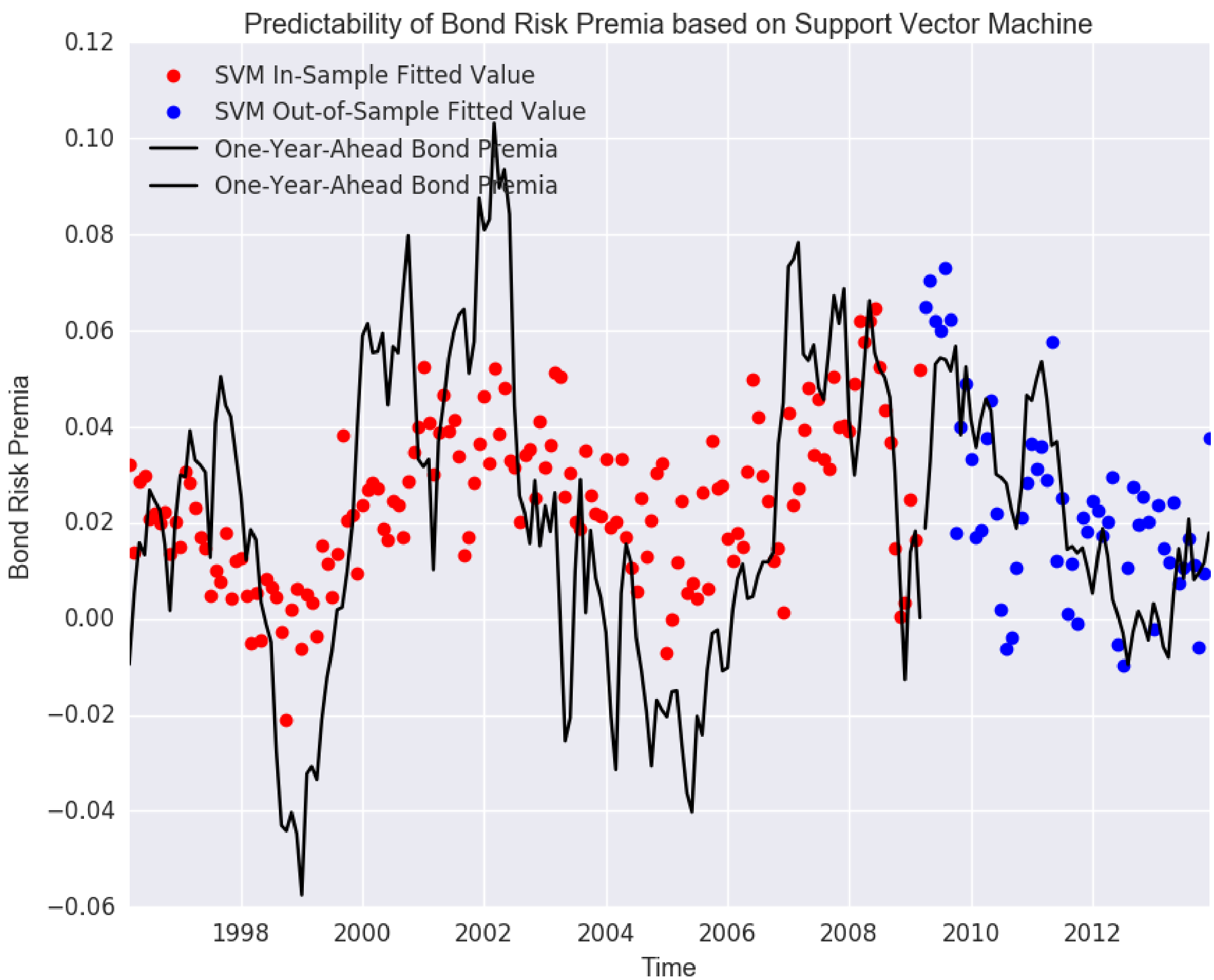


Figure 15. Predictability of Bond Risk Premia based on Support Vector Machine . Note: Bond Risk Premia from 1960.3 to 2014.12. Solid line is one-year-ahead bond risk premia. Dots are SVM Factor,fitted value from regression $ret_{1,t} = \beta_0 + \sum_{i=1...131} \beta_{1,i} X_{i,t} + \epsilon_{t+1}$, where $x_{i,t}$ are vector representation of end-of-month demand and supply macro variables and β is estimated with Support Vector Machine. The train subsample, 1996.3 to 2009.3, is used to estimate the dependency between bond risk premia and real macro-economic variables. The test subsample, 2009.4-2014.12, is used for out-of-sample tests of model fit.

Table XV: Key Macro Economic Variables Identified by LASSO

Reported are in-sample LASSO estimated non-zero coefficients on key macro economic variables. LASSO is specified as $rx_{t+1} = \beta_0 + \sum_{i=1...131} \beta_{1,i} X_{i,t} + \epsilon_{t+1}$, where $X_{1,t}, \dots, X_{131,t}$ are 131 macro-economic variables. Different columns correspond to different specifications of dependent variables.

Name	$rx^{(3)}_{t+1}$	$rx^{(4)}_{t+1}$	$rx^{(5)}_{t+1}$
IPDMAT	0.0	0.102	0.131
IPNMAT	0.0	0.049	0.159
UEMP27OV	0.023	0.049	0.062
CES1021000001	0.392	0.708	1.115
USCONS	-0.543	-0.931	-1.239
SRVPRD	0.0	-1.089	-2.042
USWTRADE	-1.568	-1.567	-1.915
USFIRE	0.0	-0.634	-1.576
CES0600000007	0.004	0.012	0.016
S&P_INDUST	-0.009	-0.013	-0.028
M2REAL	-0.253	-0.543	-0.476
TOTRESNS	0.002	0.019	0.017
FISCAL DEFICIT	0.0	-0.001	-0.004
US TOTAL PUBLIC DEBT OUTSTANDING NOTES	-0.16	-0.174	-0.202
US TOTAL PUBLIC DEBT OUTSTANDING	0.0	-0.011	-0.028

D. Stock Portfolio Sort Results based on Exposure to Demand and Supply Factors

As informative state variables for bond risk premium, both bond supply factor and demand factor should be closely related to stochastic discount factor, which is devised to price all asset returns in theory. So, it is interesting to see whether our supply factor and demand factor can also be priced in stocks. This section utilize portfolio sort method to test the pricing power of both supply factor and demand in cross-section of stocks.

Table XVI reports factor regression results across Supply factor exposure $\beta^{Supply}_{i,t}$, where $\beta^{Supply}_{i,t}$ is firm i's exposure to aggregate bond supply factor in month t. Decile portfolios are formed at the conclusion of each month, ranging from 1 to 10 with the highest (lowest) values located in the 10th (1st) decile. Equal-Weighted portfolio returns are measured in next month and regressed on contemporaneous risk factors: the three Fama-French Factors and momentum factors(UMD).The intercept in this regression(Intercept) is the portfolio alpha. Our main result is that intercepts from those regressions increase with β^{Supply}, indicating high β^{Supply} outperform low β^{Supply} firms. The row "1-10" contains a statistics test for the difference of low and high β^{Supply} decile portfolios and shows that the 1.3% difference in the four-factor alphas are statistically significant (t-statistics = 4.95). The final row "(1+2) - (9+10)" at the bottom of the table contains another statistic test for the difference of low and high β^{Supply} quintile portfolios. A strategy that goes long the quintile portfolio with the highest β^{Supply} and short the quintile portfolio with the lowest β^{Supply} yields a Fama–French–Carhart alpha of 1.8% per month (t-statistics = 4.66). Sharpe ratio for this strategy is almost 1. Analogous significant results hold for equal-weighted portfolio sort based on exposure to bond demand factor in table XVII.

Table XVI: Equity Portfolio Sort Based on Supply Factor Exposure

Reported are factor regression results across supply factor Exposure $\beta_{i,t}^{Supply}$, where $\beta_{i,t}^{Supply}$ is firm i's exposure to Supply Factor in month t. Decile portfolios are formed at the conclusion of each month, ranging from 1 to 10 with the highest (lowest) values located in the 10th (1st) decile. The sample consists of 814,997 firm-months spanning 1996 through 2015. Equal-Weighted portfolio returns are measured in next month and regressed on contemporaneous risk factors: the three Fama-French Factors and momentum factors(UMD).The intercept in this regression(Intercept) is the portfolio alpha. t-statistics are shown in the underlying row. $*$, $**$, and $***$ indicate 10%, 5%, and 1% significance levels, respectively.

	Intercept	mktrf	SMB	HML	UMD
1	0.006***	0.927***	0.972***	0.199***	-0.104***
	3.28	21.41	14.28	3.65	-3.12
2	0.005***	0.812***	0.77***	0.329***	-0.073***
	3.76	25.88	15.63	8.34	-3.01
3	0.005***	0.789***	0.542***	0.34***	-0.039*
	4.71	30.26	13.22	10.38	-1.94
4	0.005***	0.772***	0.526***	0.347***	-0.037**
	4.9	33.48	14.52	11.97	-2.09
5	0.005***	0.801***	0.503***	0.398***	-0.059***
	5.4	35.26	14.09	13.94	-3.4
6	0.006***	0.78***	0.588***	0.366***	-0.106***
	6.44	33.81	16.21	12.63	-6.01
7	0.006***	0.81***	0.651***	0.406***	-0.129***
	6.67	35.35	18.08	14.09	-7.32
8	0.007***	0.912***	0.687***	0.329***	-0.168***
	6.66	32.84	15.73	9.43	-7.87
9	0.01***	0.899***	0.86***	0.214***	-0.255***
	7.55	26.64	16.2	5.03	-9.82
10	0.017***	1.001***	1.021***	-0.012	-0.41***
	7.83	18.49	11.99	-0.18	-9.86
1-10	-0.013***	-0.07	-0.05	0.202**	0.306***
	-4.95	-1.08	-0.5	2.49	6.18
(1+2)-(9+10)	-0.018***	-0.158	-0.139	0.317**	0.488***
	-4.65	-1.61	-0.9	2.57	6.48

Table XVII: Equity Portfolio Sort Based on Demand Factor Exposure

Reported are factor regression results across demand factor Exposure $\beta_{i,t}^{Demand}$, where $\beta_{i,t}^{Demand}$ is firm i's exposure to Demand Factor in month t. Decile portfolios are formed at the conclusion of each month, ranging from 1 to 10 with the highest (lowest) values located in the 10th (1st) decile. The sample consists of 814,997 firm-months spanning 1996 through 2015. Equal-Weighted portfolio returns are measured in next month and regressed on contemporaneous risk factors: the three Fama-French Factors and momentum factors(UMD).The intercept in this regression(Intercept) is the portfolio alpha. t-statistics are shown in the underlying row. $*$, $**$, and $***$ indicate 10%, 5%, and 1% significance levels, respectively.

	Intercept	mktrf	SMB	HML	UMD
1	0.019***	0.954***	1.298***	-0.079	-0.672***
	6.42	12.86	11.14	-0.85	-11.8
2	0.011***	0.99***	0.962***	0.138**	-0.35***
	6.06	23.21	14.35	2.57	-10.7
3	0.009***	0.902***	0.826***	0.278***	-0.269***
	6.49	27.47	16.01	6.72	-10.7
4	0.007***	0.902***	0.679***	0.391***	-0.148***
	6.44	33.96	16.27	11.71	-7.27
5	0.006***	0.837***	0.6***	0.388***	-0.089***
	6.33	34.36	15.68	12.68	-4.74
6	0.005***	0.794***	0.556***	0.417***	-0.026
	5.47	34.04	15.15	14.21	-1.45
7	0.004***	0.733***	0.495***	0.389***	0.001
	4.42	29.79	12.79	12.58	0.04
8	0.003***	0.717***	0.499***	0.346***	0.028
	3.54	29.87	13.23	11.46	1.52
9	0.003***	0.761***	0.502***	0.375***	0.052**
	2.71	28.68	12.03	11.25	2.55
10	0.005***	0.914***	0.699***	0.271***	0.096***
	2.79	22.05	10.73	5.2	3.0
1-10	0.013***	0.045	0.599***	-0.359***	-0.768***
	3.86	0.53	4.55	-3.41	-11.9
(1+2)-(9+10)	0.021***	0.274**	1.059***	-0.596***	-1.171***
	4.05	2.17	5.35	-3.77	-12.1

E. VAR and Impulse Response Function

We estimate three monthly pth-order vector autoregression (VAR) models to examine impulse response functions of excess bond returns to demand ans supply shocks. We use Hamilton 's notation for VAR.

$$Y_{t+1/12} - \mu = \Phi_1(Y_t - \mu) + \Phi_2(Y_{t-1/12} - \mu) + \cdots + \Phi_p(Y_{t-\frac{p-1}{12}} - \mu) + \epsilon_{t+1/12} \tag{16}$$

The VAR (p) can be stacked in the first-order companion form to become a VAR(1):

$$\xi_{t+1/12} = A\xi_t + v_{t+1/12} \tag{17}$$

where $\xi_{t+1/12} = \begin{bmatrix} Y_t - \mu \\ Y_{t-1/12} - \mu \\ \vdots \\ Y_{t-\frac{p-1}{12}} - \mu \end{bmatrix}$

$$A = \begin{bmatrix} \Phi_1 & \Phi_2 & \Phi_3 & \cdots & \Phi_p \\ I_n & 0 & 0 & \cdots & 0 \\ 0 & I_n & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & I_n & 0 \end{bmatrix}$$

$$v_{t+1/12} = \begin{bmatrix} \epsilon_{t+1/12} \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

We usew Python Statsmodels to generate Orthogonalized Impulse Response Functions and Variance Decomposition Results for VAR(p). The first model is a 1-order VAR with five variables: supply factor, demand factor, CP factor, two-year-bond excess return, four-year-bond excess return. The second model is a 12-order VAR with the same variables. The third model is a 12-order VAR with different variables: supply factor, demand factor, CP factor, two-year-bond yield, and four-year-bond yield. We plot the monthly time series of annualized variables in model 1 from March 1996 to December 2014. It shows that all series are stationary and three explanatory factors are very different from each other. Excess bond returns are countercyclical. CP factors are also countercyclical except for financial crisis. Standardized Demand factor is also countercyclical after financial crisis but acyclical before that. In contrast, Supply factor is procyclical after financial crisis and acyclical before that. Two-year-bond excess return shows a similar time sires pattern as four-year-bond excess return. The IRF shows how each factor shocks affect four-year-bond excess return in sixty months. We find one-standard-deviation increase in supply shock result in a 10 basis point decrease in four-year-bond monthly excess return in month 3 and since then, the supply effect dissipates gradually in three years. In contrast, one-standard-deviation increase in demand shock result in a 30 basis point increase in four-year-bond monthly excess return in month 3 and since then the demand effect disappear gradually in three years. Both of effects are short to medium term effects rather than long term effects. Figure 16 is variance decomposition analysis for this VAR. It is surprising to see demand factor explains much more variations in bond excess returns than supply factor. One of potential reasons is that demand factor is extracted out of 80 macro variables while supply factor is constructed only from 30 variables. Figure 17 is VAR(12) with the same variables. This is just robustness check and IRF are still not statistically significant.

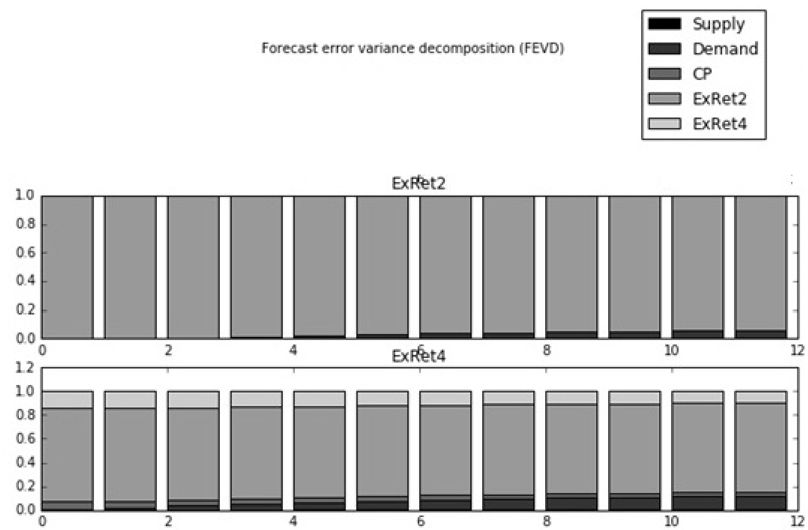


Figure 16. This figure plots the variance decomposition results for supply factor, demand factor, CP factor, two-year-bond excess return, four-year-bond excess return for a 1-order VAR.

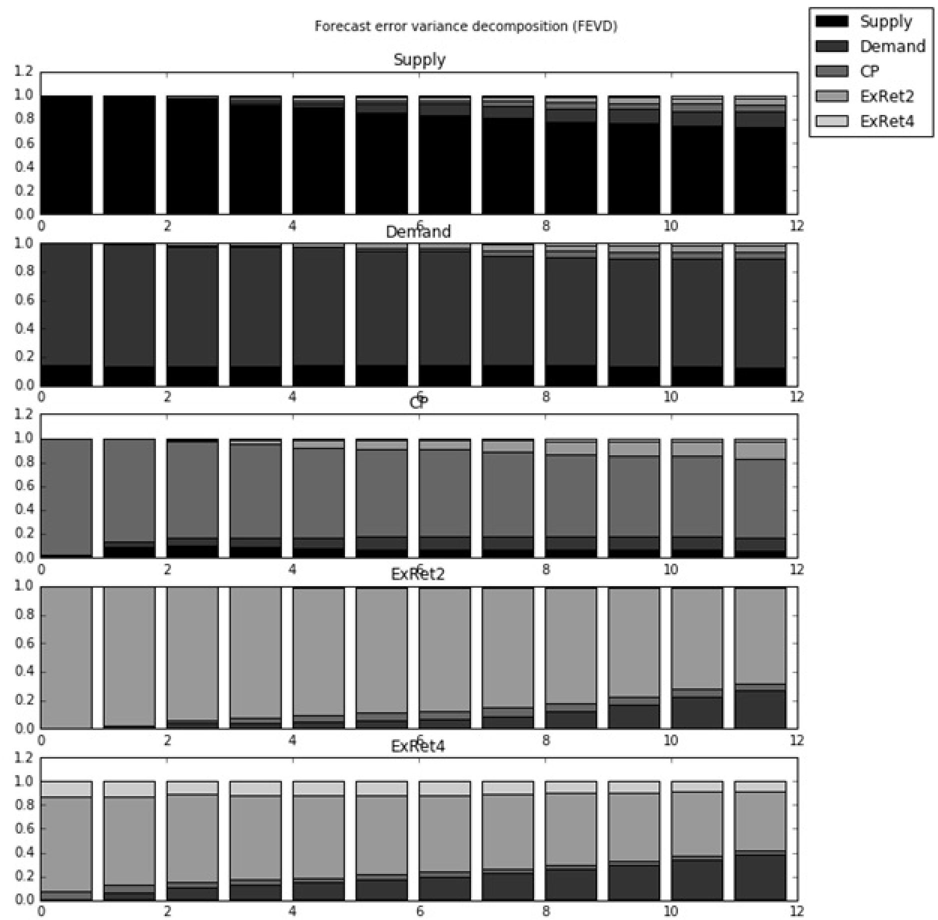


Figure 17. This figure plots the variance decomposition results for supply factor, demand factor, CP factor, two-year-bond excess return, four-year-bond excess return for a 12-order VAR.

F. Market Price of Risk on Bond Factors

From previous discussions, we find that both demand factors and supply factors are important for bond prediction no matter which method we employ, tradition PCA or machine learning LASSO/SVM. These findings are ruled out by unrestricted (and commonly employed) affine term structure models, where the forecastability of bond returns and bond yields is completely summarized by the cross-section of yields or forward rates. So, we want to extend term structure model to incorporate those macro factors and formally estimate market price of risk on those factors. Term structure model would also give us a decomposition of bond yield into expectation part and yield risk premium part.

F.1. General Setup for Continuous-time Macro Finance Affine Term Structure Model: MTSMs

Continuous Macro Finance Affine Term Structure Model (MTSMs) are typically constructed from the following three key ingredients:

First, we assume that the time-series process for state variables X under the risk-neutral pricing measure $\mathbb{Q}$, induced by numeraire price $P(t)$, which is the price of a continuously compounded risk free rate (short rate, or bank deposit), has the drift and volatility functions of the risk factors satisfying

$$\mu_X^{\mathbb{Q}}(t) = \kappa^{\mathbb{Q}}(\theta^{\mathbb{Q}} - X(t)), \tag{18}$$

where $\theta^{\mathbb{Q}}$ is an $N \times 1$ vector and $\kappa^{\mathbb{Q}}$

$$\sigma_X(t) = \Sigma\sqrt{S(t)} \tag{19}$$

where $S_{ii}(t) = \alpha_i + \beta_i X(t), S_{ij}(t) = 0, i \neq j, 1 \leq i, j \leq N$, and Σ is $N \times N$ matrix of constants.

