EXPECTED STOCK RETURNS

Ana González-Urteaga

Department of Business Management, Universidad Pública de Navarra, Campus Arrosadía, 31006 Pamplona, Spain. Phone: (34) 948169400 (ana.gonzalezu@unavarra.es)

Belén Nieto

Department of Financial Economics and Accounting, Universidad de Alicante, San Vicente del Raspeig, 03690, Alicante, Spain. Phone: (34) 965903400 (belen.nieto@ua.es)

Gonzalo Rubio

Department of Economics and Business, Universidad CEU Cardenal Herrera, Reyes Católicos 19, 03204 Elche, Alicante, Spain. Phone: (34) 965426486 (gonzalo.rubio@uch.ceu.es)

This version: April 29, 2017

Abstract

Contrary to the standard practice of using past average realized returns when testing asset pricing models, this paper analyzes the factor structure and the cross-sectional variability of expected returns. We show that the first two principal components explain 99.6% of the variability of (lower bound) expected returns. Quality, funding illiquidity, the default premium and the market-wide variance risk premium explain most of the time-varying behavior of the first principal component. The cross-sectional fit of several asset pricing models using expected returns is consistently better than the one with average realized returns. The most successful model is a multi-factor model with the market, and the four aggregate factors that explain the first principal component. The cross-sectional *R*-squares proposed by Kan, Robotti, and Shanken (2013) for the principal component and the multifactor models are 52% and 84%, respectively. Both measures of fit are (asymptotically) different from zero.

Keywords: expected returns; risk-neutral variance; factor structure of expected returns; stochastic discount factor; cross-section of expected returns

JEL classification: G12, G13

The authors acknowledge financial support from the Ministry of Economics and Competitiveness through grant ECO2015-67035-P**.** In addition, Belén Nieto and Gonzalo Rubio acknowledge financial support from the Bank of Spain, and Ana González-Urteaga from Ministry of Economics and Competitiveness through grant ECO2016-77631-R (AEI/FEDER.UE). We thank John Cochrane, and Ian Martin for helpful comments.

Corresponding author: Gonzalo Rubio (gonzalo.rubio@uch.ceu.es)

1. Introduction

Alternative empirical specifications of macroeconomic-based asset pricing models have shown that expected returns are time-varying and show a clear counter-cyclical behavior.[1] As reported by Harvey, Liu, and Zhu (2016), the traditional cross-sectional analysis of factor-based models has delivered 316 systematic factor risks capable of explaining average returns. It is somehow frustrating that after years of academic research the important and obvious empirical question remains open: How many state variables do we need to price financial assets? It is important to emphasize that this huge literature employs average realized returns as proxies for expected returns.

The key insight is raised in Cochrane (2016), who points out that the truly relevant question is simply: What is the factor structure of expected returns? This is very different from the well-known research regarding the factor structure of realized returns at the end of the period. Along this line, our paper investigates the factor structure of returns that are expected at the beginning of the period. This is a surprising unexplored issue in the empirical literature of financial economics, and it is the main research objective of this paper. In addition, we also explore the cross-sectional variability of expected returns by taking advantage of the evidence obtained from the factor analysis.

The old issue of how to estimate expected returns remains an open question. Thus, it is not surprising that researchers have employed average realized returns over long horizons to proxy the true expected returns. However, the seminal paper of Breeden and Litzenberger (1978) shows that we may use option prices to extract forward-looking risk-neutral probabilities. As Kadan and Tang (2016) point out, it remains much more controversial whether we can extract physical probabilities from option prices and be able

[1] See Cochrane (2007, 2011, and 2016) for detailed reviews.

to derive the implied expected returns. Indeed, we have recently witnessed the debate raised by the Recovery theorem of Ross (2015), and the follow up papers by Borovicka, Hansen, and Scheinkman (2016), Bakshi, Chabi-Yo, and Gao (2016), and Jackwerth and Menner (2016).

Using a less ambitious (but probably more practical) approach, Martin (2013 and 2017) obtains a lower bound for the expected market risk premium by extracting forward-looking information on the market expected excess returns from option data and, in particular, from risk-neutral variances. The cost of this approach is that Martin (2017) does not obtain full recovery but, at least, he is able to obtain a useful lower bound on the expected return. Our paper employs this lower bound to study the factor structure of expected returns and the determinants of their temporal and cross-sectional variability.

Figure 1 display the normalized (lower bound) expected market risk premium, elicited from option prices, and the expected market risk premium forecasted from the previous 12-month aggregate dividend yield as in Cochrane (2011). We employ normalized data to emphasize the striking correlation between both series. They clearly follow the same business cycle pattern with peaks in recessions and low levels in booms. On average, the (lower bound) expected market risk premium is higher and more volatile than the expected market risk premium forecasted from the dividend yield. In addition, Martin and Wagner (2016) propose a simple model of expected returns of individual assets. They empirically show that stocks' average excess returns line up with their excess stock risk-neutral variances relative to the average stock risk-neutral variance, measured as a weighted average of individual stock risk-neutral variances. It turns out that their model's predictions also hold for portfolios sorted by beta, book-to-market, and momentum. Interestingly, when Martin and Wagner (2006) sort by size, the sensitivity of portfolio returns to risk-neutral variance is even higher than that predicted by theory.

Taking together, all of this evidence suggests that we may safely employ the lower bounds extracted from Martin's (2017) approach to infer useful information about time-varying expected returns.

We employ the traditional principal component analysis and the methodology of Connor and Korajczyk (1988) to extract the factors that better explain the variability of the variance-covariance matrix of (lower bound) expected returns. Independently of the technique employed, we show that first two principal components are enough to capture approximately 99.6% of the variability of expected returns. Indeed, the first component by itself explains around 97% of their variability. In contrast, when using realized returns, the first two principal components only explain around 77% of their variability. A particularly relevant question is to uncover the underlying determinants of the two first principal components of the factor structure of expected returns. We employ a full battery of state variables, which have been shown to be relevant pricing factors in well-known and fully recognized previous research. Indeed, the first principal component is strongly explained by the default premium, the quality minus junk factor (QMJ) of Asness, Frazzini, and Pedersen (2014), the market variance risk premium, and the betting against beta factor (BAB) of Frazzini and Pedersen (2014). These variables explain 53.4% of the temporal variability of the first principal component. These are key variables revealing the time-varying behavior of the factor structure of expected returns. On the other hand, the same battery of state variables does not succeed in explaining the second factor as they do with the first principal component. The momentum factor of Carhart (1997) is the only significant factor that explains the temporal behavior of the second component with a positive sign. Moreover, the highest adjusted R-square across alternative combinations of factors is low and equal to 6.1%.

In addition, we analyze the cross-sectional variability of (lower bound) expected returns. We first employ the two first principal components to find that their betas explain 90.9% of the cross-sectional variability. We also employ the out-of-sample Information Stochastic Discount Factor recently proposed by Ghosh, Julliard, and Taylor (2016 a, b) estimated from alternative sets of portfolios. Their betas explain approximately between 66% and 81% of the cross-sectional variability of our set of expected returns. Finally, we use a multi-factor asset pricing model, which includes the excess market return and the four factors that explain the time-varying behavior of the first principal component of expected returns. The betas of these five factors are significantly priced with the correct theoretical sign, and they jointly explain 98.2% of the cross-sectional variability of (lower bound) expected excess returns. We also employ the standard errors suggested by Kan, Robotti, and Shanken (2013), which are adjusted by errors-in-variable and model misspecification, and the corresponding corrected *R*-square. The results remain valid, with *R*-squares of 52.3%, 30.1%, and 83.6% for the principal components, the Information Stochastic Discount Factor, and the multi-factor model, respectively. These modified *R*-squares are in all three cases (asymptotically) statistically different from zero.

In order to obtain these results, we extract (lower bound) expected excess returns estimating risk-neutral variances from option prices using the approach of Martin (2013, 2017). He suggests an estimation of risk-neutral variances from a portfolio of out-of-the money puts and calls equally weighted by strike. As Martin (2013, 2017) points out, this is different from the approach of Britten-Jones and Neuberger (2000), and Jiang and Tiang (2005). These authors weight each out-of-the-money option by the inverse of the square of their strikes. In fact, this approach estimates risk-neutral entropy rather than risk-neutral variance. It is important to note that we must derive (lower bound) expected excess returns from risk-neutral variances, and not from the risk-neutral entropy. In any case, in the Appendix at the end of this paper, we perform a robustness analysis to compare both procedures. Overall, we find very similar results independently of the estimation methodology employed in the analysis.