Second, We assume the time-series process for X under the actual measure $\mathbb{P}$ follows a Markovian-state Process, with

$$dX(t) = \mu_X^{\mathbb{P}}(X,t)dt + \sigma_X(X,t)dB(t), \tag{20}$$

where $\mu_X^{\mathbb{P}}(X,t)$ is an $N \times 1$ vector of drifts under $\mathbb{P}$ and $\sigma_X(X,t)$ is an $N \times N$ state dependent factor-volatility matrix. In this diffusion setting, the pricing kernel M_t can be written generically as

$$\frac{dM_t}{M_t} = -r_t dt - \Lambda_t^{'} dB(t), \tag{21}$$

where $r_t = r(X(t),t)$ is the instantaneous riskless rate, $B(t)$ is a vector of N independent Brownian motions, and $\Lambda_t = \Lambda(X(t),t)$ is the N-vector of market prices of risk. For simplicity, We take the risk factors driving M_t and X_t to be one and the same.

Given $\sigma_X(t)$ satisfying Equation (15), the requirement of Equation (14) determines the drift of X under the actual measure, $\mu_X^{\mathbb{P}}(t)$, once the market prices of risk Λ_t are specified, and vice versa, because $\sigma_X(t)$ is the same for both measure $\mathbb{Q}$ and measure $\mathbb{P}$. ($\mathbb{Q}$-drift of $X(t)$ is $\mu_X^{\mathbb{Q}}(t) = \mu_X^{\mathbb{P}} - \sigma_X(t)\Lambda(t)$.) So, We follow Duffee (2002) and propose flexible "essentially affine" specification of market price of risk $\Lambda(t)$ that has the form

$$\Lambda(t) = \sqrt{S_t}\lambda_1 + \sqrt{S_t^-}\lambda_2 X(t) \tag{22}$$

where λ_1 is $N \times 1$ vector, and λ_2 is $N \times N$ vector, and $S_{ii}^{-}(t) = (\alpha_i + \beta_i' X_t)^{-1}$, if $\inf(\alpha_i + \beta_i' X_t)^{-1} \geq 0$

Third, instantaneous riskless rate or the short rate, r(t) is a function of state variables, X(t):

$$r(t) = \delta_0 + \delta_X' X(t) \tag{23}$$

For a fixed income security with a dividend rate $h(X(t), t)$ for $t \leq T$ and terminal payoff $g(X(T))$ at date T, its price at date$t \leq T$ can be expressed in terms of the pricing kernel as

$$P(X(t), t) = E_t^{\mathbb{P}} \left[\int_t^T \frac{M(s)}{M(t)} h(X(s), s) ds \right] + E_t^{\mathbb{P}} \left[\frac{M(T)}{M(t)} g(X(t), t) \right] \tag{24}$$

where E_t denotes expectation conditioned on date t information under actual measure $\mathbb{P}$.

The solution to Equation (19) is also the solution to the following fundamental partial differential equation (PDE) [e.g., Duffie (1996)],

$$\left[\frac{\partial}{\partial t} + \mathcal{A} \right] P_t - r_t P_t + h_t = 0 \tag{25}$$

where $\mathcal{A}$ is the infinitesimal generator

$$\mathcal{A} = (\mu_t^{\mathbb{P}} - \sigma_{X_t} \Lambda_t)' \frac{\partial}{\partial X_t} + \frac{1}{2} Trace \left[\sigma_{X_t} \sigma_{X_t}' \frac{\partial^2}{\partial X_t \partial X_t'} \right]. \tag{26}$$

The expected excess return from holding a fixed income security with price P is given by

Duffie and Kan (1996) show that under these three assumptions, the solution to the PDE of Equation (20) for U.S. government zero-coupon bond $D(t,T)$ is exponentially affine in underlying state variables:

$$D(t,T) = e^{\gamma_0(T-t)+\gamma_X(T-t)'X(t)} \tag{27}$$

where γ_0 and γ_X satisfy known ordinary differential equations (ODEs).

F.2. State Variables and Expected Returns

In this section, We discrete the above MTSMs and use exactly Adrian, Crump, and Moench (2013)'s three step linear regressions to estimate the term structure. We closely follow their notation and show their derivation for data generating process for arbitrage-free excess holding period returns. Based on their model, We can derive the closed form solution for the decomposition of bond yields into expectation component and yield risk premium. We put vector auto-regression (VAR(1)) on dynamics of a $K \times 1$ vector of state variables Xt:

$$X_{t+1} = \mu + \phi X_t + v_{t+1}. \tag{28}$$

Where, the shocks v_{t+1} follow a Conditional Gaussian distribution with covariance matrix Σ:

$$v_{t+1}|\{X_s\}_{s=0}^{t} \sim N(0,\Sigma), \tag{29}$$

where $\{X_s\}_{s=0}^{t}$ denotes the path of X_t. We denote $P_t^{(n)}$ the U.S. zero coupon Treasury bond price with maturity n at time t. No-arbitrage assumption implies that there exists a Stochastic Discount Factor M_t such that

$$P_t^{(n)} = E_t\Big[M_{t+1}P_{t+1}^{(n-1)}\Big], \tag{30}$$

$$M_{t+1} = exp\Big(-r_t - \frac{1}{2}\lambda_t'\lambda_t - \lambda_t'\Sigma^{-1/2}v_{t+1}\Big) \tag{31}$$

where $r_t = lnP_t^{(1)}$ is considered as continuously compounded risk-free rate. Market prices of risk are assumed to be of the essentially affine form:

$$\lambda_t = \Sigma^{-1/2}(\lambda_0 + \lambda_1 X_t). \tag{32}$$

We denote $rx_{t+1}^{(n-1)}$ the log excess holding return of a bond maturing in n periods:

$$rx_{t+1}^{(n-1)} = \ln P_{t+1}^{(n-1)} - \ln P_t^{(n)} - r_t. \tag{33}$$

Using Eq. (17) and Eq.(19) in Eq.(16) yields

$$1 = E_t\Big[exp\Big(rx_{t+1}^{(n-1)} - \frac{1}{2}\lambda_t'\lambda_t - \lambda_t'\Sigma^{-1/2}v_{t+1}\Big)\Big]. \tag{34}$$

Assuming that $\Big\{rx_{t+1}^{(n-1)}, v_{t+1}\Big\}$ are jointly normally distributed,

$$E_t\Big[rx_{t+1}^{(n-1)}\Big] = Cov_t\Big[rx_{t+1}^{(n-1)}, v_{t+1}^{'}\Sigma^{-1/2}\lambda_t\Big] - \frac{1}{2}Var_t\Big[rx_{t+1}^{(n-1)}\Big]. \quad (35)$$

We denote

$$\beta_t^{(n-1)'} = Cov_t\Big[rx_{t+1}^{(n-1)}, v_{t+1}^{'}\Big]\Sigma^{-1}, \quad (36)$$

and using Eq. (18) We obtain

$$E_t\Big[rx_{t+1}^{(n-1)}\Big] = \beta_t^{(n-1)'}(\lambda_0 + \lambda_1 X_t) - \frac{1}{2}Var_t\Big[rx_{t+1}^{(n-1)}\Big]. \quad (37)$$

Unexpected excess return can be decomposed into one component that is correlated with v_{t+1} and another component that is conditionally orthogonal. Then We find

$$rx_{t+1}^{(n-1)} - E_t\Big[rx_{t+1}^{(n-1)}\Big] = \gamma_t^{(n-1)'}v_{t+1} + e_{t+1}^{(n-1)}. \quad (38)$$

Wet is easy to see $\gamma_t^{(n-1)} = \beta_t^{(n-1)}$ based on Eq. (22). We assume that the return pricing errors $e_{t+1}^{(n-1)}$ are conditionally independently and identically distributed (i.i.d.) with variance σ^2. β is assumed to be constant.

The data generating process for log excess holding period returns is then

$$rx_{t+1}^{(n-1)} = \beta^{(n-1)'}(\lambda_0 + \lambda_1 X_t) - \frac{1}{2}(\beta^{(n-1)'}\Sigma\beta^{(n-1)} + \sigma^2) + \beta^{(n-1)'}v_{t+1} + e_{t+1}^{(n-1)} \tag{39}$$

Stacking this system across maturities and time periods, the matrix form looks like

$$rx = \beta^{'}(\lambda_0\iota_T^{'} + \lambda_1 X_{-}) - \frac{1}{2}(B^{\star}vec(\Sigma) + \sigma^2\iota_N)\iota_T^{'} + \beta^{'}V + E \tag{40}$$

where rx is a $N*T$ matrix of excess returns, $\beta = [\beta^{(1)}\beta^{(2)}\cdots\beta^{(N)}]$ is a $K*N$ matrix of factor loadings, ι_T and ι_N are a $T*1$ and $N*1$ vectors of ones, $X_{-} = [X_0 X_1 \cdots X_{T-1}]$ is a $K*T$ matrix of lagged pricing factors, $B^{\star} = [vec(\beta^{(1)}\beta^{(1)'})\cdots vec(\beta^{(N)}\beta^{(N)'})]$ is an $N*K^2$ matrix, V is a $K*T$ matrix, and E is an $N*T$ matrix.

F.3. Unspanned Factors Estimation

Unsnapped factor model assumes that a given factor does not affect bond yields under the pricing measure, which is equivalent to imposing the restriction that the corresponding elements of $\{B_n, n = 1, \cdots, N\}$ be exactly equal to zero.

Partition the factors into spanned factors X_t^s with nonzero risk exposures and unspanned factors X_t^u, which have zero risk exposures. The factors continue to follow a joint VAR process under the P measure

$$\begin{bmatrix} X_t^s \\ X_t^u \end{bmatrix} = \mu + \phi \begin{bmatrix} X_{t-1}^s \\ X_{t-1}^u \end{bmatrix} + \begin{bmatrix} \nu_t^s \\ \nu_t^u \end{bmatrix} \tag{41}$$

where X_t^s is of dimension $K_s * 1$, X_t^u is of dimension $K_u * 1$, and μ and ϕ are partitioned accordingly. The spanning restriction implies that the upper right $K_s * K_u$ block of the risk-neutral transition matrix $\Phi^\star = (\Phi - \lambda_1)$ is zero, i.e.

$$\Phi^\star = \begin{bmatrix} \Phi_{ss} - \lambda_1^{ss} & 0 \\ \Phi_{us} - \lambda_1^{us} & \Phi_{uu} - \lambda_1^{uu} \end{bmatrix} \tag{42}$$

With this restriction, the data generating process can be rewritten as

$$\begin{aligned} rx_{t+1}^{(n-1)} &= \beta^{(n-1)\prime}(\lambda_s^0 + \lambda_1 X_t + \nu_{t+1}) - \frac{1}{2}(\beta^{(n-1)\prime}\Sigma\beta^{(n-1)} + \sigma^2) + e_{t+1}^{(n-1)} && (43) \\ &= \beta^{(n-1)\prime}(\lambda_s^0 + \lambda_1 X_t + X_{t+1} - \Phi X_t - \mu) - \frac{1}{2}(\beta^{(n-1)\prime}\Sigma\beta^{(n-1)} + \sigma^2) + e_{t+1}^{(n-1)} && (44) \\ &= -\beta_s^{(n-1)\prime}(\mu_s^\star + \Phi_{ss}^\star X_t^s) - \frac{1}{2}(\beta_s^{(n-1)\prime}\Sigma_{ss}\beta_s^{(n-1)} + \sigma^2) + e_{t+1}^{(n-1)} && (45) \end{aligned}$$

where $\mu_s^\star$ denotes the upper $k_s * 1$ sub-vector of the risk-neutral mean $\mu^{(\star)} = (\mu - \lambda_0)$, $\Phi_{ss}^\star$ denotes the upper $K_s * K_s$ block of $\Phi^\star$, and Σ_{ss} is equal to the upper left $K_s * K_s$ block of Σ. Estimation of this model proceeds with the three-step regression adopted by Adrian, Crump, and Moench (2013). Estimates of the market prices of risk and the corresponding test statistics are provided in the following table. It shows that demand factors affect level risk and supply factors affect slope risk.

Table XVIII Macro factor model: market prices of risk

This table summarizes the estimates of the market price of risk parameters λ_0 and λ_1.Panel A summarizes the estimates of the market price of risk parameters λ_0 and λ_1 for the MTSM without unspanned factors and panel B is for MTSM with unspanned factors. t-Statistics are reported in parentheses. Standard errors have been computed according to the formulas from Section 4.1. Wald statistics for tests of the rows of and of 1 being different from zero are reported along each row, with the corresponding p-values in parentheses. PC1,, PC3 denote the first through third principal components of Treasury yields. Bolded coefficients represent significance at the 5% level.

Panel A:	λ_0	$\lambda_{1,1}$	$\lambda_{1,2}$	$\lambda_{1,3}$	$\lambda_{1,4}$	$\lambda_{1,5}$
PC1	-0.0177	-0.006	-0.0253	-0.0176	0.0113	0.0236
PC2	-0.0198	0.0104	-0.0470	-0.0106	-0.0137	0.0379
PC3	0.0228	-0.0535	0.0324	-0.1285	-0.1663	-0.0919
Panel B:	λ_0	$\lambda_{1,1}$	$\lambda_{1,2}$	$\lambda_{1,3}$	$\lambda_{1,demand}$	$\lambda_{1,supply}$
PC1	-0.0180	-0.0217	-0.0160	-0.0092	-0.0234	0.0075
PC2	-0.0186	0.0113	-0.0416	-0.0130	-0.0062	-0.0105
PC3	0.0159	0.0117	-0.0387	-0.1361	0.0606	-0.0748

F.4. Unspanned Factors Inference

F.5. Affine Yields

Under this model, bond prices are exponentially affine in the vector of state variables:

$$\ln P_t^{(n)} = A_n + B^{'} X_t + u_t^{(n)}. \tag{46}$$

By substituting Eq. (32) into Eq. (19), we see that

$$rx_{t+1}^{(n-1)} = A_{n-1} + B_{n-1}^{'} X_{t+1} + u_{t+1}^{(n-1)} - A_n - B_n^{'} X_t - u_t^{(n)} + A_1 + B_1^{'} X_t + u_t^{(1)}. \tag{47}$$

Equating this expression for excess returns with the return generating expression in Eq. (25), we find

$$A_{n-1} + B_{n-1}^{'}(\mu + \phi X_t + \nu_{t+1}) + u_{t+1}^{(n-1)} - A_n - B_n^{'} X_t - u_t^{(n)} + A_1 + B_1^{'} X_t + u_t^{(1)} \tag{48}$$

$$= \beta^{(n-1)'}(\lambda_0 + \lambda_1 X_t + \nu_{t+1}) - \frac{1}{2}(\beta^{(n-1)'}\Sigma\beta^{(n-1)} + \sigma^2) + e_{t+1}^{(n-1)}. \tag{49}$$

This equation has to hold state by state. Let $A_1 = -\delta_0$ and $B_1 = -\delta_1$. Matching terms, we obtain the following system of recursive linear restrictions for the bond pricing parameters:

$$A_n = A_{n-1} + B_{n-1}^{'}(\mu - \lambda_0) + \frac{1}{2}(B_{n-1}^{'}\Sigma B_{n-1} + \sigma^2) - \delta_0, \tag{50}$$

$$B_n^{'} = +B_{n-1}^{'}(\phi - \lambda_1) - \delta_1^{'}, \tag{51}$$

$$A_0 = 0, B_0^{'} = 0, \beta^{(n)'} = B_n^{'}. \tag{52}$$

F.6. Decomposition of Bond Yields into Expectation and Yield Risk Premium

Based on Campbell and Ammer (1993) decomposition, bond yields could be decomposed into two parts: expectation component and yield risk premium component.

$$y_t^{(n)} = \underbrace{\frac{1}{n}E_t\big(y_t^{(1)} + y_{t+1}^{(1)} + \cdots + y_{t+n-1}^{(1)}\big)}_{\text{Expectation Component}} + \underbrace{\frac{1}{n}E_t\big(rx_{t+1}^{(n-1)} + rx_{t+2}^{(n-2)} + \cdots + rx_{t+n}^{(0)}\big)}_{\text{Yield Risk Premium}} \tag{53}$$

Using Eq. (32), we can calculate expectation component as follows:

$$EC_t^{(n)} = -A_1 - \frac{1}{n}B_1'(I-\phi)^{-1}(I-\phi^n)X_t \tag{54}$$

Accordingly, yield risk premium is the residual in Eq (27).

$$\text{Yield Risk Premium} \equiv \Xi_t^{(n)} = y_t^{(n)} - EC_t^{(n)} \tag{55}$$

The following two graphs plot the decomposition results for unspanned factor model and spanned factor model. Figure 11 plots decomposition results for only spanned factor model. It is clear that yield risk premium follows a opposite path against bond yield pre-2009 and tracks exactly the bond yield path post-2009. In 2001 and 2007, yield risk premium reached zero level, which is definitely underestimating bond risk. In contrast, the model with unspanned factors would generate a higher level of yield risk premium with counter-cyclical path pattern.

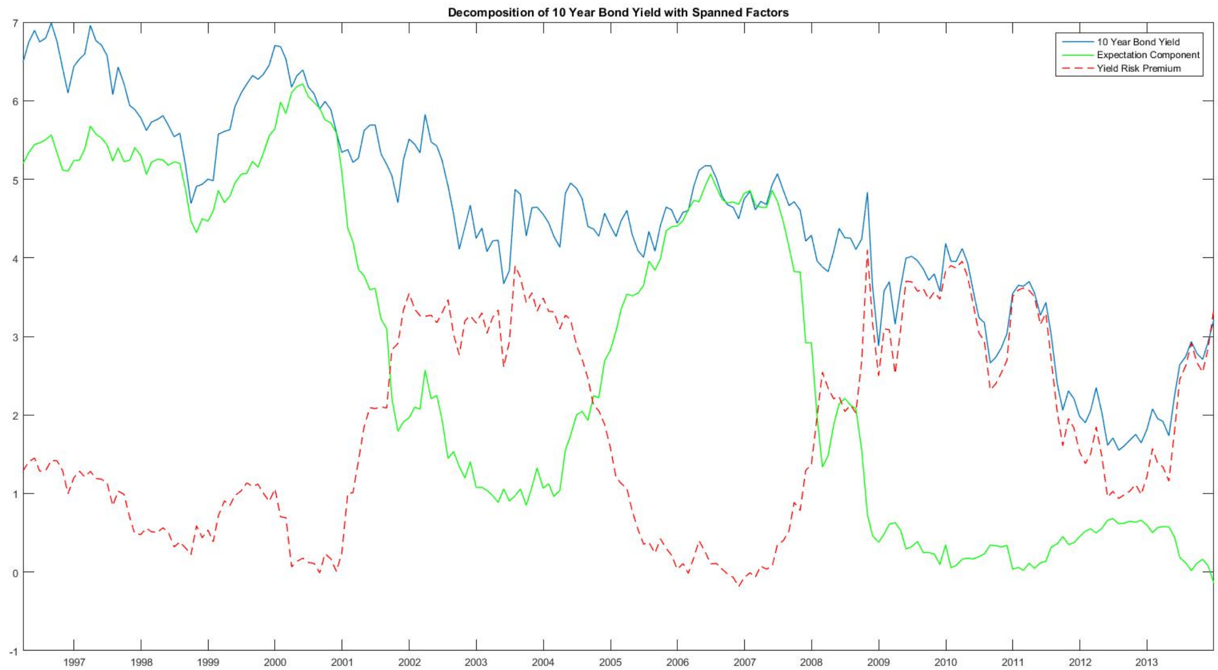


Figure 18. This figure plots the decomposition of 10 year US zero coupon treasury bond yields using unspanned factors.

G. Conclusion

We use asymptotic PCA to construct seven demand factors and four supply factors out of 133 monthly time series of economic activity spanning a period from 1996:1 to 2015:12. We find both demand factors and supply factors are unspanned by bond yields and have stronger predictability power for future bond excess returns even than CP factors. The longer maturity the bond have, the more predictive power these factors show. These predictive results are robust to or even stronger for machine learning technique LASSO. More interestingly, we can use LASSO to identify 15 most important economic variables out of original 133 variables and give direct economic explanation for bond demand factors and supply factors. Regarding to the equity market, we find both of demand and supply factors are priced by the cross-section of stock returns. In particular, portfolios with highest exposure to supply factor outperform portfolios with lowest exposure to supply factor 1.8% per month while portfolios with lowest exposure to demand factor outperform portfolios with highest exposure to demand factor 2.1% per month. This is consistent with ”fly to safety” explanation. Furthermore, variance decomposition from VAR including supply factor, demand factor, excess return, and bond yield shows that demand factors are much more important than supply factors in explaining bond excess returns. Finally, we use macro-finance affine term structure models (MTSMs) to estimate market price of risk of factors and find unspanned factors affect significantly the market price of risk of spanned factors. Demand factors affect level risk and supply factors affect slope risk. Yield risk premium for MTSMs with unspanned factor model and MTSMs with only spanned factor model are very different from each other and MISMs with only spanned factor model underestimate yield risk premium.