Finally, we also evaluate exact expected returns rather than lower bounds. Under the exact relation, expected returns incorporate the covariance between returns and the product of returns and the stochastic discount factor. In order to avoid a model dependent stochastic discount factor, we employ the non-parametric estimation procedure of the Information Stochastic Discount Factor suggested by Ghosh, Julliard, and Taylor (2016 a, b). We find parallel but somewhat weaker results.

This paper proceeds as follows. Section 2 describes the theoretical framework to extract information about expected returns from option data, and Section 3 discusses the data employed in the research. Section 4 presents the estimation and descriptive statistics of (lower bound) expected returns. The estimation of the out-of-sample Information Stochastic Discount is discussed in Section 5. Section 6 presents the analysis of the factor structure of (lower bound) expected returns, and Section 7 contains the cross-sectional variability of (lower bound) expected returns. Section 8 displays a robustness analysis by estimating expected returns under the equality condition. Finally, Section 9 presents our conclusions. An Appendix at the end of the paper contains the results under the alternative estimation procedure of risk-neutral entropy of Jiang and Tian (2005).

2. The Lower Bound on the Expected Risk Premia

Martin (2017) obtains a lower bound on the expected market risk premium from the fundamental pricing equation for the market portfolio return:

$$1 = E_t^P\left(M_{t+1}R_{mt+1}\right) = \frac{E_t^Q\left(R_{mt+1}\right)}{R_{ft}} \quad , \tag{1}$$

where R_{mt+1} is the gross market return at time $t+1$, M_{t+1} is the stochastic discount factor (SDF) at time $t + 1$, R_{ft} is the gross risk-free rate available at time t, $E_t^P(\cdot)$ is the expectation operator under the physical probability as of time t, and $E_t^Q(\cdot)$ is the risk-neutral expectation operator as of time t.

The risk-neutral variance is

$$Var_t^Q\left(R_{mt+1}\right) = E_t^Q\left(R_{mt+1}^2\right) - \left(E_t^Q\left(R_{mt+1}\right)\right)^2. \tag{2}$$

Combining expressions (1) and (2), after a simple manipulation, the expected market risk premium is

$$E_t^P\left(R_{mt+1}\right) - R_{ft} = \frac{1}{R_{ft}} Var_t^Q\left(R_{mt+1}\right) - Cov_t^P\left(M_{t+1}R_{mt+1}, R_{mt+1}\right). \tag{3}$$

Martin (2017) points out that, as long as the relative risk aversion coefficient is greater or equal than one, under mild conditions and for most theoretical models, the following negative correlation condition (NCC) holds for the market portfolio return

$$Cov_t^P\left(M_{t+1}R_{mt+1}, R_{mt+1}\right) \leq 0. \tag{4}$$

Therefore, the risk-neutral variance normalized by the risk free rate constitutes a lower bound for the expected market risk premium:

$$E_t^P\left(R_{mt+1}\right) - R_{ft} \geq \frac{1}{R_{ft}} Var_t^Q\left(R_{mt+1}\right). \tag{5}$$

Kadan and Tang (2016) point out that the argument used by Martin (2017) does not extend to the case of individual stocks. For any given asset *j*, equation (4) can be written as[2]

$$Cov_t^P = \left(U'\left(\sum_{i=1}^{N} \omega_i R_{it+1} \right) R_{jt+1}, R_{jt+1} \right) \leq 0 \ , \tag{6}$$

where $U'(\cdot)$ is marginal utility, and ω_i is the weight given to asset *i* in a portfolio composed of *N* assets. Therefore, the sign of the covariance in (6) depends on the entire correlation structure between R_{jt+1} and all other stocks. Kadan and Tang (2016) obtain a sufficient condition under which the NCC holds for individual stocks. Let γ be the relative risk aversion. Then, the sufficient condition for equation (6) to hold is

$$\gamma \geq \frac{Var_t(R_{jt+1})}{Cov_t(R_{j+1}, R_{mt+1})}. \tag{7}$$

As in Kadan and Tang (2016), we denote this ratio by

$$\delta_{jt} \equiv \frac{Var_t(R_{jt+1})}{Cov_t(R_{j+1}, R_{mt+1})}, \tag{8}$$

and the sufficient condition is

$$\gamma \geq \delta_{jt} . \tag{9}$$

Thus, the NCC holds for any individual stock *j* with a positive beta if risk aversion is high enough to be larger than the stock's delta. The obvious problem with this condition is that risk aversion is not observable. However, we have ample evidence about the typical values of risk aversion, which for most individuals is between 1 and 10. Depending upon