III. Chapter 3: Making Better Use of Options to Predict Stock Returns

A typical stock has several hundred listed options of different types, moneyness levels, and expirations. Thus, a rich multi-dimensional set of option information is available for a given stock on a given day. In this paper, we present a parsimonious method for summarizing this option information in just a few numbers using conventional dimensionality reduction techniques. In particular, we perform a principal component analysis (PCA) on the cross-section of implied volatility surfaces to extract its factor structure. We then study how this condensed option information represented by the main principal components predicts future stock and option returns. This topic has not been explored and is important since the literature documents that (combinations of) implied volatilities predict stock returns. Our main goals are to document the existence of a factor structure in the cross-section of implied volatility surfaces, to examine if these factors predict future stock returns, and to establish the information carried out by these factors.

We study the factor structure of implied volatility surfaces using the entire universe of U.S optionable stocks for the period 1996-2014. Given that the number of options varies across stocks, we use Optionmetrics data that contains the interpolated volatility surfaces for each security on each day. In particular, we extract 112 implied volatilities for each stock that span across option types (calls and puts), specified option maturities (from 30 to 365 days), and moneyness levels (absolute option deltas from 0.2 to 0.8). For each date, we perform PCA to the cross-stock correlation matrix of the demeaned volatility surfaces. PCA has been used to understand the term structure of interest rates (Litterman and Scheinkman (1991)), the term structure of credit and CDS spreads (Collin-Dufresn, Goldstein, and Martin (2001) and Pan and Singleton (2008)), and more recently, the equity volatility levels, skews, and term structures (Christoffersen, Fournier, and Jacobs (2017)).

The cross-section of implied volatility surfaces reveals a strong factor structure. The first five principal components (PCs) factors explain more than 70% of the variation of the cross-section of implied volatilities. The first factor alone explains 32% of the variation of the data. Since the volatility surfaces are demeaned, the first PC factor is not a level factor. Instead, the first PC factor captures two simultaneous effects: 1) call put spread: implied volatilities for calls and puts move in opposite directions, and 2) option skew: the differential between implied volatilities for OTM puts (ITM calls) and ITM puts (OTM calls). The second PC factor explains 18% of the data and captures the slope of the term structure effect: short-term volatilities and long-term volatilities move in opposite direction. The third PC explains 11% of the variation and captures the option skew, the volatility slope between OTM puts (ITM calls) and ITM puts (OTM calls). The fourth and fifth PCs explain 8% and 4% of the variation of the data and they capture non-linear aspects of the volatility surface.

Next we examine the relation between the first five PC factors and future stock returns. Fama-Macbeth regressions show that the first three PCs predict one-week and one-month stock returns. The relation of stock returns and the first three principal components is positive and significant for PC1, PC2, and PC3. Note that the sign of the relation between PCs and returns as well as the sign of the PCs is arbitrary since changing the sign of the PCs does not change the variance that is contained on each component. Using this information and the fact that the PC factors are orthogonal by construction, we create an aggregate PC factor equal to sum of the first three PCs.

The aggregate PC factor predicts weekly and monthly stock returns. At the weekly level, a strategy that buys the portfolio with the highest aggregate PC factor and sells the portfolio with the lowest aggregate factor has a return of 0.38% for equal weighted returns and 0.28% for value-weighted returns with corresponding t-statistics of 6.01 and 2.44. The risk-adjusted Fama-French-Carhart alphas remain of the same magitude and significance than the raw returns. At the monthly level, the long-short returns are 1.36% and 1.00% for equal and value weighted portfolios with t-statistics above 4.7. The Fama-French-Carhart alphas remain unchanged at 1.32% for equal-weighted and 0.94% for value-weighted portfolios but with higher t-statistics of 12.6 and 5.03 respectively. These results are confirmed with Fama-MacBeth regressions. The results remains unchanged when we perform the PCA decomposition using volatility surface data prior to the returns predicted date. We conclude that the aggregate PC factor predicts returns in- and out-of-sample.

Our paper is related to the literature that uses option market information to predict future stock returns. These predictors can be grouped in four categories: 1) call-put implied volatility spread (Bali and Hovakimian (2009), Cremers and Weinbaum (2010), and Yan (2011)), 2) risk-neutral skewness (Xing, Zhang, and Zhao (2010), Rehman and Vilkov (2012), Conrad, Dittmar, and Ghysels (2013), Stilger, Kostakis, and Poon (2016), and Bali, Hu, and Murray (2016)), 3) option to stock volume ratio (Roll, Schwartz, and Subrahmanyam (2010), and Johnson and So (2012)) and 4) volatility of implied volatility (Baltussen, Van Bekkum, and Van Der Grient (2012)). How are these predictors related to the PC factors from the cross-section of the implied volatility surfaces? The correlation matrix reveals that PC1 is highly correlated with the call-put implied volatility spread (-39% correlation) and the option skew used in Xing, Zhang, and Zhao (2010) (66% correlation). The second PC shows no strong correlation with any of the existing predictors, and the third PC as a correlation of 34% with the call-put implied volatility spread and -23% with risk-neutral skewness. Finally, the aggregate PC factor ($PC1 + PC2 + PC3$) is highly correlated with the call-put implied volatility spread (-50%) and the option skew from Xing, Zhang, and Zhao (2010) (52%), and mildly correlated with the option to stock volume ratio from Johnson and So (2012) (-11%).

Univariate and multivariate Fama-MacBeth regressions uncover the true power of the aggregate

PC factor. In univariate regressions, the call-put implied volatility spread, risk-neutral skewness, and the option to stock volume ratio successfully predict future stock returns. However, bivariate regressions reveal that only the aggregate PC factor predicts stock returns at the expense of existing option predictors. The aggregate PC factor seems to carry the predictive information of existing predictors from the option market.

We contribute to the vast literature that studies why option prices predict future returns in several ways. First, we show that the first five principal components explain more than 70% of variation in the implied volatility surfaces across stocks. Thus, the entire IV surface can be replaced with these few PCs with little information loss. In contrast with Christoffersen, Fournier, and Jacobs (2017) who also study the factor structure of implied volatilities for 30 stocks, we use the entire cross-section of volatility surfaces. While they develop a new option pricing model, we examine the stock return predictability of the PC factors.

Second, it is well-known that option prices and implied volatilities predict future stock returns. These predictors can be grouped in four categories: 1) the call-put implied volatility spread , 2) risk-neutral moments such as skewness , 3) option to stock volume ratio, and 4) volatility of implied volatility. Obviously, these ad hoc ways to aggregate the IV surface can lose or miss relevant information. We could include the entire surface in the predictive return regression. However the IV surface contains more than one-hundred volatilities and we might incur in overfitting and data mining. This is one of the main reasons the literature uses ad hoc IV aggregations in the first place. Instead we replace the surface with its main principal components and use them as return predictors. We go even further and aggregate the main PCs into a single factor: "the aggregate PC." Not only this aggregate PC factor strongly predicts future stock returns but it also completely wipes out the predictability of existing option-based predictors mentioned above. The fact that a single variable aggregates the information of all option-based predictors is striking.

To uncover the economic driver of the return predictability, we study how the aggregate PC interacts with proxies for alternative hypotheses. The three main explanations that we test are 1) informed trading (Cremers and Weinbaum (2010), Roll, Schwartz, and Subrahmanyam (2010), and Xing, Zhang, and Zhao (2010)), 2) jump risk (Bali and Hovakimian (2009), and Yan (2011)), and 3)short sale constraints (Johnson and So (2012), and Stilger, Kostakis, and Poon (2016)). We find that measures of short-sale constraints such as the stock lending fee drive out the predictability of the aggregate PC making it insignificant in stock return regressions. Perhaps, option prices predict returns simply because they reflect short-sale constraints.

Given that we work with the factor structure of the implied volatility surface, we also examine its relation with the cross-section of *option* returns. Univariate regressions reveal a negative and significant correlation between the aggregate PC factor and future option returns at the weekly level. Fama-MacBeth regressions confirm the negative relation after we control for the slope of the volatility term structure (Vasquez (2017)), historical minus implied volatility (Goyal and Saretto (2009)), and idiosyncratic volatility (Cao and Han (2013)).

The paper is organized as follows. Section 2 describes the data and the methodology we follow to study the factor structure of the cross-section of implied volatilities. Section 3 explores the predictability of the principal components on future stock returns. Section 4 tests the robustness of the results as well as the predictability of option returns. Section 5 concludes.

A. Data and Methodology

In this section we first describe the data. We then explain how the principal component analysis is performed on the volatility surfaces to construct the aggregate principal component factor. Finally, we form portfolios by sorting stocks into deciles based on the exposure to the principal components, and then report on the characteristics of these portfolios.

A.1. Data

We use the cross-section of volatility surfaces from the Optionmetrics Ivy database which provides end-of-day summary statistics on all exchange-listed option on U.S. equities from 1996 to 2014. The Optionmetrics volatility surface data contain the interpolated volatility surfaces for each security on each day, using a methodology based on a kernel smoothing algorithm. A standardized option is included only if there exists option price data to properly interpolate the required values. The volatility surface data contains implied volatility data for calls and puts across standardized maturities and deltas for each stock. The standardized expirations are 30, 60, 91, 122, 152, 182, 273, and 365 calendar days at absolute deltas 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, and 0.8. Each stock has 112 standardized implied volatilities across different option types, option maturities, and option moneyness levels. We eliminate close-end funds, real estate investment trusts, American depository receipts, and stock with price below $1.

We report variables obtained from the option market that are related with future stock returns. We compute the call put parity volatility spread as the open-interest weighted difference between call and put implied volatilities as in Cremers and Weinbaum (2010).[1] Risk-neutral volatility, skewness, and kurtosis is calculated as in Bakshi, Kapadia, and Madan (2003) for the last trading day before the testing period as used in Rehman and Vilkov (2012) and Stilger, Kostakis, and Poon (2016). Option skewness is the difference between OTM put and ATM call implied volatilities as computed by Xing, Zhang, and Zhao (2010). The option to stock volume ratio is the total volume in option contracts across all strikes for options with less than 30 days to expiration over the total volume in the stock as in Johnson and So (2012) and Roll, Schwartz, and Subrahmanyam (2010). The volatility of implied volatility is the standard deviation of the previous month ATM implied volatilities as in Baltussen, Van Bekkum, and Van Der Grient (2012). The slope of the implied

[1] Bali and Hovakimian (2009) and Yan (2011) define the call-put spread as the negative difference between one-month implied volatilities of a call and a put with absolute delta of 0.5.

volatility term structure is defined as the difference between long and short term ATM implied volatilities.

We also include firm characteristics in the analisys. These variables are extracted from Center for Research and Security Prices (CRSP) and Compustat. The final data sample is formed by the intersection of Optionmetrics, CRSP, and Compustat data. From CRSP, we extract the stock market capitalization (size) and we use daily returns to calculate weekly returns from Tuesday close to Tuesday close, and monthly returns to compute the 6-month return for all firms. From Compustat we extract book values to calculate book-to-market ratios of individual firms. The predictive variables are computed skipping one day before the trading period.

A.2. Methodology

Using the implied volatility surfaces for all stocks and for every Monday over the 1996 to 2014 period, we compute the correlation matrix of the 112 demeaned implied volatilities. To ensure valid volatility surfaces, we only include stocks that traded at least 50 calls and 50 put contracts. After decomposing the correlation matrix using principal component analysis (PCA), we find that the first five principal components (PCs) explain 78% of the total variance of the demeaned implied volatilities. Figure 1 plots the relative contribution of the first ten principal components. The first five principal components explain 32%, 21%, 13%, 8%, and 4% of the variation of the data. By construction these PC factors are orthogonal and independent of each other.

**Wenzhi Comment: Using the implied volatility surfaces for all stocks and for every week over the 1996 to 2014 period (we collpase daily data obtained from Optionmetrics to weekly data by averaging the daily oberservations for each week), we compute the correlation matrix of the 112 demeaned implied volatilities. To ensure valid volatility surfaces, we only include stocks that traded at least 50 calls and 50 put contracts. After decomposing the correlation matrix using principal component analysis (PCA), we find that the first five principal components (PCs) explain 78% of the total variance of the demeaned implied volatilities. Figure 1 plots the relative contribution of the first ten principal components. The first five principal components explain 32%, 21%, 13%, 8%, and 4% of the variation of the data. By construction these PC factors are orthogonal and independent of each other. **

[Figure 1 about here]

Every Monday, we compute the exposure of each stock to each of the five PCs. The implied volatility exposure of a stock to a principal component is computed as the multiplication of the vector of 112 demeaned implied volatilities with the PC vector. Using this procedure we obtain a weekly measure of the exposure of each stock to each of the five PCs. Overall we obtain $1,897,536$ firm-weeks corresponding to 936 weeks and $1,982$ unique firms.

**Wenzhi Comment: Every week, we compute the exposure of each stock to each of the five PCs. The implied volatility exposure of a stock to a principal component is computed as the multiplication of the vector of 112 demeaned implied volatilities with the PC vector. Using this procedure we obtain a weekly measure of the exposure of each stock to each of the five PCs. Overall we obtain $1,897,536$ firm-weeks corresponding to 936 weeks and $1,982$ unique firms. **

A.3. Summary Statistics

[Table 1 about here]

Table 1 reports summary statistics for the first five principal components, the option market measures, and the underlying stock measures. The first five principal components have a mean of zero given that we work with demeaned volatility surfaces. Based on the predictability results from

Table 3, we construct an aggregate PC factor that adds up the first three PCs.[2] The aggregate PC factor also has zero mean.

[Figure 2 about here]

Figures 2 plots the first 5 principal components for the cross-section of implied volatility surfaces. To facilitate interpretation, we plot one graph for calls and one for puts across moneyness, defined as strike over stock price, and maturities. Each surface contains 56 datapoints. Panel A of Figure 1 reports the first principal component (PC1) that explains 32% of the variance of the data. PC1 can be interpreted as the call-put volatility spread factor combined with the option skew. While the factor loadings for calls are all positive, those for puts are all negative. According to PC1, put and call implied volatilities move in opposite directions. In addition, the loadings for OTM (ITM) volatilities for puts (calls) are lower than those of ITM (OTM) volatilities which reflects the well know option skew. When the first PC factor increases, call implied volatilities go up and put volatilities go down; but OTM (ITM) volatility for puts (calls) increases more than those of ITM (OTM) volatilities for puts (calls). This interpretation is confirmed by the correlations reported in Table 2. PC1 has a correlation of 66% and −39% with the call-put volatility spread defined in Cremers and Weinbaum (2010) and the option skew defined in Xing, Zhang, and Zhao (2010).

The second principal component (PC2) is reported in Panel B of Figure 1. PC2 explains 18% of the total variance of the data and represents a term-structure factor. While the factor loadings are positive for short-term maturities for both calls and puts, they are negative for long-term maturities. The factor loading changes signs at about 150 days to expiration. Panel C of Figure 1 reports the third principal component (PC3) that explains 11% of the data variation and corresponds to the option skew. For puts (calls), the factor loadings are positive (negative) for ITM

[2] The rationale to add PC1, PC2, and PC3 is that these PCs have a positive and significant relation with future stock returns. We only include the first three PCs since they explain most of the variation of the cross-section of volatility surfaces. Note that the sign on the principal component loadings is arbitrary, hence the predictability could be negative as well.

options and negative (positive) for OTM options. PC3 has a correlation of 28% with the option skew and -22% with risk-neutral skewness. Finally, PC4 and PC5 are reported in Panel D and E of Figure 1, and they explain 8% and 4% of the variation of the data. These factors are non-linear and their interpretation is not straightforward. The highest absolute correlations are between PC4 and risk-neutral skewness at -15%, and between PC5 and the implied volatility spread at -26%.

[Figure 3 about here]

Figure 3 reports the time series average of the first three principal components. Panel A plots the time series for PC1. The time-series revolves around zero and increases its variability after 2008. PC1 does not seem to be persistent since the autocorrelation of lag 1 is only 7%. Panel B plots the time series for PC2 and PC3. Both PCs also revolve around zero and they appear to be more persistent than PC1. The lag 1 autocorrelations for PC2 and PC3 are 90% and 95%, respectively. Both reach their maximum value in 2008.

[Figure 4 about here]

Figure 4 contains the loadings and the time series average of the aggregate PC factor defined as the sum of the first three PCs. As displayed in Panel A, the aggregate PC factor loadings are the most negative for low levels of moneyness, defined as strike over stock price, and short maturities for both calls and puts. As the maturity and the moneyness increase, so does the factor loadings. The factor loadings are the most positive for high levels of moneynes and long term maturities. The factor loadings change sign for a moneyness of 1 for calls with maturity below 90 days. For puts, the sign of the loading changes for long term maturities (above 240 days) with moneyness above 1.1.

Figure 4, Panel B graphs the time series average of the aggregate PC factor. The aggregate PC takes positive and negative values. In calm periods, the aggregate PC remains positive. However, in crash periods such as September 2001 or the crisis in 2008, the aggregate PC factor displays big negative spikes similar to those of the volatility index VIX.

[Table 2 about here]

Table 2 reports the correlation matrix of the principal components, option measures, and firm characteristics. PC1 is highly correlated with two variables that predict stocks returns: the implied volatility spread by Cremers and Weinbaum (2010) with a correlation of 71% and the option skew by Xing, Zhang, and Zhao (2010) with a correlation of -46%. PC2 is not correlated with any existing variables that predicts stock returns. PC3 is mildly correlated with the option skew and the risk-neutral skewness with correlations of 28% and -22%, respectively. The aggregate PC factor is a linear combination of the first three PCs ($PC1 + PC2 + PC3$) and the correlations are 74% with PC1, 53% with PC2, and 41% with PC3. The correlation of the aggregate PC factor with the volatility spread and the option skew is 55% and -51%, and it is not highly correlated with any other variable. The correlation matrix in Table 2 shows that existing factors such as size, book-to-market, and illiquidity are not related with the PC factors computed from the implied volatility surfaces. For this reason, any stock return predictability coming from the implied volatility surface PC factors is not likely to be related to firm characteristics such as size, book-to-market, momentum, reversal, or illiquidity.

We conclude that the first five principal components explain most of the variation of the data and they are highly correlated with existing return predictors extracted from options such as the implied volatility spread, the option skew, and the risk neutral skewness. We now turn to explore the ability of these principal components to predict stock returns.

B. Principal Components and Future Stock Returns

In this section, we first analyze the relationship between the current week's returns and the previous week's principal components using the Fama and MacBeth (1973b) methodology. Next we explore the predictability of the aggregate PC factor and we then run a horse race between the aggregate PC factor and existing predictors of stock returns extracted from options to see which one is the best predictor. We also form portfolios by sorting stocks based on an aggregate PC factor.

B.1. Fama-MacBeth Regressions

B.2. First Five Principal Components

To assess the relationship between future returns and principal components, we carry out various cross-sectional regressions using the method proposed in Fama and MacBeth (1973b). Each week t, we compute the principal components for firm i and estimate the following cross-sectional regression:

$$r_{i,t+1} = \gamma_{0,t} + \gamma_{1,t}PC1_{i,t} + \gamma_{2,t}PC2_{i,t} + \gamma_{3,t}PC3_{i,t} + \gamma_{4,t}PC4_{i,t} + \gamma_{5,t}PC5_{i,t} + \phi_t' Z_{i,t} + \varepsilon_{i,t+1}, \quad (56)$$

where $r_{i,t+1}$ is the weekly return of the ith stock for week $t+1$ (from Tuesday close to Tuesday close), PC1 to PC5 are the first five principal components for firm i at the end of week t (Friday), and $Z_{i,t}$ represents a vector of characteristics and controls for the ith firm observed at the end of week t (Friday). Note that we skip one day between portfolio formation and testing period.

[Table 3 about here]

Table 3 presents the results of the Fama-MacBeth regressions from regressing one-week and one-month stock returns on the volatility surface PCs factors, option variables, and firm characteristics.

In the first regression we include all PCs and the contemporaneous return to control for return-reversal. The first three PCs significantly predict stock returns at the 1% level. PC1, PC2, and PC3 are positively and significantly related with future stock returns. The coefficient for PC4 is not significant, while that of PC5 is negative and significant at the 10% level. Regression 4 presents the results for monthly returns which are similar to those of one-week returns.

In regression 2, we include firm related control variables such as lagged one-week return, lagged six-month return, size, book-to-market ratio, and the illiquidity measure by Amihud. The results remain unchanged compared to regression 1: PC1, PC2, PC3, and PC5 continue to predict stock returns significantly. A similar pattern is observed for monthly returns in regression 5.

Finally, we include three option related variables that are commonly used to charaterize the implied volatiltiy surface. The first variable is the slope of the implied volatility surface defined as the difference between long-term and short-term ATM implied volatilities. The second variable is the option skew as Xing, Zhang, and Zhao (2010), and the third variable is the implied volatility spread as defined by Cremers and Weinbaum (2010). We also include the option to stock volume ratio as defined by Johnson and So (2012). Regression 3 and 6 present the results for weekly and monthly returns. The predictability of the first two PCs is confirmed in the two regressions after we include all control variables.

We now proceed to aggregate the information of the first three principal components into a single factor and explore the predictability of that aggregate factor.

B.3. Aggregate Principal Component Factor

Given that the first three principal components predict stock returns and that by construction the PC factors are uncorrelated, we create an aggregate principal component (PC) factor. The goal of the aggregate PC factor is to combine the information from the three PCs that predict stock returns into a single variable. The aggregate PC factor is a linear combination of the first three PCs. Since PC1, PC2, and PC3 have a positive relation with future stock returns, we define the aggregate PC factor as $PC1 + PC2 + PC3$.

[Table 4 about here]

In Table 4, we present the results of regressing one-week and one-month future stock returns on the aggregate PC factor. In regression 1 and 4, we control for the reversal effect. The aggregate PC factor successfully predicts one-week and one-month returns. Next we include firm characteristics such as lagged six-month return, size, book-to-market ration and the Amihud illiquidity meaure. As reported in regressions 2 and 5, the coefficient of the aggregate PC factor remains unchanged and the significance increases.

Finally, we include in regressions 3 and 6 three variables that are normally used to describe the implied volatility surface: 1) the slope of the implied volatility surface defined as the difference between long-term and short-term ATM implied volatilities, 2) the option skew as Xing, Zhang, and Zhao (2010), and 3) the implied volatility spread as defined by Cremers and Weinbaum (2010). We also include the option to stock volume ratio as defined by Johnson and So (2012). The aggregate PC factor continues to predict future weekly and monthly returns. The coefficient is positive and highly significant.

We conclude that the aggregate PC factor successfully combines the information of the first three PCs to predict future stock returns.

B.4. Aggregate PC Factor and Existing Predictors from Options

Several papers show that there is a connection between the option and the stock markets. The four main variables extracted from the option market that predict stock returns are the call-put implied volatility spread, risk-neutral skewness, the option to stock volume ratio, and the volatility of implied volatility. Various definitions of the call-put implied volatility spread are used to predict returns in Cremers and Weinbaum (2010), Bali and Hovakimian (2009), and Yan (2011). Risk-neutral skewness predictability is documented in Bali, Hu, and Murray (2016), Conrad, Dittmar, and Ghysels (2013), Rehman and Vilkov (2012), Stilger, Kostakis, and Poon (2016), and Xing, Zhang, and Zhao (2010). The predictability of the option to stock volume ratio has been documented by Johnson and So (2012) and Roll, Schwartz, and Subrahmanyam (2010). The volatility of implied volatility predicts future stock returns according to Baltussen, Van Bekkum, and Van Der Grient (2012).

In this section we first confirm the predictability of these variables extracted from the option market. Then we explore how these variables predict returns in the presence of the aggregate PC factor. Given that the aggregate PC factor incorporates the most relevant information from the variation in the implied volatility surfaces, we expect that it subsumes the predictability of existing option based predictors.

[Table 5 about here]

Table 5 reports the results from regressing future stock returns on predictors extracted from the option market and on the aggregate PC factor. We explore six option based predictors: 1) the implied volatility spread, the average difference between call and put options, as defined by Cremers and Weinbaum (2010), 2) the volatility smirk or option skew, the difference beweeen OTM put and ATM call implied volatilities, as defined by Xing, Zhang, and Zhao (2010), 3) the risk-neutral skewness and 4) risk-neutral kurtosis as defined in Conrad, Dittmar, and Ghysels (2013), 5) the O/S option to stock volume ratio as defined in Johnson and So (2012), and 6) the volatility of implied volatility as defined in Baltussen, Van Bekkum, and Van Der Grient (2012).

First, we perform univariate regressions of future stock returns on each of these variables extracted from the option market. We confirm the predictability of the volatility spread, the option skew, risk-neutral skewness, the O/S ratio predict stock returns, and the volatility of volatility. The volatility spread and risk neutral skewness have a positive and significant relation with future stock returns, while the option skew, the O/S ratio, and the volatility of volatility have a negative and significant relation with stock returns. The results hold for weekly and monthly returns as reported in Panel A and Panel B.

In the second set of regressions, we explore the impact that the aggregate PC factor has on the predictability of these 6 option based variables. After including the aggregate PC factor as a control variable, none of the option-based variables predicts stock returns anymore. The predictability is driven by aggregate PC factor whose coefficient is positive and significant in all regressions. These results hold for weekly and monthly returns.

We conclude that the aggregate PC factor outperforms existing stock return predictors that use combinations of implied volatilities such as the volatility spread, the option skew, risk-neutral skewness, and volatility of implied volatility. The aggregate PC factor seems to successfully embed the information of all of the existing option-based predictors.

B.5. Portfolio Sorts by the Aggregate PC Factor

We have shown that the aggregate PC factor predicts stock returns above and beyond existing predictors extracted from options. We now explore the predictability of the aggregate PC factor using the portfolio sort methodology. Every week we sort stocks by the aggregate PC factor and form ten portfolios. Portfolios 1 (10) contains stocks with the lowest (highest) level in the aggregate PC factor. We report the long-short portfolio which buys portfolio 10 and sells portfolio 1. Additionally we risk-adjust all portfolio returns with the Fama-French and Carhart factors.

[Table 6 about here]

Table 6 reports average returns and risk-adjusted returns of decile portfolios sorted by the aggregate PC factor. We report equal and value weighted weekly and monthly returns. In the first column, we present the results for equal-weighted weekly returns. Returns for decile 1 are lower than those for decile 10. The long-short portfolio has a weekly return of -0.38% with a t-statistic of 6.01. In the second column we regress the average portfolio returns on the Fama-French-Carhart factors. The risk-adjusted return for portfolio 1 is negative while the return for portfolio 10 is positive. The alpha of the long-short portfolio is positive and significant. In the third and fourth column we report the results for value-weighted weekly returns. The positive relation between the aggregate PC factor and stock returns is confirmed.

We repeat the portfolio sorts for monthly returns. The equal-weighted monthly returns are positive and significant and the magnitude of the long-short return is about 4 times that of the weekly return. The t-statistic is twice as big as the one of weekly returns. The Fama-French-Carhart alpha and the t-statistic of the long-short returns are almost identical to that of the raw returns. The results for value-weighted monthly returns confirm the results.

We conclude that the aggregate PC factor extracted from the cross-section of implied volatility surfaces predicts stock returns at the weekly and monthly horizons for equal-weighted and value weighted portolios. The results cannot be explained by the Fama-French-Carhart model.

B.6. Persistence of the Aggregate PC Factor

We have shown that the aggregate PC factor predicts future stock returns at the weekly and monthly frequencies. To assess how persistent the aggregate PC factor is, we now analyse the transition probabilities across decile portfolios. Table 7 reports the probability that a stock moves up or down one portfolio, that it remains in the same portfolio or that it moves to any other portfolio. These probabilities are reported for all portfolios and for portfolio 10 at the weekly and monthly frequencies.

When analysing all 10 decile portfolios, we find that the probability of remaining in the same portfolios is much higher than the probability of moving either up or down one portfolio. The probability of remaining in the same portfolio is 24% at the weekly horizon and 36% at the monthly horizon. Note that the unconditional probability of moving to any portfolio is 10%.

We report the same probabilities for decile portfolio 10 since most of the long-short stock return predictability at the weekly and monthly frequencies comes from that portfolio. We find that the probability of remining in portfolio 10 is 41% and 59% for weekly and monthly frequencies. These numbers show that the aggregate PC factor is highly persistent and that portfolio 10 returns are generated by similar stocks from week to week.

[Table 7 about here]

B.7. Long-Term Predictability

Given that the aggregate PC factor is persistent, we explore the long-term predictability of the aggregate PC factor. So far we showed that the aggregate PC factor can predict stock returns over 1 week and 1 month horizons. Xing, Zhang, and Zhao (2010) and Johnson and So (2012) show that the option smirk and the implied volatility spread can predict stock returns at longer horizon than one week. Because the aggregate PC drives out the predictability of the option smirk and the implied volatility spread, the aggregate PC factor might predict stock returns beyond one month. For this reasons, the aggregate PC factor might predict stock returns well beyond one month.

[Figure 5 about here]

Figure 5 plots the long-short value-weighted return of the portfolios sorted by the aggregate PC factor over 30 weeks (about 7 months). Panel A reports the average return of the value-weighted long-short portfolio at week t, where t varies from 1 to 30. The average return is surrounded by error bars that represent the 95% confidence interval. We observe that the returns are consistently positive and significant up to week 15. From week 16 to week 26, the returns are positive and significant 7 times out of 11. After week 27, the returns are not significant anymore.

Panel B graphs the cumulative weekly returns of the long-short portfolio. The cummulative returns are positive and increasing up to week 26. On the first week, the long-short return is 0.38% and it accummulates to 2.6% after 26 weeks. We conclude that the aggregate PC factor predicts stock returns for long horizons up to six months.

B.8. Potential Explanantions

In this section we explore potential explanations of the sources of the return predictability by the aggregate PC factor from the cross-section of implied volatility surfaces. To do so we first look at existing predictors from the option market and the explanations provided to explain their results. There are several explanations of the predictability such as short-sale constraints in the equity market, informed traders deciding to trade in options to profit from leverage or to profit from mispriced stocks, jump risk, firm misvaluation, and risk-return trade off. The most popular explanations are short sale constraints (Johnson and So (2012), and Stilger, Kostakis, and Poon (2016)), informed trading (Cremers and Weinbaum (2010), Roll, Schwartz, and Subrahmanyam (2010), and Xing, Zhang, and Zhao (2010)), and jump risk (Bali and Hovakimian (2009), and Yan (2011)).

[Table 8 about here]

In Table 8 we study the relation between short sale constraints, the aggregate PC factor, and future stock returns. We use the lending fee as a proxy of short-sale constraints. In the first regressions we confirm the relation between the aggregate PC factors. In the second regression we use the lending fee to predict future stock returns. The stock lending fee has a negative and significant relation with returns. In the last regressions we include both the aggregate PC factor and the stock lending fee. Surprisingly the stock lending fee wipes out the predictability power of the aggregate PC factor. Therefore, the aggregate PC factor might be a proxy of short sale constraints. Stock that are harder to short sale must pay a premium to the investor and the aggregate PC factor is just a proxy of these short-sale constraints.

C. Robustness

We now look at the robustness of our results. First we compute the aggregate PC factor out-of.sample, that is, using information before the period to be predicted. Second, we study the relation between the aggregate PC factor and future *option* returns in the cross-section.

C.1. Out of Sample Aggregate PC Factor

In this study we document a positive relation between the aggregate PC factor extracted from the cross section of volatility surfaces and future stock returns. To perform the PCA analysis we use the full data period, so we incur in a look-ahead bias. To avoid this concern, we perform the PCA analysis using a 3-year rolling window of volatility surfaces. We use volatility surfaces up to time t to forecast returns on week t+1. Then we use the first three PCs to contruct the aggregate PC factor that is used to predict weekly and monthly returns. Table 9 presents the results of the Fama-MacBeth regressions of next week and next month returns on the aggregate PC factor. The positive and significant relation between the aggregate PC factor and future stock returns is confirmed in all regressions.

[Table 8 about here]

C.2. Option Returns

In this paper we extract the aggregate PC factor information from the cross-section of equity implied volatility surfaces and document that it predicts stock returns. A more natural experiment is to study the relation between the aggregate PC factor and future *option* returns in the cross-section. In this section, we study the relation between the aggregate PC factor and weekly straddle returns. The straddle is an option strategy that simultaneously buys an at–the–money call option and an at–the–money put option of the same underlying stock, with the same strike price, and the same expiration date. This position is almost delta neutral so that the stock price is not the main driver of the straddle return. We study portfolio sorts and Fama-MacBeth regressions of one-week straddle returns formed with options that expire in 3 to 6 weeks.

C.3. Portfolio Sorts

Straddle returns are mainly driven by changes in volatility since its delta is close to zero. The aggregate PC factor has two qualities that make it a ideal candidate to predict straddle returns. First, it is extracted from implied volatilities of equity options. Therefore it captures the drivers of volatilities which are the main variables that affect straddle returns. Second, since the aggregate PC factor is highly persistent it should be related with contemporaneous as well as future returns.

[Table 10 about here]

Every week we form decile portfolios based on the aggregate PC factor and report next-week straddle returns. Panel A of Table 10 shows that there is a negative relation between the aggregate PC factor and future straddle returns. The portfolio that buys decile 10 and sells decile 1 has a weekly return of –4.2% with a t-statistic of –14.83. The returns of the long-short portfolio are negatively skewned and fat tailed. The predictability remains very similar for value-weighted returns.

C.4. Fama-MacBeth Regressions

To further confirm that the aggregate PC factor predicts straddle returns, we run Fama-MacBeth regressions on weekly straddle returns. Panel B of Table 10 presents the results. The first column reports univariate regressions that confirm the robustness of the negative relation between the aggregate PC factor and straddle returns. In the second column, we include three variables that predict option returns: the slope of the volatility term structure (Vasquez (2017)), historical minus implied volatility (Goyal and Saretto (2009)), and idiosyncratic volatility (Cao and Han (2013)). The multivariate regression confirms that the aggregate PC factor is related with option returns. The coefficient of the aggregate PC factor and its significance are higher in the multivariate than in the univariate regression. Moreover, the significance of the aggregate PC factor is the highest among all regressors.

We conclude that portfolio sorts and Fama-MacBeth regressions support the negative relation between the aggregate PC factor and future straddle returns.

D. Conclusion

In this paper we study the factor structure of the cross-section of implied volatility surfaces from Optionmetrics for the period 1996-2014. We find that five principal components (PC) explain more than 70% of the variability of the volatility surfaces. The PC factors contain information related to existing stock return predictors obtained from the option market such as the difference between put and call implied volatility, the option smirk, and the term structure of implied volatilities.

Next, we show that the first 3 PC factors predict stock returns. We construct an aggregate PC factor based on these 3 PC factors, and show that not only it explains stock returns but also drives out the predictability of existing predictors extracted from the option market.

We examine potential explanations of our results. Short-sale constraints seems the most plausible explation as it wipes out the predictability of the aggregate PC factor. Finally, we find that the aggregate PC factor also predicts the cross-section of option returns.

REFERENCES

Acharya, Viral V, Lasse H Pedersen, Thomas Philippon, and Matthew Richardson, 2017, Measuring systemic risk, *The Review of Financial Studies* 30, 2–47.

Adrian, Tobias, and Markus K Brunnermeier, 2016, Covar, *The American Economic Review* 106, 1705–1741.

Alanis, Emmanuel, Sudheer Chava, and Praveen Kumar, 2016, Shareholder bargaining power, debt overhang, and investment, .

Allen, Linda, Turan G Bali, and Yi Tang, 2012, Does systemic risk in the financial sector predict future economic downturns?, *The Review of Financial Studies* 25, 3000–3036.

Amihud, Yakov, 2002, Illiquidity and stock returns: cross-section and time-series effects, *Journal of financial markets* 5, 31–56.

Andrieu, Christophe, Arnaud Doucet, and Roman Holenstein, 2010, Particle markov chain monte carlo methods, *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 72, 269–342.

Ang, Andrew, Robert J Hodrick, Yuhang Xing, and Xiaoyan Zhang, 2006, The cross-section of volatility and expected returns, *The Journal of Finance* 61, 259–299.

Bakshi, Gurdip, Nikunj Kapadia, and Dilip Madan, 2003, Stock return characteristics, skew laws, and the differential pricing of individual equity options, *The Review of Financial Studies* 16, 101–143.

Bali, Turan G, Stephen Brown, Qiusha Peng, and Yi Tang, 2016, Is economic uncertainty priced in the cross-section of stock returns?, .

Bali, Turan G, and Armen Hovakimian, 2009, Volatility spreads and expected stock returns, *Management Science* 55, 1797–1812.

Bali, Turan G, Jianfeng Hu, and Scott Murray, 2016, Option implied volatility, skewness, and kurtosis and the cross-section of expected stock returns, .

Baltussen, Guido, Sjoerd Van Bekkum, and Bart Van Der Grient, 2012, Unknown unknowns: Vol-of-vol and the cross section of stock returns, *Erasmus University Rotterdam Working Paper*.

Beber, Alessandro, Michael W Brandt, and Maurizio Luisi, 2015, Distilling the macroeconomic news flow, *Journal of Financial Economics* 117, 489–507.

Billio, Monica, Mila Getmansky, Andrew W Lo, and Loriana Pelizzon, 2012, Econometric measures of connectedness and systemic risk in the finance and insurance sectors, *Journal of Financial Economics* 104, 535–559.

Bisias, Dimitrios, Mark Flood, Andrew W Lo, and Stavros Valavanis, 2012, A survey of systemic risk analytics, *Annu. Rev. Financ. Econ.* 4, 255–296.

Brownlees, Christian T, and Robert F Engle, 2012, Volatility, correlation and tails for systemic risk measurement, *Available at SSRN 1611229*.

Brunnermeier, Markus K, and Lasse H Pedersen, 2009, Funding liquidity and market liquidity, *Review of Financial Studies* 22, 6.

Campbell, John Y, Jens Hilscher, and Jan Szilagyi, 2008, In search of distress risk, *The Journal of Finance* 63, 2899–2939.

Cao, Jie, and Bing Han, 2013, Cross-section of stock option returns and stock volatility risk premium, *Journal of Financial Economics*.

Chava, Sudheer, 2014, Environmental externalities and cost of capital, *Management Science* 60, 2223–2247.

——— , and Robert A Jarrow, 2004, Bankruptcy prediction with industry effects, *Review of Finance* 8, 537–569.

Chen, Zhuo, and Andrea Lu, 2016, A market-based funding liquidity measure, .

Christoffersen, Peter, Mathieu Fournier, and Kris Jacobs, 2017, The factor structure in equity options, *The Review of Financial Studies* 31, 595–637.

Cochrane, John H, and Monika Piazzesi, 2002, Bond risk premia, Discussion paper, National Bureau of Economic Research.

Collin-Dufresn, Pierre, Robert S Goldstein, and J Spencer Martin, 2001, The determinants of credit spread changes, *The Journal of Finance* 56, 2177–2207.

Conrad, Jennifer, Robert F Dittmar, and Eric Ghysels, 2013, Ex ante skewness and expected stock returns, *The Journal of Finance* 68, 85–124.

Cremers, Martijn, and David Weinbaum, 2010, Deviations from put-call parity and stock return predictability, *Journal of Financial and Quantitative Analysis* 45, 335–367.

Diebold, Francis X, and Kamil Yilmaz, 2009, Measuring financial asset return and volatility spillovers, with application to global equity markets, *The Economic Journal* 119, 158–171.

Duffee, Gregory R, 2011, Information in (and not in) the term structure, *The Review of Financial Studies* 24, 2895–2934.

Fama, Eugene F, and James D MacBeth, 1973a, Risk, return, and equilibrium: Empirical tests, *Journal of political economy* 81, 607–636.

———, 1973b, Risk, return, and equilibrium: Empirical tests, *Journal of political economy* 81, 607–636.

Fontaine, Jean-Sébastien, René Garcia, and Sermin Gungor, 2016, Funding risk, market liquidity, market volatility and the cross-section of asset returns, .

Giglio, Stefano, Bryan Kelly, and Seth Pruitt, 2016, Systemic risk and the macroeconomy: An empirical evaluation, *Journal of Financial Economics* 119, 457–471.

Gilchrist, Simon, and Egon Zakrajšek, 2012, Credit spreads and business cycle fluctuations, *The American Economic Review* 102, 1692–1720.

Goyal, Amit, and Alessio Saretto, 2009, Cross-section of option returns and volatility, *Journal of Financial Economics* 94, 310–326.

Greenwood, Robin, Samuel G Hanson, and Jeremy C Stein, 2015, A comparative-advantage approach to government debt maturity, *The Journal of Finance* 70, 1683–1722.

Greenwood, Robin, and Dimitri Vayanos, 2014, Bond supply and excess bond returns, *The Review of Financial Studies* 27, 663–713.