[2] Kadan and Tang (2016) extend this simple analysis to the cases of a dynamic model with separable utility, and a dynamic consumption model with recursive utility.

the values of δ_j we obtain in our sample, we will decide how binding the NCC is in the sample. In any case, Martin and Wagner (2016) argue that it is not obvious that the NCC condition should hold for individual stocks. Indeed, their approach assumes that the covariance term is approximately zero for individual stocks.

3. Data

If the NCC is satisfied for individual stocks, the lower bound on expected returns for any given asset *j* is

$$E_t^P\left(R_{jt+1}\right)-R_{ft}\geq\frac{1}{R_{ft}}Var_t^Q\left(R_{jt+1}\right). \tag{10}$$

Therefore, we have to extract the risk-neutral variance for any given asset *j*. As we explain in the next Section, we calculate risk-neutral variances by integrating option prices for alternative strike prices. We employ daily data from OptionMetrics for the S&P 100 Index options and for individual options on all stocks included in the S&P 100 Index at some point during the sample period from January 1996 to August 2015. This yields 201 stocks used in our estimations. From the OptionMetrics database, we obtain all put and call options on the individual stocks and on the index with time to maturity τ between six days and 60 days. Given that the options are American style, it is convenient to work with short-term maturity options, for which the early exercise premium tends to be negligible.[3] We select the best bid and ask closing quotes to calculate the mid-quotes as the average of bid and ask prices, not actual transaction prices, to avoid the well-known bid-ask bounce problem described by Bakshi, Cao, and Chen (1997). In selecting our final option sample, we apply the usual filters. We discard options with zero open interest, zero bid prices, missing option delta or implied volatility, and negative implied volatility.

[3] See the evidence reported by Driessen, Maenhout, and Vilkov (2009), who employ a similar database.

Regarding the exercise level, we follow Jiang and Tian (2005), Driessen, Maenhout, and Vilkov (2009), and Martin (2017) and exclude in-the-money options. In addition, we ignore options with extreme moneyness, that is, puts with a delta higher than -0.05 and calls with a delta lower than 0.05.

Fama and French (2015) (FF hereafter) show that a five-factor model, that expands their popular three-factor model with profitability (robust minus weak, RMW) and investment (aggressive minus conservative, CMA) factors, explains anomalies associated with low betas, low share repurchases, and low volatility assets relative to high betas, high repurchases, and high volatility securities.[4] On the other hand, they are not able to explain the cross-section variability of momentum portfolios unless the momentum factor (MOM) of Carhart (1997) is included in the cross-section. Thus, from Kenneth French's website (http://mba.tuck.darmouth.edu), we collect monthly data on the FF five factors, the value-weighted stock market portfolio return, the risk-free rate, the MOM factor, the 25 FF portfolios by size and book-to-market, the 32 FF portfolios sorted by size, book-to-market and profitability, and the 10 portfolios sorted by investment aggressiveness. We also collect daily and monthly data on the 25 FF portfolios by size and investment aggressiveness, the 25 FF portfolios by book-to-market and profitability and the 10 portfolios sorted by momentum.

We also use the QMJ factor of Asnes, Frazzini and Pedersen (2014). These authors define a quality stock as an asset for which an investor would be willing to pay a higher price. These are stocks that are safe (low required rate of return), profitable (high return on equity), growing (high cash flow growth), and well managed (high dividend payout ratio). Asnes, Frazzini and Pedersen (2014) show that the QMJ factor, that buys high-

[4] Novy-Marx (2013) also discusses the relevance of the profitability factor in pricing the cross-section of average stock returns.

quality stocks and shorts low-quality (junk) stocks, earns significant risk-adjusted returns not only in the U.S. market, but also in 24 other countries. The QMJ factor is downloaded from AQR Capital Management Database (www.aqr.com).