Herbst, Edward P, and Frank Schorfheide, 2015, *Bayesian Estimation of DSGE Models* (Princeton University Press).

Johnson, Travis L, and Eric C So, 2012, The option to stock volume ratio and future returns, *Journal of Financial Economics* 106, 262–286.

Joslin, Scott, Marcel Priebsch, and Kenneth J Singleton, 2014, Risk premiums in dynamic term structure models with unspanned macro risks, *The Journal of Finance* 69, 1197–1233.

Krishnamurthy, Arvind, and Annette Vissing-Jorgensen, 2011, The effects of quantitative easing on interest rates: channels and implications for policy, Discussion paper, National Bureau of Economic Research.

Kritzman, Mark, and Yuanzhen Li, 2010, Skulls, financial turbulence, and risk management, *Financial Analysts Journal* 66, 30–41.

———, Sebastien Page, and Roberto Rigobon, 2011, Principal components as a measure of systemic risk, *The Journal of Portfolio Management* 37, 112–126.

Lee, Jaehoon, 2013, Funding liquidity and its risk premium, .

Lindsten, Fredrik, Michael I Jordan, and Thomas B Schön, 2014, Particle gibbs with ancestor sampling., *Journal of Machine Learning Research* 15, 2145–2184.

Litterman, Robert, and Jose Scheinkman, 1991, Common factors affecting bond returns, *Journal of fixed income* 1, 54–61.

Ludvigson, Sydney C, and Serena Ng, 2009, Macro factors in bond risk premia, *The Review of Financial Studies* 22, 5027–5067.

Manela, Asaf, and Alan Moreira, 2017, News implied volatility and disaster concerns, *Journal of Financial Economics* 123, 137–162.

McLean, R David, and Jeffrey Pontiff, 2016, Does academic research destroy stock return predictability?, *The Journal of Finance* 71, 5–32.

Merton, Robert C, 1973, An intertemporal capital asset pricing model, *Econometrica: Journal of the Econometric Society* pp. 867–887.

Mikolov, Tomas, Ilya Sutskever, Kai Chen, Greg S Corrado, and Jeff Dean, 2013, Distributed representations of words and phrases and their compositionality, in *Advances in neural information processing systems* pp. 3111–3119.

Novy-Marx, Robert, 2013, The other side of value: The gross profitability premium, *Journal of Financial Economics* 108, 1–28.

Pan, Jun, and Kenneth J Singleton, 2008, Default and recovery implicit in the term structure of sovereign cds spreads, *The Journal of Finance* 63, 2345–2384.

Rehman, Zahid, and Grigory Vilkov, 2012, Risk-neutral skewness: Return predictability and its sources, .

Roll, Richard, Eduardo Schwartz, and Avanidhar Subrahmanyam, 2010, O/s: The relative trading activity in options and stock, *Journal of Financial Economics* 96, 1–17.

Stilger, Przemysław S, Alexandros Kostakis, and Ser-Huang Poon, 2016, What does risk-neutral skewness tell us about future stock returns?, *Management Science* 63, 1814–1834.

Stock, James H, and Mark W Watson, 1989, New indexes of coincident and leading economic indicators, *NBER macroeconomics annual* 4, 351–394.

Tibshirani, Robert, 1996, Regression shrinkage and selection via the lasso, *Journal of the Royal Statistical Society. Series B (Methodological)* pp. 267–288.

Vasquez, Aurelio, 2017, Equity volatility term structures and the cross section of option returns, *Journal of Financial and Quantitative Analysis* pp. 1–28.

Xing, Yuhang, Xiaoyan Zhang, and Rui Zhao, 2010, What does the individual option volatility smirk tell us about future equity returns?, *Journal of Financial and Quantitative Analysis* 45, 641–662.

Yan, Shu, 2011, Jump risk, stock returns, and slope of implied volatility smile, *Journal of Financial Economics* 99, 216–233.

Appendix A. Appendix: Two-Stage Learning of a Systemic Risk Factor

Appendix .1. Definition of Individual Systemic Risk Measures

Table 1: Variable Definition

Key Financial Systemic Risk Variables and Definition:
Absorption: fraction of return variance of a set of N financial institutions explained by the first k PCs
Aim: a weighted average of stock-level illiquidity measures
Book_lvg: aggregate book leverage for large financial institutions
Catfin : time-varying value at risk (VaR) of financial institutions at the 99% confidence level
Covar: value-at-risk (VaR) of the financial system as a whole conditional on an institution being in distress
Dci : how interconnected a set of financial institutions is by computing the fraction of significant Granger-causality relationships among their returns
Def_spr: baa bond yield minus aaa bond yield
Delta_absorption: difference between absorption ratios calculated for long and short estimation windows
Delta_covar: the difference between the conditional value at risk (CoVaR) of the financial system conditional on an institution being in distress and the CoVaR conditional on the median state
Gz: Gilchrist-Zakrajsek measure of credit spread
Intl_spillover: measures co-movement in macroeconomic variables across countries.
Mes : the expected return of a firm conditional on the system being in its lower tail
Mes_be: a version of marginal expected shortfall proposed by Brownlees and Engle (2011)
Mkt_lvg: aggregate market leverage for large financial institutions
Real_vol: average equity volatility of the largest financial institutions
Size_conc: Herfindal index of the size distribution among financial firms:
Ted_spr: 3-month LIBOR minus the 3-month T-bill rate
Term_spr: slope of Treasury Yield Curve
Turbulence: a measure of excess volatility that compares the realized squared returns of financial institutions with their historical volatility

Appendix .2. Factor Stochvol Model of 19 Individual Risk Measures

This section mainly follows Natesh's notes. Factor analysis is a method for investigating whether a number of variables of interest $Y_1, Y_2, ..., Y_d$ are related to a smaller number of unobservable factors $F_1, F_2, ..., F_k$ where $k \ll d$ (we use 3 in our analysis, similar to the PCA; worth adding one or two more factors). The fact that the factors are not observable makes it harder to use regression and other standard statistical models. Kannan (and others) use orthogonal regression. This seems interesting; we have not studied this yet for this data set, and comparing this with what we do is on our list of things to do next.

We introduce a model which performs factor analysis that aims at extracting the latent factors relating to the observed variables. This model, initiated by **?** is a multi-factor stochastic volatility model or simply, Factor Stochvol. We assume discrete time point $t \in \{1, \ldots, T\}$, although noting this assumption can be relaxed to incorporate varying time points. Denoting the observed variables of M dimensions at each time point t as $\mathbf{y}_t$, and the full observation as $\mathbf{Y}$, we have

$$\mathbf{y}_t := \begin{bmatrix} y_{1t} \\ y_{2t} \\ \vdots \\ y_{Mt} \end{bmatrix}, \qquad \mathbf{Y} := [\mathbf{y}_1\,, \mathbf{y}_2\,, \cdots, \mathbf{y}_T] = \begin{bmatrix} y_{11} & y_{12} & \cdots & y_{1T} \\ y_{21} & y_{22} & \cdots & y_{2T} \\ \vdots & \vdots & \ddots & \vdots \\ y_{M1} & y_{M2} & \cdots & y_{MT} \end{bmatrix}, \qquad t = 1, \ldots, T\,. \tag{A1}$$

The observation $\mathbf{y}_t$ follows the multivariate factor model with stochastic volatility:

$$\mathbf{y}_t \sim \text{Normal}(\mathbf{\Lambda f}_t\,, \mathbf{U}_t) \tag{A2}$$

$$\mathbf{f}_t \sim \text{Normal}(0, \mathbf{V}_t) \tag{A3}$$

where $\mathbf{\Lambda}$ an $M \times R$ matrix called the factor loadings and $\mathbf{f}_t$ is an $R \times 1$ vector of factors. The number of factors, R, is a parameter we choose. We have M is about 20 and we choose $R = 3$ (the data analysis showed that perhaps one or two factors might be present as well, but they contribute very little compared to the first three).

$$\mathbf{\Lambda} := \begin{bmatrix} \Lambda_{11} & \Lambda_{12} & \cdots & \Lambda_{1R} \\ \Lambda_{21} & \Lambda_{22} & \cdots & \Lambda_{2R} \\ \vdots & \vdots & \ddots & \vdots \\ \Lambda_{M1} & \Lambda_{M2} & \cdots & \Lambda_{MR} \end{bmatrix}, \qquad \mathbf{f}_t := \begin{bmatrix} f_{1t} \\ f_{2t} \\ \vdots \\ f_{Rt} \end{bmatrix}. \tag{A4}$$

We let the noises to be independent given a fix time t, thus the matrices $\mathbf{U}_t$ and $\mathbf{V}_t$ are diagonal matrices. In particular, we formulate them as

$$\mathbf{U}_t := \begin{bmatrix} \exp(h_{1t}) & 0 & \cdots & 0 \\ 0 & \exp(h_{2t}) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \exp(h_{Mt}) \end{bmatrix}, \ \dim = (M \times M), \tag{A5}$$

$$\mathbf{V}_t := \begin{bmatrix} \exp(h_{M+1,t}) & 0 & \cdots & 0 \\ 0 & \exp(h_{M+2,t}) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \exp(h_{M+R,t}) \end{bmatrix}, \ \dim = (R \times R), \tag{A6}$$

where the log variances (or also known as the latent volatilities) h_{it}, $i = 1, \ldots, M + R$, follow an autoregressive process of order one, *i.e.*, AR(1) process. This also means that each of the variances i.e. $\mathbb{V}[y_{mt}], \mathbb{V}[f_{rt}]$ respects the discretized Black-Karasinski model (**?**). The AR(1) process for each

h_{it} is given as

$$h_{it}\,|\,h_{i,t-1} \sim \text{Normal}\Big(\mu_i + \phi_i\,(h_{i,t-1} - \mu_i)\,, \sigma_i^2\Big)\,, \qquad i = 1, \ldots, M+R;\; t = 1, \ldots, T\,. \tag{A7}$$

where μ_i , ϕ_i , and σ_i are parameters to be learned, and we define h_{i0} as the initial log variance corresponding to time $t = 0$. Here, the starting value h_{i0} is treated as a parameter to be learned.

The inference to estimate the latent factors is via Bayesian inference based on a set of carefully selected proper priors:

$$\mu_i \sim \text{Normal}(b_\mu, B_\mu),\; \frac{(\phi_i + 1)}{2} \sim \text{Beta}(a_0, b_0)\,, \sigma_i^2 \sim \text{Gamma}\Big(\frac{1}{2}, \frac{1}{2B_\sigma}\Big)\,, i = 1, \ldots, M+R. \tag{A8}$$

Further choose the conditional prior for starting values h_{i0} , which is the stationary distribution of the AR(1) process:

$$h_{i0}\,|\,\mu_i, \phi_i, \sigma_i \sim \text{Normal}\big(\mu_i\,, \sigma_i^2/(1-\phi_i^2)\big)\,. \tag{A9}$$

Finally, for each of the elements in the loading factors $\mathbf{\Lambda}$, we choose a conjugate zero-mean Gaussian distribution prior, that is,

$$\Lambda_{ij} \sim (0, B_\Lambda)\,, \qquad i = 1, \ldots, M;\; j = 1, \ldots, R\,. \tag{A10}$$

For the estimation algortihm, we refer to ?. Running the `factorstochvol.R` script on the Risk dataset to determine the unobserved factors yields factor loadings of three latent factors based on Factor Stochastic Volatility model In Figure 19,

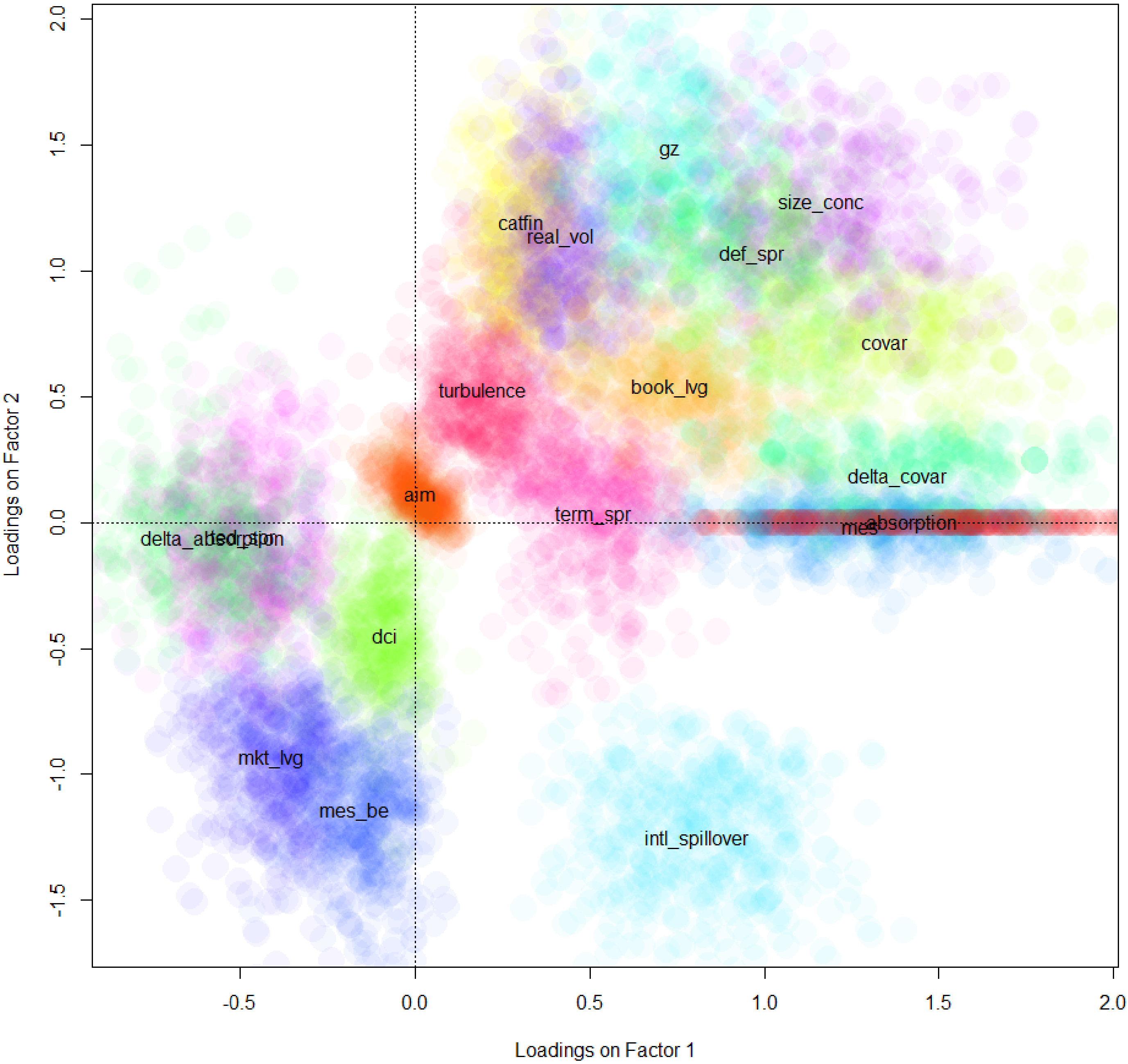


Figure 19. This figure plots factor loadings of all the systemic risk measures.

Appendix .3. Principal Component Analysis of 19 Individual Systemic Measures

The first step is to construct PCA of all individual measures. As is standard practice in PCA, we demean the data and standardize it to have unit empirical variance. We describe PCA analysis a little bit here. Recall that the eigenvectors of covariance $\mathbf{Y}\mathbf{Y}^T$ are closely connected to the singular value decomposition (SVD) of $\mathbf{Y} \in \mathbb{R}^{d\times n}$. First, recall the eigendecomposition of the real and symmetric matrix $\mathbf{Y}\mathbf{Y}^T$:

$$\mathbf{Y}\mathbf{Y}^T = \mathbf{P}\mathbf{\Lambda}\mathbf{P}^T, \tag{A11}$$

where $\mathbf{P}^T\mathbf{P} = \mathbf{I}$ and $\mathbf{\Lambda}$ is a diagonal with non-increasing entries. Next recall that the singular value decomposition (SVD) of any $\mathbf{Y} \in \mathbb{R}^{d\times n}$,

$$\mathbf{Y} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T \tag{A12}$$

where $\mathbf{U} \in \mathbb{R}^{d\times d}, \mathbf{\Sigma} \in \mathbb{R}^{d\times n}, \mathbf{V} \in \mathbb{R}^{n\times n}$, with $\mathbf{U}^T\mathbf{U} = \mathbf{V}^T\mathbf{V} = \mathbf{I}$ and $\mathbf{\Sigma}$ is a square diagonal matrix. Using the SVD, we proceed to find that:

$$\mathbf{Y}\mathbf{Y}^T = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T(\mathbf{U}\mathbf{\Sigma}\mathbf{V}^T)^T = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T\mathbf{V}\mathbf{\Sigma}^T\mathbf{U}^T = \mathbf{U}(\mathbf{\Sigma}\mathbf{\Sigma}^T)\mathbf{U}^T \text{ since } \mathbf{V}^T\mathbf{V} = \mathbf{I}. \tag{A13}$$

Further note that $\mathbf{\Sigma}\mathbf{\Sigma}^T \in \mathbb{R}^{n\times n}$ is a diagonal matrix. Thus this has *precisely* the same form as the eigendecomposition of $\mathbf{Y}\mathbf{Y}^T$, c.f. with equation (A11), i.e. set $\mathbf{P} = \mathbf{U}, \mathbf{\Lambda} := \mathbf{\Sigma}\mathbf{\Sigma}^T$. It follows that the left singular vectors $\mathbf{U}$ of $\mathbf{Y}$ will be the eigenvectors of $\mathbf{Y}\mathbf{Y}^T$, with corresponding eigenvalues $\mathbf{\Sigma}\mathbf{\Sigma}^T$. Now let $\mathbf{U}_k \in \mathbb{R}^{d\times k}$ be a truncated version of $\mathbf{U}$, where we just take the first k columns. Then $\mathbf{U}_k \in \mathbb{R}^{d\times k}$ equivalently consists of an orthonormal basis for the eigenspace of the k eigenvalues of $\mathbf{Y}\mathbf{Y}^T$. So the solution to the PCA problem is expressible as: Eigenvectors : $\hat{\mathbf{W}} = \mathbf{U}_k$, Projection :

$\hat{\mathbf{Z}} = \mathbf{U}_k^T\mathbf{Y}$, Reconstruction : $\hat{\mathbf{Y}} = \mathbf{U}_k\mathbf{U}_k^T\mathbf{Y}$, Reconstruction error : $\hat{\mathbf{RE}} = ||\hat{\mathbf{Y}} - \mathbf{Y}||_F^2$ where $||\cdot||_F$ denotes the Frobenius norm.

Figure 10 shows variances explained by first 10 principal components extracted out of cross-section. The first PCA explains 25% of variations in all series and the second PCA explains nearly 23% of total variations. All together,the first two PCAs explain almost 50% of total variations. So we will focus on the first two PCAs in the remainder of this paper. Although we plot PCA loadings on each specific risk measure in Figure 20, it is hard to give economic explanation for each PCA. However, if we can project PCs onto news information, there would be textual explanation for those latent factors. This is part of reasons why we bring text data into picture.

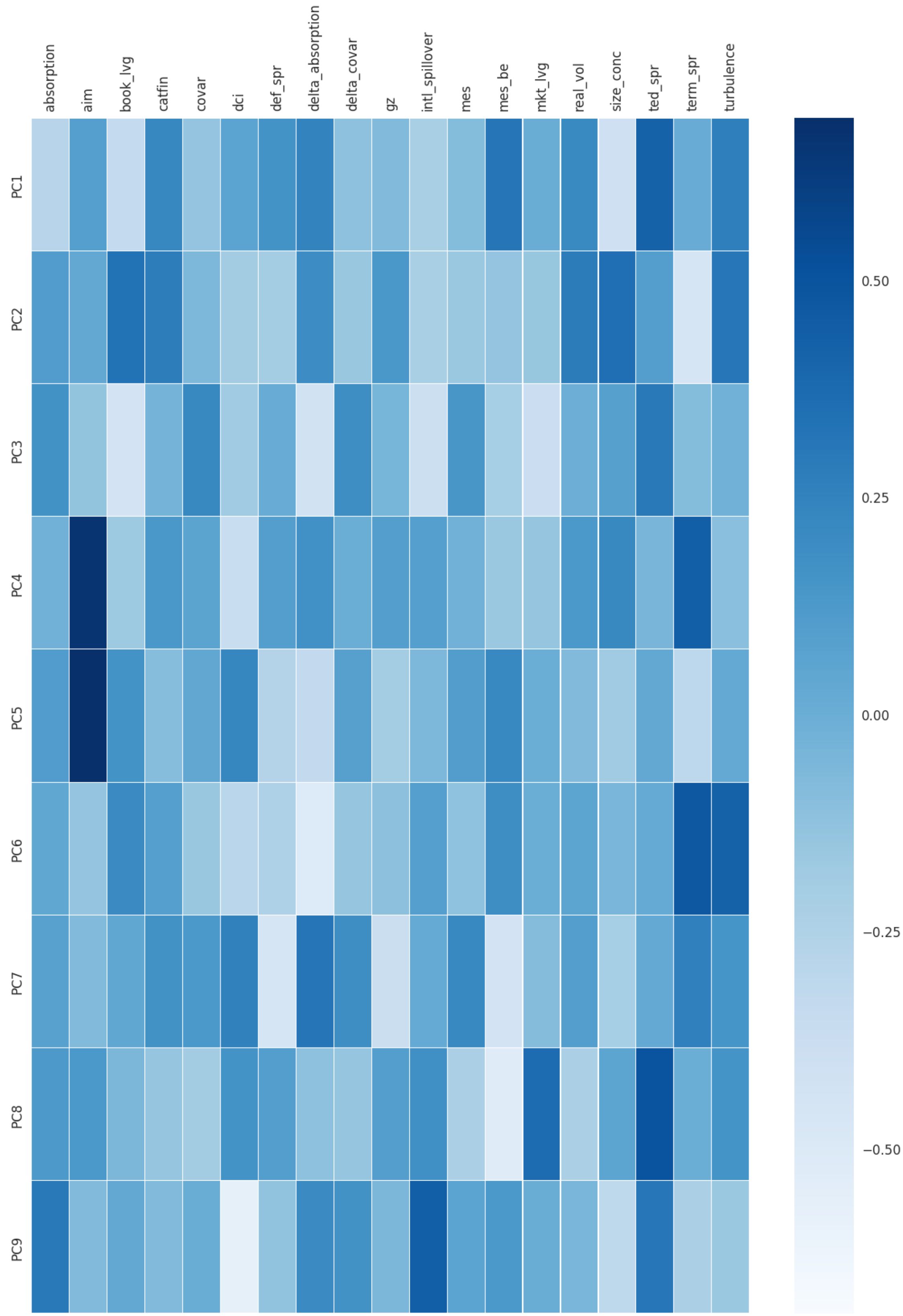


Figure 20. This figure plots heat map of all principal components.

Table 2: Market Price of PC1

Reported are market price risk for PC1 across different portfolios. Each row is estimated as $E[R^e] = \lambda_0 + \beta_{fac}\lambda_{fac}$. $E[R^e]$ are expected excess returns for each portfolios. VWM^e, SMB and HML are Fama-French three factors. HKM denotes intermediary capital factor from He, Kelly and Manela(2016). ΔPC1 is the first difference of our Principal Component 1. The first row estimates market price of risk using monthly Fama-French 25 Portfolios Formed on Size and Book-to-Market from Jan 1995 to Dec 2009. The second row estimates market price of risk using monthly Fama-French 25 Portfolios Formed on Size and Momentum. The third row estimates market price of risk using both Fama-French 25 Portfolios Formed on Size and Book-to-Market and Fama-French 10 Portfolios Formed on Momentum from Jan 1995 to Dec 2009. J statistic that tests whether the pricing errors are jointly zero is shown in the last column of each row. Fama-MacBeth t-statistics are reported in parenthesis and $*$, $**$, and $***$ indicate 10%, 5%, and 1% significance levels, respectively.

	VWM^e	SMB	HML	HKM	ΔPC1	
25SizeBe	3.602***	2.504**	5.039***	0.256***	−1.409**	J-stat: 74.8
	(2.827)	(2.403)	(5.023)	(6.997)	(−2.319)	(0.000)
	VWM^e	SMB	HML	HKM	ΔPC1	
25SizeMoM	3.383***	3.977***	6.258***	−0.321***	3.564***	J-stat: 58.4
	(2.626)	(3.638)	(5.254)	(−6.052)	(5.629)	(0.000)
	VWM^e	SMB	HML	HKM	ΔPC1	
35SizeBeMoM	3.445***	3.066***	4.417***	−0.151***	0.188	J-stat: 91.6
	(2.707)	(2.926)	(4.372)	(−3.224)	(0.325)	(0.000)

Appendix .4. Market Price of Risk

Table 3: Market Price of Textual PC1

Reported are market price risk for the Textual PC1 across different portfolios. Each row is estimated as $E[R^e] = \lambda_0 + \beta_{fac}\lambda_{fac}$. $E[R^e]$ are expected excess returns for each portfolios. VWM^e, SMB and HML are Fama-French three factors. HKM denotes intermediary capital factor from He, kelly and Manela(2016). ΔTPC1 is the first difference of our Textual Based Principal Component 1. The first row estimates market price of risk using monthly Fama-French 25 Portfolios Formed on Size and Book-to-Market from Jan 1995 to Dec 2009. The second row estimates market price of risk using monthly Fama-French 25 Portfolios Formed on Size and Momentum. The third row estimates market price of risk using both Fama-French 25 Portfolios Formed on Size and Book-to-Market and Fama-French 10 Portfolios Formed on Momentum from Jan 1995 to Dec 2009. J statistic that tests whether the pricing errors are jointly zero is shown in the last column of each row. Fama-MacBeth t-statistics are reported in parenthesis and $*$, $**$, and $***$ indicate 10%, 5%, and 1% significance levels, respectively.

	VWM^e	SMB	HML	HKM	ΔTPC1	
25SizeBe	3.803***	2.361**	5.201***	0.263***	−1.494***	J-stat: 72.8
	(2.988)	(2.284)	(5.231)	(7.221)	(−3.619)	(0.000)
	VWM^e	SMB	HML	HKM	ΔTPC1	
25SizeMoM	2.867**	4.237***	6.926***	−0.291***	2.633***	J-stat: 58.2
	(2.212)	(3.880)	(5.727)	(−5.426)	(4.987)	(0.000)
	VWM^e	SMB	HML	HKM	ΔTPC1	
35SizeBeMoM	3.392***	2.983***	4.605***	−0.147***	1.139**	J-stat: 91.1
	(2.661)	(2.859)	(4.623)	(−3.226)	(2.551)	(0.000)

Appendix .5. Predictability of Time Series of Government Bond Excess Returns

Table VI shows monthly bond return predictability regressions based on TPC1, TPC2, original PC1 , original PC2 and Bayesian Factor. The dependent variables are annualized one year holding period returns on three year US government bond (Results are robust to different US government bonds). Each column represents a different regression. The in sample columns examine period from 1990.1 to 2006.12, and the out of sample columns focus on the financial crisis period from 2007.1 to 2014.12. t_{NW} are Newey and West corrected t-statistics with 24 lags. $*$, $**$, and $***$ indicate 10%, 5%, and 1% significance levels, respectively. Panel A and Panel B show results for PCAs and Bayesian Factor respectively. In Panel A, before financial crisis, neither of textual PCs nor original PCs have any significant effects on future bond risk premia while CP factor has very significant effect on bond risk premia. In contrast, after financial crisis, CP factor can not affect significantly future bond risk premia while systemic risk measures can predict future bond risk premia very significantly. More interestingly, textual PCs have more predictive power than original PCs. For example, the adj R^2 of regressing bond risk premia onto Textual PC2 is 27.1%, which is much larger than the adj R^2 5% from regression of bond risk premia onto PC2. This means that news embed additional information beyond number information and combing news and number information gives a larger predictive power for future bond risk premia. In Panel B, Bayesian Factor shows predictive power not only after financial crisis but also before financial crisis. This is very important because a good pricing kernel measure should have pricing power both in crisis period and normal period and our Bayesian Factor satisfies this requirement. Although we didn't show the result, PQR factor does not have a predictive power for future bond excess returns in normal period.

Textual Code: Data: CitiGroup AND NOT Morning Agenda AND NOT Paid Notice AND NOT Profit Scoreboard AND NOT In Quote AND NOT Residential Sales AND NOT Treasury Auction AND NOT Morning Takeout AND NOT Wedding AND NOT Business Digest AND NOT Executive Change AND NOT Brief

Appendix B. Appendix: Combining Bond Demand and Supply Factors

Appendix C. Appendix: Making Better Use of Option Prices to Predict Stock Returns

Figure 1: Principal Components of Implied Volatility Surfaces

This figure plots the relative contribution of the first ten principal components of the correlation matrix of the cross-section of implied volatility surfaces. Every Monday, we extract the implied volatility surfaces for all

Table 4: Bond Predictability of PCs, Textual PCs, and Bayesian Factor

Reported are monthly bond return predictability regressions based on Financial Distress Factor. Panel A show regressions for Textual PC1, Textual PC2, PC1 and PC2. Panel B show regressions for Bayesian Factor. The dependent variables are annualized one year holding period returns on three year US government bond. The independent variables are each distress measure plus CP factor. In Panel B, we did not report CA factor estimation. Each column represents a different regression. The in sample columns examine period from 1990.1 to 2006.12, and the out of sample columns focus on the financial crisis period from 2007.1 to 2014.12. t_{NW} are Newey and West corrected t-statistics with 24 lags. $*$, $**$, and $***$ indicate 10%, 5%, and 1% significance levels, respectively.

$$r_{t+1}^{bond} = \beta_0 + \beta_1 X_t + CP_t + \epsilon_{t+1}$$

Panel A	In Sample :1990 - 2006				Out of Sample 2007-2014			
TPC1	-0.0035				0.0230***			
t_{NW}	(-0.531)				(2.586)			
TPC2		0.0037				0.0309***		
t_{NW}		(0.755)				(4.785)		
PC1			-0.0028				0.009***	
t_{NW}			(-0.743)				(2.838)	
PC2				0.0055				0.0026
t_{NW}				(1.422)				(0.412)
CP	0.0126***	0.0131***	0.0126***	0.0135***	0.0001	0.0044	0.013	0.0102
t_{NW}	(5.137)	(5.722)	(5.095)	(6.096)	(0.397)	(0.939)	(1.647)	(1.452)
Adj. R^2	36.7%	36.5%	36%	38%	21.5%	27.1%	19%	5%
Obs.	204	204	204	204	96	96	96	96

Panel B	In Sample :1990 - 2006	Total Sample :1990 - 2014
Bayesian Factor	-0.0977***	-0.0935***
t_{NW}	(-2.257)	(-4.927)
Adj. R^2	38.6%	38.5%
Obs.	204	300

Table 5 Macro Variable Description

The column tcode denotes the following data transformation for a series x: (1) no transformation; (2)Δx_t ; (3) $\Delta^2 x_t$; (4) $log(x_t)$; (5) $\Delta log(x_t)$; (6) $\Delta^2 log(x_t)$. (7) $\Delta(x_t/x_{t1}1.0)$. The FRED column gives mnemonics in FRED followed by a short description. The comparable series in Global Insight is given in the column GSI.

firms from Optionmetrics from 1996 to 2014. Each implied volatility surface contains 112 volatilities defined across calls and puts, absolute deltas of 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, and 0.8 at maturities 30, 60, 91, 122, 152, 182, 273, and 365 calendar days. After demeaning each volatility surface, we compute the correlation matrix of the cross-section of implied volatility surfaces and decompose it using principal component analysis.

(2).png

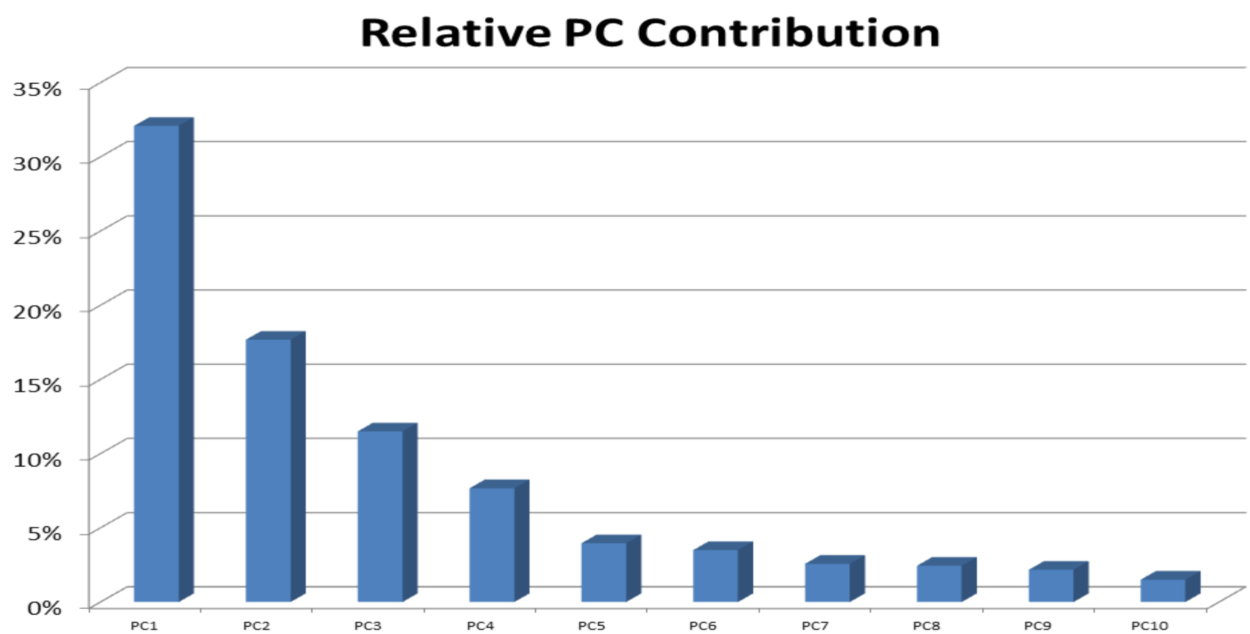
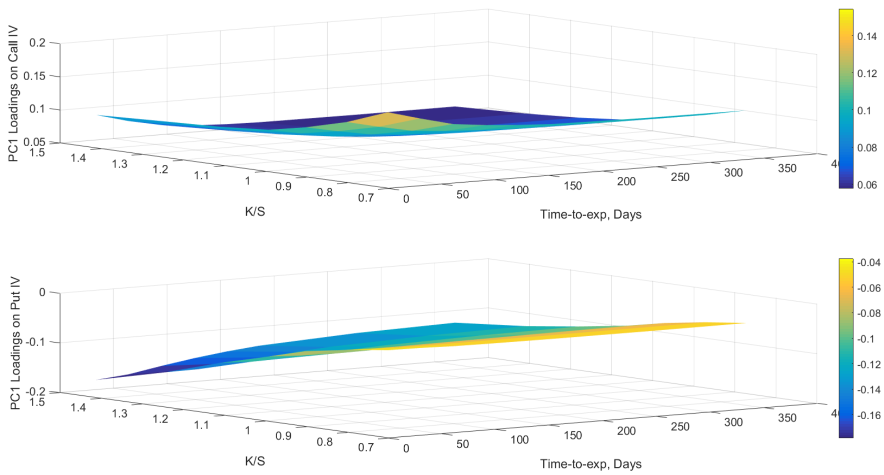


Figure 2: Principal Components of Implied Volatility Surfaces

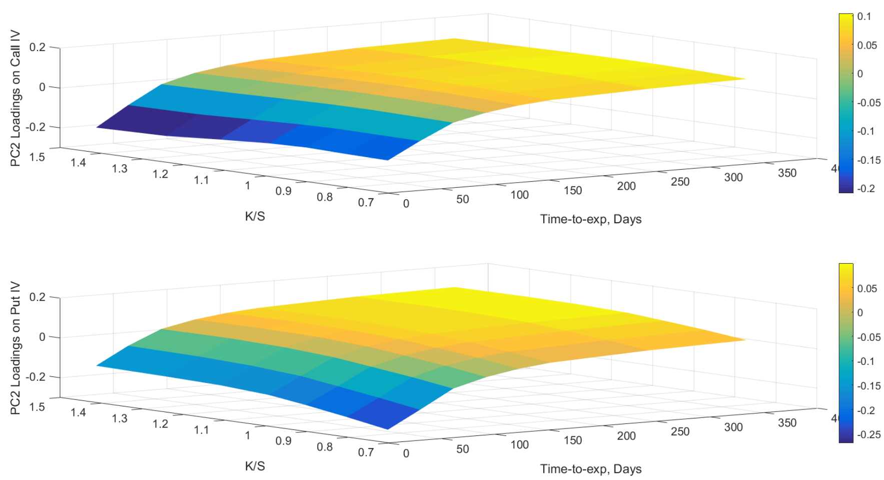
This figure plots the loadings for the first five principal components of the correlation matrix of the cross-section of implied volatility surfaces across maturity (in days) and moneyness (Strike over stock price). Every Monday, we extract the implied volatility surfaces for all firms from Optionmetrics from 1996 to 2014. Each implied volatility surface contains 112 volatilities defined across calls and puts, absolute deltas of 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, and 0.8 at maturities 30, 60, 91, 122, 152, 182, 273, and 365 calendar days. After demeaning each volatility surface, we compute the correlation matrix of the cross-section of implied volatility surfaces and decompose it using principal component analysis. Panel A to E report the loadings of the first five principal components (out of 112).

Panel A: First Principal Component (PC1)



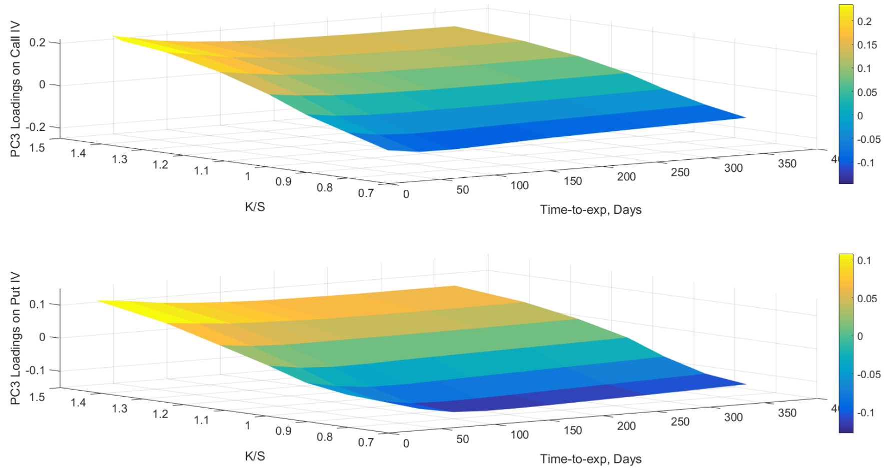
(2).png

Panel B: Second Principal Component (PC2)



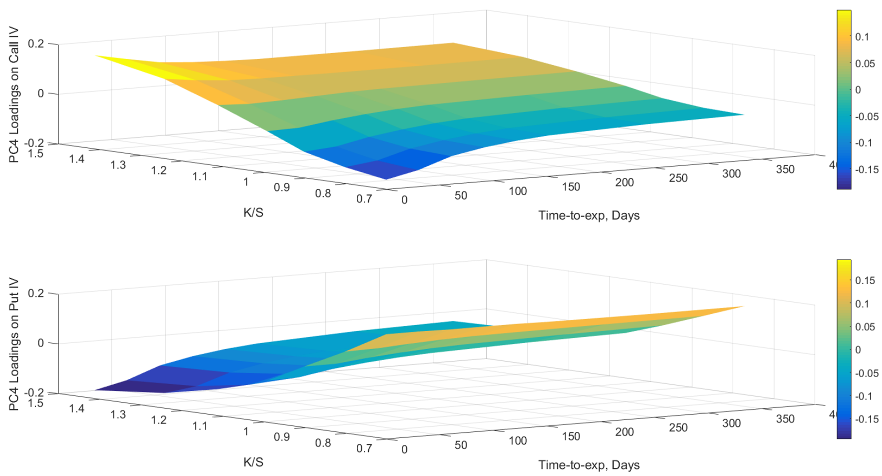
(2).png

Panel C: Third Principal Component (PC3)



(2).png

Panel D: Fourth Principal Component (PC4)



Panel E: Fifth Principal Component (PC5)

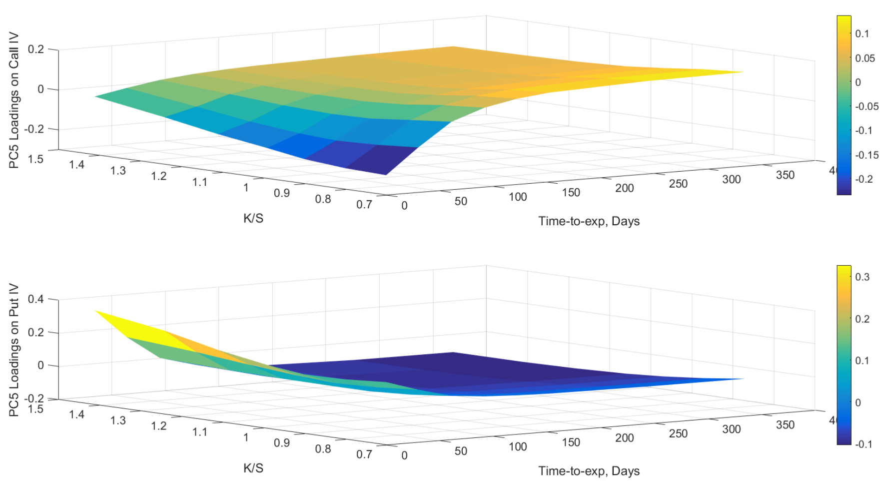
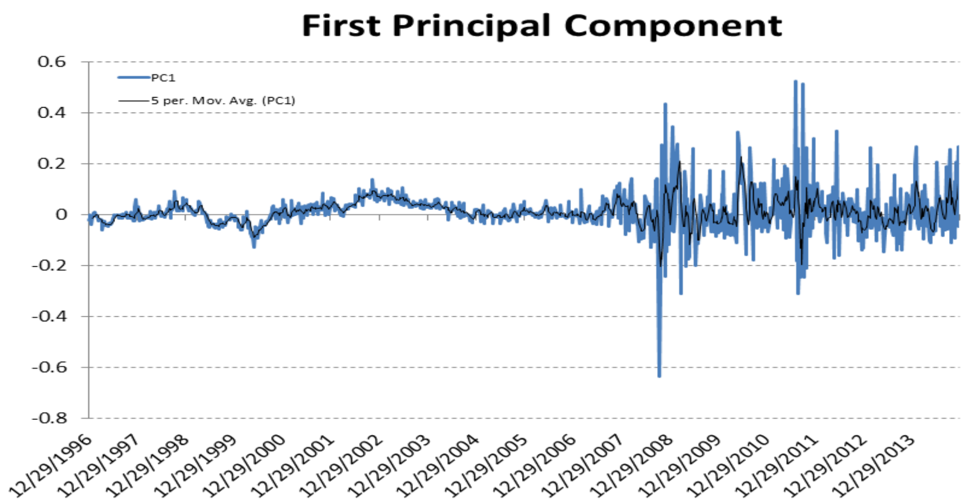


Figure 3: Time-series Average of Principal Components of Implied Volatility Surfaces

This figure plots the time-series average of the first three principal components (PC1, PC2, and PC3) of the correlation matrix of the cross-section of implied volatility surfaces. Every Monday, we extract the implied volatility surfaces for all firms from Optionmetrics from 1996 to 2014. Each implied volatility surface contains 112 volatilities defined across calls and puts, absolute deltas of 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, and 0.8 at maturities 30, 60, 91, 122, 152, 182, 273, and 365 calendar days. After demeaning each volatility surface, we compute the correlation matrix of the cross-section of implied volatility surfaces and decompose it using principal component analysis. Panel A reports the time-series average of the first principal component (out of 112). Panel B reports the time-series average of the second and third principal components (out of 112).