Recent empirical evidence supports the presence of funding liquidity across a wide range of securities. Frazzini and Pedersen (2014) show that leverage constraints are strong and significantly reflected in the return differential between leveraged low-beta stocks and de-leveraged high-beta stocks. These authors argue that the positive and highly significant risk-adjusted returns relative to traditional asset pricing models shown by portfolios sorted by the level of market beta are explained by shadow cost-of-borrowing constraints. The authors illustrate their argument by proposing a market neutral BAB factor consisting of the difference between long levered low-beta stocks and short de-levered high-beta securities. The authors provide convincing evidence that the BAB factor generates high and consistent performance in each of the major global markets and asset classes, and that the results are independent of the asset pricing model employed in the analysis of performance. The BAB factor is downloaded from AQR Capital Management Database. We also employed the market-wide illiquidity factor of Pastor and Stambaugh (2003), which is obtained from Lubos Pastor's website (http://faculty.chicagobooth.edu/lubos.pastor/research/).

We define the default premium (DEF) as the difference between Moody's yield on Baa corporate bonds and the 10-year government bond yield. Both yields are downloaded from the Federal Reserve Statistical Release.

Finally, we estimate the variance risk premium (VRP) for the S&P 100 Index as the logarithm of the ratio between the realized variance and the risk-neutral variance on the index. The estimation details of risk-neutral variances for both individual stocks and the market are in the next section.

4. Estimation and Descriptive Statistics of (Lower Bound) Expected Risk Premia

We follow Martin (2013, 2017) to estimate risk-neutral variances and (lower bound) expected risk premia. Martin (2013) argues that, under stress market conditions, like October 1987 and the fall of 2008, there is no known way to replicate the payoff of a variance swap. This may be particularly severe for individual stocks, which may experience more frequent and larger jumps than market indices. Martin (2013) proposes the "simple variance swap", which can be hedged at discrete points even if the underlying's asset price jumps. This author develops a risk-neutral variance as an equally-weighted portfolio of options rather than a portfolio of options weighted by the inverse of the square of their strike price, and proposes SVIX, as an alternative to VIX. Martin (2017) shows that the risk-neutral variance is given by

$$Var_{jt,t+\tau}^{Q} = \frac{2R_{ft,t+\tau}}{S_{jt}^{2}} \left[\int_{0}^{F_{jt,t+\tau}} P_{jt,t+\tau}(K)\,dK + \int_{F_{jt,t+\tau}}^{\infty} C_{jt,t+\tau}(K)\,dK \right], \quad (11)$$

where $P_{jt,t+\tau}(K)$ and $C_{jt,t+\tau}(K)$ are the prices at time t of τ-maturity put and call options with strike K on either an asset or an index j with price S_{jt}, and $F_{jt,t+\tau}$ is the price of a future contract on the asset with the same maturity such that

$$F_{jt,t+\tau} = R_{ft,t+\tau}\left(S_{jt} - d_{jt}\right), \quad (12)$$

and d_{jt} represents the present value of dividends paid during the life of the contract.

We approximate expression (11) following the same steps carried out by Jiang and Tian (2005) to solve for their model free implied variance. Thus, we approximate the integrals of expression (11) by the following sums over a finite number of strikes:

$$I_x = \sum_{h=1}^{m} \left[g_{jt,t+\tau}\left(K_h^x\right) + g_{jt,t+\tau}\left(K_{h-1}^x\right) \right] \Delta K^x \ , \ x = C, P, \qquad (13)$$

where m equals 100, and ΔK and $g_{jt,t+\tau}$ are given by

$$\Delta K^x = \frac{\left(K_{\max}^x - K_{\min}^x\right)}{m} \ , \ K_h^x = K_{\min}^x + h\Delta K^x; \ h = 1,\ldots,m \qquad (14)$$

$$g_{jt,t+\tau}\left(K_h^x\right) = \begin{cases} C_{jt,t+\tau}\left(K_h^x\right), x = C \\ P_{jt,t+\tau}\left(K_h^x\right), x = P \end{cases} \qquad (15)$$

For each time-to-maturity τ from six to 60 days, we calculate the risk-neutral variance each day for each underlying asset that has at least three available options outstanding, using all the available options at time t.[5] For the risk-free rate, we use the T-bill rate of appropriate maturity (interpolated when necessary) from OptionMetrics, namely, the zero-coupon curve. For the dividend rate for the index, we employ the daily data on the index dividend yield from OptionMetrics. To infer the continuously compounded dividend rate for each individual asset, we combine the forward price with the spot rate used for the forward price calculations. We obtain the mean continuously compounded dividend rate by averaging the implied OptionMetrics dividends.

In practice, we only observe options at some finite sample set of strikes. We transform the prices of listed options into implied volatilities using the Black and Scholes (1973) model, and we fit a smooth function to the implied volatilities using cubic splines. Then, we extract implied volatilities at strikes K_h^x from the fitted function. Finally, we

[5] The window from six days to 60 days corresponds to the maximum range of time to maturity we allow in the necessary interpolation to have enough options every day in the sample with 30 days to maturity.

employ equation (13) to calculate the risk-neutral variance using the extracted out-of-the-money option prices. At each time *t*, we focus on a 30-day horizon maturity, interpolated when necessary following the procedure of Carr and Wu (2009).

We also calculate the market variance risk premium for each day in the sample. We first estimate the realized market variance over the same period as that for which risk-neutral variance is obtained for that day:

$$RV_{mt,t+\tau} = \frac{1}{\tau}\sum_{s=1}^{\tau} R_{mt+s}^{2} \ , \tag{16}$$

where $RV_{mt,t+\tau}$ is realized market variance. As in Carr and Wu (2009), we define the market variance risk premium as,

$$VRP_{mt,t+\tau} = \ln\left(\frac{RV_{mt,t+\tau}}{Var_{mt,t+\tau}^{Q}}\right). \tag{17}$$

Once we have a time-series of daily risk-neutral variances for each asset and the market from equation (11), we calculate the average risk-neutral variance across all days in each month for every available asset and compute the lower bound of expected excess returns following equation (10). Next we construct 20 risk-neutral variance-sorted portfolios including approximately the same number of assets in each portfolio. Portfolio 1 (P1) contains the assets with the lowest risk-neutral variance (and thus with the lowest expected return), while Portfolio 20 (P20) includes the stocks with the highest risk-neutral variance. The lower bound of the expected excess return and the realized excess return for each portfolio are computed imposing equal weights for the individual assets within the portfolio.

Panel A of Table 1 contains the descriptive statistics for the (lower bound) expected excess return or expected risk premia (ERP) for each of the 20 portfolios and the market, and Panel B reports the realized returns of the same portfolios. The first column shows that the average ERP goes from 0.56% for P1 to 5.87% for P20. The average (lower bound) expected market risk premium is 0.53% (6.36% on annual basis). Note that the average (lower bound) ERP across portfolios is higher than that for the market, which follows from Jensen´s inequality. The volatility of ERP maintains the same monotonic cross-sectional increase that we observe for the mean. Thus, P1 presents the lowest volatility of expected returns and P20 the highest volatility. Figure 2.1 shows these descriptive statistics. Moreover, in the third column of Panel A, we observe that the sensitivity of the (lower bound) ERP of each portfolio with respect to the market ERP increases monotonically with the level of the risk-neutral variance. In contrast, in the first column of Panel B, we show that there is very little dispersion of average realized returns across all 20 portfolios, although the return volatility increases with the average lower bound ERP. This pattern is displayed in Figure 2.2. Finally, in the third column of Panel B and Figure 2.3, we show that the realized return market betas of the 20 portfolios increase almost monotonically with the lower bounds. In other words, our sorting procedure is practically equivalent to sort assets by market portfolio betas. Portfolio P1, with the lowest risk-neutral variance, has the lowest average market beta of 0.47, and portfolio P20 has the highest average beta of 2.07. This is an important result. To all practical effects suggests that we are using the standard beta-sorted portfolios, which is one of the most popular sorting procedures in the traditional asset pricing literature.

Figure 3 represents the time-varying (lower bound) ERP for representative portfolios. All portfolios tend to have a strong counter-cyclical behavior with high peaks during bad economic times. As expected, this is especially the case for portfolio P20, which is the one with the highest market beta and the highest average lower bound of expected returns. Of course, this highly sensitive portfolio has a very poor performance during those months in which its expected return is especially high.