Panel A: Average First Principal Component



Panel B: Average Second and Third Principal Component

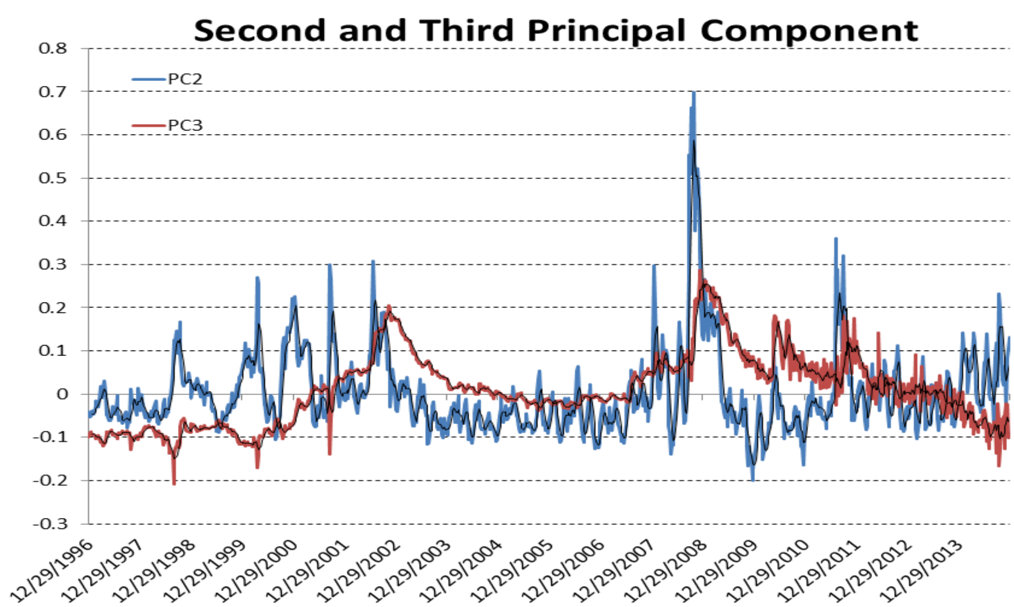
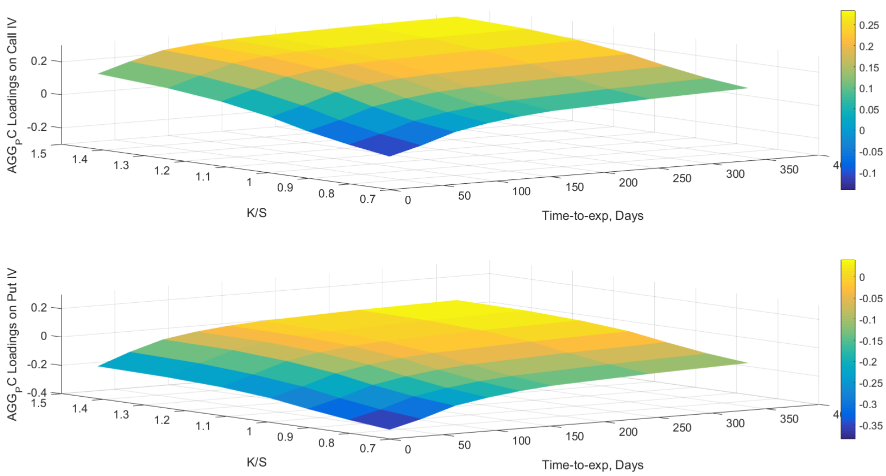


Figure 4: Aggregate PC Factor

This figure plots the loadings of the aggregate principal component (PC) factor across maturity (in days) and moneyness (Strike over stock price) in Panel A and the time-series average (along with the 5-day moving average) of the aggregate PC factor in Panel B. To construct the aggregate PC factor, every Monday we extract the implied volatility surfaces for all firms from Optionmetrics from 1996 to 2014. Each implied volatility surface contains 112 volatilities defined across calls and puts, absolute deltas of 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, and 0.8 at maturities 30, 60, 91, 122, 152, 182, 273, and 365 calendar days. After demeaning each volatility surface, we compute the correlation matrix of the cross-section of implied volatility surfaces and decompose it using principal component analysis. The agregate PC factor is computed with the first three PCs and is equal to $PC1 + PC2 + PC3$.

Panel A: Aggregate PC Factor Loadings



Panel B: Time-Series Average

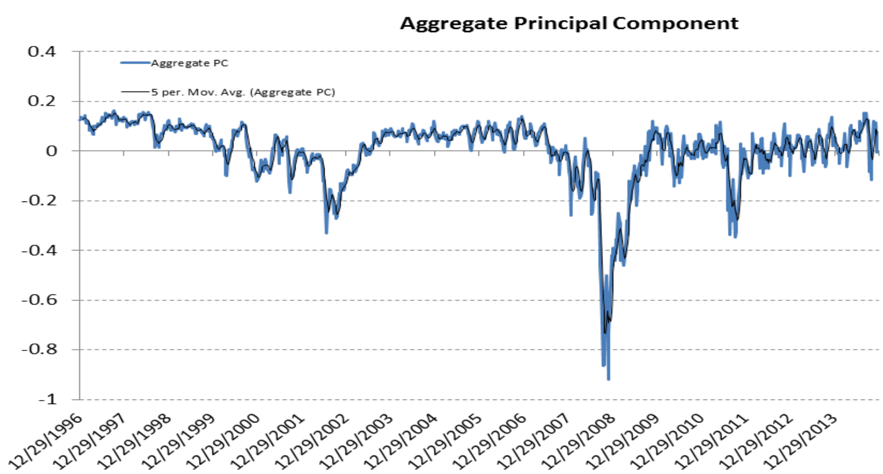
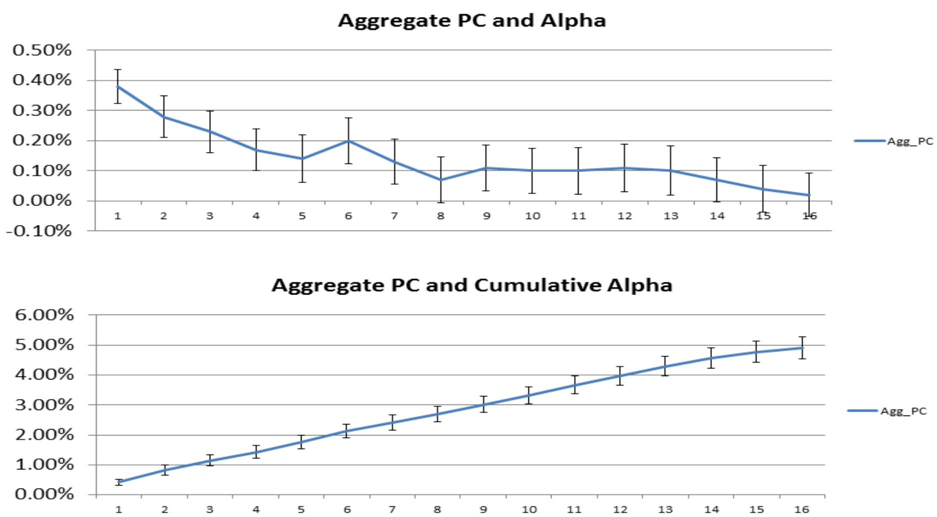


Figure 5: Long Term Return Predictability of the Aggregate PC Factor

This figure plots the long-term return predictability of the aggregate principal component (PC) factor. Panel A reports the economic significance of the aggregate PC factor in predicting week t value-weighted return and Panel B reports cumulative week t value-weighted returns, where t goes from 1 to 30 weeks. The surrounding error bars represent the 95% confidence interval. To construct the aggregate PC factor, every Monday we extract the implied volatility surfaces for all firms from Optionmetrics from 1996 to 2014. Each implied volatility surface contains 112 volatilities defined across calls and puts, absolute deltas of 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, and 0.8 at maturities 30, 60, 91, 122, 152, 182, 273, and 365 calendar days. After demeaning each volatility surface, we compute the correlation matrix of the cross-section of implied volatility surfaces and decompose it using principal component analysis. The agregate PC factor is computed with the first three PCs and is equal to $PC1 + PC2 + PC3$.

Panel A: Future Weekly Returns



Panel B: Cummulative Future Weekly Returns

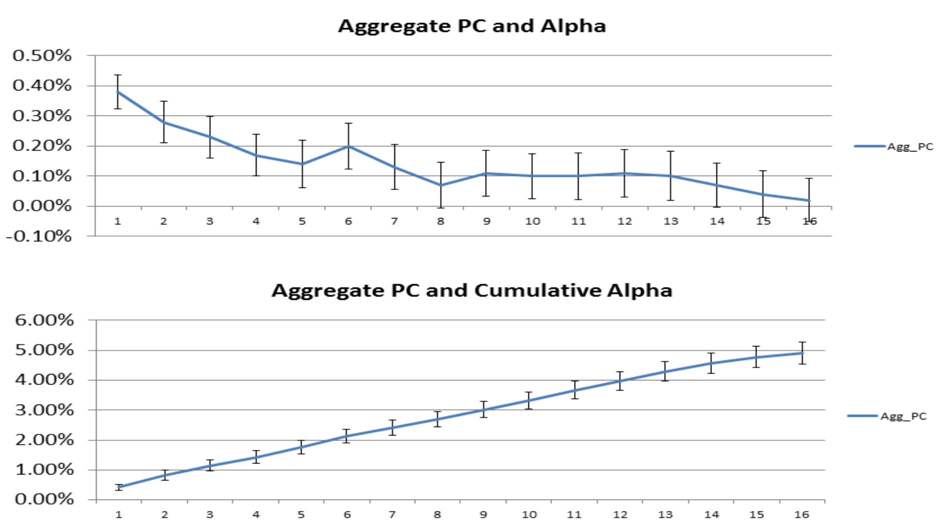


Table 1: Summary Statistics

This table reports summary statistics of the variables. We report the first five principal components (out of 112), $PC1$ to $PC5$, of the cross-section of volatility surfaces. Every Monday, we extract the implied volatility surfaces for all firms from Optionmetrics from 1996 to 2014. Each implied volatility surface contains 112 volatilities defined across calls and puts, absolute deltas of 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, and 0.8 at maturities 30, 60, 91, 122, 152, 182, 273, and 365 calendar days. After demeaning each volatility surface, we compute the correlation matrix of the cross-section of implied volatility surfaces and decompose it using principal component analysis. We report an agregate PC factor that is equal to $PC1 + PC2 + PC3$, implied volatility skew as in Xing, Zhang, and Zhao (2010), implied volatility spread as in Cremers and Weinbaum (2010), the slope of the implied volatility term structure defined as the difference between long and short term ATM implied volatilities, volatility of implied volatilities as in Baltussen, Van Bekkum, and Van Der Grient (2012), risk-neutral moments as defined in Conrad, Dittmar, and Ghysels (2013), O/S option to stock volume ration as defined in Johnson and So (2012), contemporaneous weekly return (Ret_1w), lagged six-month return (Ret_6M), the logarithm of the market-capitalization (Size), book-to-market (BE/ME), and the Amihud measure of illiquidity.

Variables	Mean	Std	5% Percentile	50%Percentile	95%Percentile
PC1	0.000	0.287	-0.292	0.019	0.256
PC2	0.000	0.246	-0.401	0.044	0.243
PC3	0.000	0.179	-0.218	-0.006	0.248
PC4	0.000	0.136	-0.149	-0.009	0.185
PC5	0.000	0.096	-0.092	-0.016	0.153
Agg. PC Factor	0.000	0.419	-0.547	0.048	0.395
IVskew	0.055	0.074	-0.009	0.040	0.182
IVspread	-0.008	0.067	-0.078	-0.005	0.055
SlopeIV_TS	-0.029	0.079	-0.149	-0.017	0.046
VoV	0.085	0.065	0.028	0.067	0.207
RNvol	0.469	0.153	0.253	0.452	0.745
RNskew	-0.361	0.306	-0.893	-0.336	0.087
RNkurt	3.064	0.649	2.193	3.002	4.150
O/S	3.374	7.707	0.000	1.000	13.907
Ret_1w	0.002	0.065	-0.089	0.000	0.099
Ret_6M	0.151	0.801	-0.571	-0.002	1.325
Size	7.261	1.584	4.927	7.105	10.134
BE/ME	0.549	0.634	0.067	0.447	1.370
Amihud	0.032	0.345	0.000	0.003	0.075

Table 2: Correlation Matrix

This table reports the correlation matrix of the principal components and firm variables. We report the first five principal components (out of 112), $PC1$ to $PC5$, of the cross-section of volatility surfaces. Every Monday, we extract the implied volatility surfaces for all firms from Optionmetrics from 1996 to 2014. Each implied volatility surface contains 112 volatilities defined across calls and puts, absolute deltas of 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, and 0.8 at maturities 30, 60, 91, 122, 152, 182, 273, and 365 calendar days. After demeaning each volatility surface, we compute the correlation matrix of the cross-section of implied volatility surfaces and decompose it using principal component analysis. The remaining variables are an agregate PC factor that is equal to $PC1 + PC2 + PC3$, implied volatility spread as in Cremers and Weinbaum (2010), implied volatility skew as in Xing, Zhang, and Zhao (2010), risk-neutral skewness and risk-neutral kurtosis as defined in Conrad, Dittmar, and Ghysels (2013), O/S option to stock volume ration as defined in Johnson and So (2012), contemporaneous weekly return (Ret_1w), lagged six-month return (Ret_6M), the market-capitalization (Size), book-to-market (BE/ME), and the Amihud measure of illiquidity.

	PC1	PC2	PC3	PC4	PC5	Agg. PC Factor	IVskew	IVspread	RNskew	RNkurt	O/S	Ret_1w	Ret_6M	Size	BE/ME	Amihud
PC2	0.9															
PC3	-1.4	-4.2														
PC4	1.0	1.7	-1.7													
PC5	-0.9	-2.7	0.7	-6.8												
Agg. PC Factor	71.1	-55.6	-40.8	0.4	0.6											
IVskew	-39.2	16.4	34.4	13.5	22.1	-49.7										
IVspread	66.1	-8.5	-2.1	-9.1	-3.7	51.7	-56.3									
RNskew	4.7	9.8	-23.2	-16.0	-0.4	6.1	-31.5	9.1								
RNkurt	0.7	-9.6	1.9	10.3	5.5	5.9	11.0	0.7	-69.7							
O/S	-8.1	3.3	7.3	-7.2	-3.9	-10.8	8.0	-6.5	-9.2	8.2						
Ret_1w	-5.7	-2.0	1.0	2.0	0.0	-3.3	5.9	-10.7	-2.5	1.7	1.4					
Ret_6M	-0.8	0.0	3.6	-1.3	-3.1	-2.1	-2.3	-0.6	5.3	-5.4	5.3	0.0				
Size	3.6	-15.8	16.9	-6.2	-10.2	4.6	-4.3	4.1	-38.2	34.1	18.6	3.4	3.0			
BE/ME	3.2	5.1	-4.8	0.9	6.9	1.4	9.8	1.2	1.3	-2.6	-6.6	-4.6	-10.9	-14.3		
Amihud	-6.5	1.6	-2.8	1.3	-2.6	-4.4	6.5	-7.2	3.9	-4.6	-0.5	0.0	5.9	-7.3	-0.9	

Table 3: Fama-Macbeth Regressions and Principal Components

This table presents the Fama-Macbeth results from regressing next-week and next-month stock returns on the first five principal components (out of 112), $PC1$ to $PC5$, of the cross-section of volatility surfaces. Every Monday, we extract the implied volatility surfaces for all firms from Optionmetrics from 1996 to 2014. Each implied volatility surface contains 112 volatilities defined across calls and puts, absolute deltas of 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, and 0.8 at maturities 30, 60, 91, 122, 152, 182, 273, and 365 calendar days. After demeaning each volatility surface, we compute the correlation matrix of the cross-section of implied volatility surfaces and decompose it using principal component analysis. The remaining variables are an agregate PC factor that is equal to $PC1 + PC2 + PC3$, the slope of the implied volatility term structure defined as the difference between long and short term ATM implied volatilities, implied volatility spread as in Cremers and Weinbaum (2010), implied volatility skew as in Xing, Zhang, and Zhao (2010), O/S option to stock volume ration as defined in Johnson and So (2012), contemporaneous weekly return (Ret_1w), lagged six-month return (Ret_6M), the market-capitalization (Size), book-to-market (BE/ME), and the Amihud measure of illiquidity.