We estimate the deltas of Kadan and Tang (2016) for each of the 20 portfolios according to expression (8). We have to check whether they satisfy the sufficient condition under which the NCC of expression (9) holds for every portfolio. Figure 4 shows the average deltas of the 20 portfolios estimated with the full sample period. The highest delta corresponds to portfolio P20 and is equal to 4.14, and the lowest average delta is 1.19 of portfolio P1. We can safely employ the NCC to extract (lower bound) ERP of all 20 portfolios.

5. The Out-of-Sample Information Stochastic Discount Factor

For a given set of assets, Ghosh, Julliard, and Taylor (2016 a, b) propose a non-parametric estimation of an out-of-sample SDF. This is known as the Information Stochastic Discount Factor (ISDF). The basic idea is to minimize the relative entropy of the risk-neutral measure with respect to the physical measure. It turns out that this can be done by the following maximization problem for $\bar{M} = 1$

$$\underset{\{M_t\}_{t=1}^{T}}{arg\,min} \frac{1}{T}\sum_{t=1}^{T} M_t \ln M_t \quad s.t. \frac{1}{T}\sum_{t=1}^{T} M_t R_t^e = 0 \ , \tag{18}$$

where M_t is the SDF that prices a given set of asset returns at time *t*, and R_t^e is an *N*-vector of excess returns over the risk-free rate. Ghosh, Julliard, and Taylor (2016 a, b) argue that the solution can be obtained by the corresponding duality

$$\hat{M}_t \equiv M_t\left(\hat{\theta}_T, R_t^e\right) = \frac{e^{\hat{\theta}'_T R_t^e}}{\sum_{t=1}^{T} e^{\hat{\theta}'_T R_t^e}} \quad , \tag{19}$$

where $\hat{\theta}$ is the vector of Lagrange multipliers that solve the unconstrained convex problem

$$\hat{\theta}_T := \underset{\theta}{arg\,min} \sum_{t=1}^{T} e^{\hat{\theta}'_T R_t^e} \quad , \tag{20}$$

which is the dual formulation of the entropy minimization problem. Note that the normalization $\bar{M} = 1$ produces the demeaned SDF. In order to obtain the monthly ISDF for an economically reasonable magnitude, we actually employ the following expression:

$$\hat{M}_t \equiv M_t\left(\hat{\theta}_T, R_t^e\right) = \frac{T}{R_{ft}} \frac{e^{\hat{\theta}'_T R_t^e}}{\sum_{t=1}^{T} e^{\hat{\theta}'_T R_t^e}} \quad , \forall t\,. \tag{21}$$

Note that $\hat{\theta}$ is a $N\,x\,1$ vector that will depend on the returns of the assets employed in the estimation. Therefore, the ISDF will be different for different sets of assets. In our formal tests in the following section, we employ several ISDFs to check the robustness of our results. We follow the out-of-sample rolling estimation procedure suggested by Ghosh, Julliard, and Taylor (2016 b). The first rolling sample contains monthly returns for the previous 30 years to the month of the first estimated ISDF. With this set of data, we obtain the first vector $\hat{\theta}_T$. These estimated parameters are used to compute the values of the ISDF for the next 12 months. Then, we roll the estimation window 12 months to generate a series of ISDF for the full sample period. Figure 5 show the ISDF estimated with the 25 FF portfolios from January 1996 to July 2015. It tends to be high and more volatile during NBER recession and financial disturbance periods than in other times. The mean and volatility of the ISDF are 0.999 and 0.512, respectively, and the covariance of the NCC given by (4) is equal to -0.31.

6. The Factor Structure of (Lower Bound) Expected Risk Premia

This section discusses the main research question in this paper: What is the factor structure of expected returns?

We first extract the principal components from the standard approach, which uses the *N x N* sample variance-covariance matrix of the expected returns of our sample of 20 risk-neutral variance-sorted portfolios. In addition, we employ the approach of Connor and Korajczyk (1988). They propose using the eigenvectors associated with the *K* largest eigenvalues of the *T x T* cross product matrix given by

$$\Omega = \frac{1}{N}\left[E_t\left(R_{t+1}^e\right)\right]'\left[E_t\left(R_{t+1}^e\right)\right], \qquad (22)$$

where $E_t\left(R_{t+1}^e\right)$ is the 20-dimensional vector of (lower bound) ERP at time *t*. It can be shown that as the cross-section becomes large, the *K x T* matrix with rows consisting of the *K* eigenvectors of Ω will converge to the matrix of factor realizations. Therefore, the estimated *K x T* dimensional matrix represents the excess returns that replicate the realizations of non-observable systematic factors.