	Weekly Return			Monthly Return		
	1	2	3	4	5	6
Intercept	0.0018	0.0016	0.0019	0.0071*	0.0089	0.0096
	(1.49)	(0.61)	(0.71)	(1.75)	(0.98)	(1.05)
PC1	0.0041***	0.0041***	0.0037***	0.0120***	0.0120***	0.0127***
	(7.41)	(8.31)	(5.11)	(7.22)	(7.96)	(5.65)
PC2	0.0031***	0.0035***	0.0018	0.0101***	0.0117***	0.0068**
	(3.62)	(5.23)	(1.48)	(3.57)	(5.40)	(1.99)
PC3	0.0036***	0.0031***	0.0028***	0.0111***	0.0085***	0.0083**
	(3.70)	(3.17)	(2.89)	(3.34)	(2.60)	(2.56)
PC4	-0.0002	-0.0006	-0.0011	0.0031	0.0014	-0.0002
	(-0.15)	(-0.46)	(-0.86)	(0.58)	(0.32)	(-0.04)
PC5	-0.0025*	-0.0032***	-0.0040***	-0.0110**	-0.0135***	-0.0168***
	(-1.65)	(-2.86)	(-3.31)	(-2.18)	(-3.69)	(-4.16)
SlopeIV_TS			0.0060*			0.0171*
			(1.81)			(1.78)
IVskew			-0.0009			-0.0022
			(-0.43)			(-0.32)
IVspread			0.0022			-0.0045
			(0.58)			(-0.38)
O/S			-0.0001			-0.0002
			(-1.58)			(-1.43)
Ret_1w	-0.0193***	-0.0204***	-0.0204***	-0.0235***	-0.0277***	-0.0280***
	(-5.40)	(-6.30)	(-6.35)	(-2.80)	(-3.84)	(-3.92)
Ret_6M		-0.0002	-0.0002		-0.0017	-0.0015
		(-0.45)	(-0.39)		(-0.95)	(-0.88)
Size		-0.0000	-0.0000		-0.0005	-0.0004
		(-0.21)	(-0.07)		(-0.63)	(-0.47)
B/M		0.0008*	0.0007		0.0022	0.0020
		(1.69)	(1.60)		(1.24)	(1.15)
Amihud		-0.0021	-0.0022		-0.0076	-0.0084
		(-0.35)	(-0.36)		(-0.37)	(-0.40)
$Adj.R^2$	0.0245***	0.0500***	0.0532***	0.0253***	0.0539***	0.0575***

Table 4: Fama-Macbeth Regressions and the Aggregate PC Factor

This table presents the Fama-Macbeth results from regressing next-week and next-month stock returns on the aggregate principal component (PC) factor. To construct the aggregate PC factor, every Monday we extract the implied volatility surfaces for all firms from Optionmetrics from 1996 to 2014. Each implied volatility surface contains 112 volatilities defined across calls and puts, absolute deltas of 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, and 0.8 at maturities 30, 60, 91, 122, 152, 182, 273, and 365 calendar days. After demeaning each volatility surface, we compute the correlation matrix of the cross-section of implied volatility surfaces and decompose it using principal component analysis. The agregate PC factor that is equal to $PC1 + PC2 + PC3$. The remaining variables are the agregate PC factor that is equal to $PC1 + PC2 + PC3$, the slope of the implied volatility term structure defined as the difference between long and short term ATM implied volatilities, implied volatility spread as in Cremers and Weinbaum (2010), implied volatility skew as in Xing, Zhang, and Zhao (2010), O/S option to stock volume ration as defined in Johnson and So (2012), contemporaneous weekly return (Ret_1w), lagged six-month return (Ret_6M), the market-capitalization (Size), book-to-market (BE/ME), and the Amihud measure of illiquidity.

	Weekly Return			Monthly Return		
	1	2	3	4	5	6
Intercept	0.0017	0.0016	0.0015	0.0071*	0.0082	0.0080
	(1.48)	(0.60)	(0.57)	(1.76)	(0.91)	(0.87)
Agg. PC Factor	0.0035***	0.0036***	0.0031***	0.0115***	0.0118***	0.0117***
	(5.98)	(7.10)	(4.42)	(5.84)	(6.94)	(4.84)
SlopeIV_TS			0.0001			-0.0004
			(0.05)			(-0.05)
IVskew			-0.0018			-0.0015
			(-0.90)			(-0.24)
IVspread			0.0043			-0.0085
			(1.24)			(-0.72)
O/S			-0.0001			-0.0002
			(-1.57)			(-1.45)
Ret_1w	-0.0195***	-0.0209***	-0.0207***	-0.0232**	-0.0282***	-0.0296***
	(-5.17)	(-6.22)	(-6.37)	(-2.54)	(-3.67)	(-4.07)
Ret_6M		-0.0002	-0.0002		-0.0017	-0.0015
		(-0.44)	(-0.31)		(-0.91)	(-0.82)
Size		-0.0001	0.0000		-0.0004	-0.0002
		(-0.24)	(0.04)		(-0.50)	(-0.29)
B/M		0.0008*	0.0007		0.0023	0.0019
		(1.70)	(1.57)		(1.22)	(1.07)
Amihud		-0.0028	-0.0026		-0.0117	-0.0129
		(-0.43)	(-0.42)		(-0.54)	(-0.60)
$Adj.R^2$	0.0151***	0.0432***	0.0490***	0.0149***	0.0467***	0.0528***

Table 5: Aggregate PC Factor vs. Option Implied Predictors

This table presents the Fama-Macbeth results from future stock returns on existing option implied predictors and the aggregate principal component (PC) factor. Panel A reports next-week returns (Ret_{1W}) and Panel B next-month returns (Ret_{1M}). To construct the aggregate PC factor, every Monday we extract the implied volatility surfaces for all firms from Optionmetrics from 1996 to 2014. Each implied volatility surface contains 112 volatilities defined across calls and puts, absolute deltas of 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, and 0.8 at maturities 30, 60, 91, 122, 152, 182, 273, and 365 calendar days. After demeaning each volatility surface, we compute the correlation matrix of the cross-section of implied volatility surfaces and decompose it using principal component analysis. The aggregate PC factor that is equal to $PC1 + PC2 + PC3$. The option implied predictors are the implied volatility spread as in Cremers and Weinbaum (2010), implied volatility skew as in Xing, Zhang, and Zhao (2010), risk-neutral skewness and risk-neutral kurtosis as defined in Conrad, Dittmar, and Ghysels (2013), O/S option to stock volume ration as defined in Johnson and So (2012), and the volatility of implied volatility as defined in Baltussen, Van Bekkum, and Van Der Grient (2012). The control variables are contemporaneous weekly return, lagged six-month return, the market-capitalization, book-to-market, and the Amihud measure of illiquidity.

Panel A: Weekly Return Regressions

	1		2		3		4		5		6	
X:	IVspread		IVskew		RNSkew		RNkurt		O/S		VoV	
Intercept	0.001	0.001	0.0015	0.0020	0.000	0.001	0.001	0.002	0.0011	0.0014	0.0029	0.0023
	(0.33)	(0.56)	(0.66)	(0.85)	(0.11)	(0.42)	(0.26)	(0.49)	(0.41)	(0.55)	(1.19)	(0.91)
X	0.009	-0.003	-0.0116	-0.0031	0.003	0.001	-0.000	-0.000	-0.0001	-0.0001	-0.0053	-0.0053
	(4.69)	(-0.83)	(-3.85)	(-0.93)	(3.27)	(1.53)	(-0.60)	(-0.57)	(-2.41)	(-1.87)	(-2.70)	(-2.81)
Agg. PC		0.003		0.0026		0.003		0.003		0.0024		0.0031
		(4.47)		(2.89)		(3.60)		(4.29)		(6.52)		(4.23)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adj.R^2	0.049	0.055	0.0584	0.0654	0.057	0.063	0.054	0.06	0.0424	0.0449	0.0482	0.0496
# Obs.	1,290,234	1,290,234	716,525	716,525	822,451	822,451	822,451	822,451	1,883,069	1,883,069	1,899,437	1,899,437

Panel B: Monthly Return Regressions

	1		2		3		4		5		6	
X:	IVspread		IVskew		RNSkew		RNkurt		O/S		VoV	
Intercept	0.006	0.008	0.0069	0.0085	0.004	0.007	0.007	0.009	0.0062	0.0076	0.0125	0.0101
	(0.68)	(0.93)	(0.89)	(1.11)	(0.45)	(0.69)	(0.56)	(0.79)	(0.75)	(0.92)	(1.45)	(1.12)
X	0.031	-0.010	-0.0374	-0.0071	0.009	0.005	-0.001	-0.001	-0.0003	-0.0002	-0.0208	-0.0200
	(6.18)	(-1.14)	(-4.43)	(-0.78)	(3.35)	(1.77)	(-1.17)	(-1.14)	(-2.52)	(-1.93)	(-3.72)	(-3.69)
Agg. PC		0.011		0.0091		0.010		0.011		0.0081		0.0119
		(4.99)		(3.49)		(3.71)		(4.49)		(7.21)		(4.76)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
$Adj.R^2$	0.049	0.055	0.0549	0.0549	0.06	0.066	0.058	0.063	0.0488	0.0486	0.0512	0.053
# Obs.	1,280,309	1,280,309	711,015	711,015	819,569	819,569	819,569	819,569	1,865,482	1,865,482	1,881,809	1,881,809

Table 6: Portfolio Sorts by Aggregate PC Factor

This table forms decile portfolios based on the aggregate principal component (PC) factor and reports weekly and monthly equal-weighted and value-weighted returns. The portfolio alpha, α_{FFC}, is the intercept of the regression of each portfolio on the Fama-French-Carhart factors. To construct the aggregate PC factor, every Monday we extract the implied volatility surfaces for all firms from Optionmetrics from 1996 to 2014. Each implied volatility surface contains 112 volatilities defined across calls and puts, absolute deltas of 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, and 0.8 at maturities 30, 60, 91, 122, 152, 182, 273, and 365 calendar days. After demeaning each volatility surface, we compute the correlation matrix of the cross-section of implied volatility surfaces and decompose it using principal component analysis. The aggregate PC factor that is equal to $PC1 + PC2 + PC3$. We report the t-statistics in parenthesis and the significance levels are indicated with * (10%), ** (5%), *** (1%).

	Weekly Returns				Monthly Returns			
	EW		VW		EW		VW	
Decile	Return	α_{FFC}	Return	α_{FFC}	Return	α_{FFC}	Return	α_{FFC}
1	0.0002	-0.0011	0.001	-0.001	0.001	-0.0034**	0.003	-0.001
	(0.1)	(-0.98)	(0.34)	(-0.35)	(0.38)	(-2.25)	(0.79)	(-0.64)
2	0.002	0.0007	0.002	0.001	0.007**	0.0022*	0.0053**	0.001
	(1.36)	(0.65)	(1.22)	(0.69)	(2.44)	(1.7)	(2.01)	(0.86)
3	0.0019	0.0005	0.001	0.0	0.0075***	0.0024**	0.0058**	0.001
	(1.41)	(0.57)	(0.88)	(0.0)	(2.93)	(2.13)	(2.5)	(0.86)
4	0.0021*	0.0008	0.001	0.0	0.0082***	0.003***	0.005**	0.0
	(1.75)	(0.91)	(1.25)	(0.3)	(3.51)	(2.8)	(2.43)	(0.28)
5	0.0019	0.005	0.002	0.0	0.008***	0.0027***	0.0062***	0.002
	(1.64)	(0.69)	(1.49)	(0.5)	(3.73)	(2.73)	(3.38)	(1.51)
6	0.0021**	0.0008	0.001	0.0	0.0083***	0.003***	0.0046***	0.0
	(1.97)	(1.07)	(1.38)	(0.27)	(4.07)	(3.12)	(2.73)	(0.08)
7	0.0022**	0.0009	0.001	0.0	0.0081***	0.0027***	0.0053***	0.001
	(2.08)	(1.18)	(1.48)	(0.34)	(4.13)	(2.95)	(3.35)	(0.87)
8	0.0018*	0.0005	0.001	0.0	0.0077***	0.0024***	0.0052***	0.001
	(1.78)	(0.76)	(1.63)	(0.55)	(3.91)	(2.6)	(3.43)	(0.72)
9	0.0021**	0.0009	0.0019**	0.001	0.0091***	0.0039***	0.0069***	0.0025***
	(2.01)	(1.2)	(2.23)	(1.36)	(4.34)	(4.01)	(4.35)	(2.66)
10	0.0035**	0.0022**	0.0029***	0.0018**	0.0129***	0.008***	0.0107***	0.0063***
	(2.56)	(2.35)	(2.62)	(2.14)	(4.71)	(6.25)	(5.15)	(5.3)
10-1	0.0038***	0.0038***	0.0028**	0.0027***	0.0136***	0.0132***	0.01***	0.0094***
	(6.01)	(6.6)	(2.44)	(2.58)	(11.3)	(12.6)	(4.76)	(5.03)

Table 7: Portfolios Transition Probabilities

This table reports the transition probabilities from one period to the next of decile portfolios formed based on the aggregate principal component (PC) factor. Every Monday, we extract the implied volatility surfaces for all firms from Optionmetrics from 1996 to 2014. Each implied volatility surface contains 112 volatilities defined across calls and puts, absolute deltas of 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, and 0.8 at maturities 30, 60, 91, 122, 152, 182, 273, and 365 calendar days. After demeaning each volatility surface, we compute the correlation matrix of the cross-section of implied volatility surfaces and decompose it using principal component analysis. The agregate PC factor that is equal to $PC1 + PC2 + PC3$. We report the probabilities that a stock moves one portofolio up, one portfolio down, that it remains in the same portfolio or that it moves to any other portfolios. The probabilities are reported for all portfolios and for portfolio 10 at the weekly and monthly frequency.

	Weekly		Monthly	
	All portfolios	Portfolio 10	All portfolios	Portfolio 10
Move up one portfolio	17%	0%	14%	0%
Remain in same portfolio	36%	41%	24%	59%
Move down one portfolio	17%	15%	14%	17%
Move to other portfolios	30%	44%	48%	24%

Table 8: Lending Fee and the Aggregate PC Factor

This table forms decile portfolios based on the aggregate principal component (PC) factor and reports weekly and monthly equal-weighted and value-weighted returns. Every Monday, we extract the implied volatility surfaces for all firms from Optionmetrics from 1996 to 2014. Each implied volatility surface contains 112 volatilities defined across calls and puts, absolute deltas of 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, and 0.8 at maturities 30, 60, 91, 122, 152, 182, 273, and 365 calendar days. After demeaning each volatility surface, we compute the correlation matrix of the cross-section of implied volatility surfaces and decompose it using principal component analysis. The agregate PC factor that is equal to $PC1+PC2+PC3$. Each portfolio is regressed on contemporaneous risk factors: the Fama-French and the momentum factors. The intercept of the regression is the portfolio alpha, α. We report the t-statistics in parenthesis and the significance levels are indicated with * (10%), ** (5%), *** (1%).

	Weekly Returns			Monthly Returns		
	1	2	3	1	2	3
Intercept	0.0017	0.0040	0.0041	0.0071*	0.0183*	0.0176*
	(1.48)	(1.46)	(1.44)	(1.76)	(1.91)	(1.75)
Agg. PC Factor	0.0035***		0.0003	0.0115***		0.0029
	(5.98)		(0.52)	(5.84)		(1.23)
Stock Lending Fee		-0.0184***	-0.0181***		-0.0841***	-0.0783***
		(-3.84)	(-3.79)		(-4.82)	(-4.63)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Adj. R^2	0.0151	0.0467	0.048	0.0151	0.0427	0.044
# Obs.	1,914,768	878,087	878,087	1,914,768	867,802	867,802

Table 9: Out-of-Sample Aggregate PC Factor

This table presents the Fama-Macbeth results from regressing next-week and next-month stock returns on the aggregate principal component (PC) factor. To construct the aggregate PC factor, every Monday we use a 3-year rolling window of implied volatility surfaces for all firms. Each implied volatility surface contains 112 volatilities defined across calls and puts, absolute deltas of 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, and 0.8 at maturities 30, 60, 91, 122, 152, 182, 273, and 365 calendar days. After demeaning each volatility surface, we compute the correlation matrix of the cross-section of implied volatility surfaces and decompose it using principal component analysis. The agregate PC factor that is equal to $PC1 + PC2 + PC3$. The remaining variables are the slope of the implied volatility term structure defined as the difference between long and short term ATM implied volatilities, implied volatility spread as in Cremers and Weinbaum (2010), implied volatility skew as in Xing, Zhang, and Zhao (2010), O/S option to stock volume ration as defined in Johnson and So (2012), contemporaneous weekly return (Ret_1w), lagged six-month return (Ret_6M), the market-capitalization (Size), book-to-market (BE/ME), and the Amihud measure of illiquidity. We use Optionmetrics data from 1996 to 2014.

	Weekly Return			Monthly Return		
	1	2	3	4	5	6
Intercept	0.0018	0.0015	0.0013	0.0073*	0.0080	0.0072
	(1.53)	(0.56)	(0.47)	(1.73)	(0.88)	(0.77)
Agg. PC Factor	0.0027***	0.0029***	0.0029***	0.0092***	0.0095***	0.0104***
	(3.96)	(5.19)	(3.84)	(3.84)	(5.00)	(3.93)
SlopeIV_TS			-0.0020			-0.0082
			(-0.72)			(-0.89)
IVskew			-0.0015			-0.0018
			(-0.70)			(-0.24)
IVspread			0.0068**			0.0007
			(2.18)			(0.06)
O/S			-0.0001			-0.0002
			(-1.58)			(-1.53)
Ret_1w		-0.0211***	-0.0206***		-0.0292***	-0.0295***
		(-6.34)	(-6.35)		(-3.85)	(-4.06)
Ret_6M		-0.0002	-0.0002		-0.0017	-0.0014
		(-0.44)	(-0.32)		(-0.90)	(-0.80)
Size		-0.0001	0.0000		-0.0004	-0.0003
		(-0.25)	(0.02)		(-0.55)	(-0.32)
B/M		0.0009*	0.0007		0.0024	0.0020
		(1.77)	(1.59)		(1.30)	(1.12)
Amihud		-0.0026	-0.0023		-0.0115	-0.0120
		(-0.41)	(-0.37)		(-0.53)	(-0.55)
$Adj.R^2$	0.0066***	0.0439***	0.0492***	0.0067***	0.0473***	0.0531***

Table 10: Aggregate PC Factor and Option Returns

This table reports the results on the relation between the aggregate PC factor and weekly straddle returns computed from Optionmetrics for the period 1996 to 2012. Panel A forms decile portfolios based on the aggregate principal component (PC) factor and reports next-week equal-weighted and value-weighted straddle returns. Panel B presents the Fama-Macbeth results from regressing next-week straddle returns on the aggregate principal component (PC) factor, the slope of the volatility term structure as in Vasquez (2017), historical minus implied volatility as in Goyal and Saretto (2009), and idiosyncratic volatility as in Cao and Han (2013). Every Monday, we extract the implied volatility surfaces for all firms from Optionmetrics from 1996 to 2014. Each implied volatility surface contains 112 volatilities defined across calls and puts, absolute deltas of 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, and 0.8 at maturities 30, 60, 91, 122, 152, 182, 273, and 365 calendar days. After demeaning each volatility surface, we compute the correlation matrix of the cross-section of implied volatility surfaces and decompose it using principal component analysis. The agregate PC factor that is equal to $PC1 + PC2 + PC3$. We report the t-statistics in parenthesis.

Panel A: Weekly Straddle Returns

	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P10-P1
Agg. PC Factor	-0.772	-0.22	-0.1	-0.029	0.024	0.069	0.111	0.156	0.219	0.541	
Equal Weighted Straddle Returns											
Mean	0.017	0.006	-0.002	-0.004	-0.009	-0.01	-0.016	-0.018	-0.02	-0.026	-0.042
t-stat	(6.20)	(2.43)	(-0.92)	(-1.66)	(-3.72)	(-4.32)	(-6.76)	(-7.22)	(-7.83)	(-7.89)	(-14.83)
StDev	0.076	0.073	0.072	0.072	0.068	0.069	0.065	0.07	0.074	0.093	0.081
Skewness	0.873	0.68	0.747	0.99	0.634	0.969	0.393	0.352	1.251	1.341	0.386
Kurtosis	4.10	5.18	7.49	6.99	7.02	8.45	5.10	5.30	11.21	10.48	2.26
Min	-0.244	-0.26	-0.362	-0.281	-0.344	-0.337	-0.338	-0.339	-0.377	-0.386	-0.305
Max	0.489	0.523	0.541	0.498	0.479	0.497	0.378	0.436	0.651	0.777	0.335
Value Weighted Straddle Returns											
Mean	0.018	0.007	-0.002	-0.003	-0.009	-0.01	-0.016	-0.018	-0.021	-0.026	-0.045
t-stat	(6.73)	(2.73)	(-0.60)	(-1.32)	(-3.65)	(-4.20)	(-6.80)	(-7.36)	(-8.25)	(-8.14)	(-15.64)
StDev	0.078	0.074	0.072	0.073	0.068	0.069	0.065	0.07	0.073	0.092	0.081
Skewness	0.944	0.797	0.809	1.067	0.636	1.071	0.409	0.455	1.266	1.496	0.268
Kurtosis	4.26	5.74	7.50	6.78	6.27	8.28	4.80	5.51	11.08	11.74	1.83
Min	-0.242	-0.258	-0.358	-0.273	-0.338	-0.325	-0.333	-0.336	-0.366	-0.395	-0.346
Max	0.515	0.56	0.545	0.489	0.453	0.485	0.363	0.453	0.639	0.792	0.322

Panel B: Weekly Straddle Return Regressions

	1	2
Intercept	-0.0082	0.0052
	(-3.47)	(1.67)
Agg. PC Factor	-0.0324	-0.0654
	(-15.37)	(-21.41)
Slope of VTS		0.2552
		(18.71)
HV–IV		0.0564
		(10.61)
Idio. Volatility		-0.4315
		(-9.85)
Adj. R^2	0.0053	0.0229
	(11.10)	(20.93)