Table 2 contains the percentage explained by the first five principal components estimated from both (lower bound) ERP and realized excess returns of our 20 risk-neutral variance-sorted portfolios. Panel A shows the results obtained by the traditional approach. It turns out that the first two principal components of lower bound expected returns explain 99.4% of the variability of expected returns. It seems that two factors may be sufficient to explain the cross-sectional variability of expected returns. On the contrary, the first two principal components of realized returns explain only 76.8% of the variability. The time-varying behavior of the first two principal components of (lower bounds) expected returns during our sample period is displayed in Figure 6.1. The first principal component, which explains 96.6% of the variability of expected returns, follows closely the counter-cyclical patter of expected returns shown in the previous section of the paper. The second principal component, which only explains an additional 2.8%, tends to be negative during bad economic times and its variability is much lower relative to the first principal component. Figure 6.2 shows the betas estimated from a multiple OLS regression of realized returns of each of the 20 portfolios on the two principal components of (lower bound) expected returns. Most factor loadings are negative and, in the case of the first principal component, they become (almost monotonically) more negative the higher the risk-neutral variance of the portfolio is. The behavior of betas relative to the second principal component is not as smooth as the in the first case, but it also tends to be more negative the higher the risk-neutral variance is. The surprising exception is the beta of portfolio P20, which becomes positive and relatively much higher than the rest of betas. This illustrates the importance of the second principal component to explain the high expected risk premium of the portfolio with the highest risk-neutral variance. These results are consistent with the strong counter-cyclicality of the first principal component of lower bounds. The principal component moves negatively with respect to realized returns and positively relative to expected returns.

Panel B of Table 2 displays the results using the principal components estimated with the Connor and Korajczyk (1988) procedure. The results are very similar to the ones obtained under the first approach. In this case, the first two principal components explain 99.8% of the variability of lower bounds. The first component explains a slightly higher variability reaching 98.7% instead of 96.6% of the first approach. In any case, it seems clear that two principal components are enough to capture most of the variability of (lower bound) expected risk premia. The percentage of realized returns explained by the first five principal components is practically equal to the percentage obtained by the traditional procedure.

Next we address the key issue of understanding the underlying state variables that explain the temporal behavior of these two first principal components. We employ the components estimated under the traditional approach. We select a full battery of candidates that have been shown to have explanatory power in the time-series and cross-sectional variability of returns in previous literature. Panel A of Table 3 shows the time-series determinants of the first principal component. Below the regression-estimated coefficients, we report in parentheses the *t*-statistic based on the traditional OLS standard errors, and in brackets the *t*-statistic based on HAC standard errors. We first analyze the explanatory capacity of the FF five factor risks employed in their five-factor model. It turns out that, the excess market return and the HML factor move negatively with the first principal component, but they lose statistical significance when we employ HAC standard errors. When we add the MOM factor, all FF factors become statistically significant except SMB. The adjusted *R*-square is 12%. Then, we analyze the individual explanatory power of the QMJ and BAB factors, the market-wide illiquidity factor, the DEF premium, and the market variance risk premium (VRP). All of them are statistically significant, and the DEF premium by itself has an *R*-square of 44.1%, which is much higher than the *R*-square of the augmented (with MOM) FF five-factor model. The next regression includes all variables together. Neither the MOM factor, nor the FF factors are statistically significant. The market-wide illiquidity is also estimated with noise. However, the estimates of the QMJ and the BAB factors, and the DEF premium are statistically different from zero with relatively high HAC *t*-statistics. The market VRP is statistical significant but it losses significance when we employ HAC standard errors. We obtain similar results when we run regressions with these four factor plus the market excess return, which presents a non-statistical significant negative coefficient. When we drop the market portfolio, all factors significantly explain the behavior of the first principal component, as we estimate the market VRP with more precision.

The positive relation between the first principal component of (lower bound) ERP and the QMJ, DEF premium and market VRP suggests that the variables tend to be high in bad economic times. Indeed, Asness, Frazini, and Pedersen (2014) show that the QMJ factor displays large realized returns during downturns, which indicates that the quality-based factor does not exhibit bad-times risk. In particular, they plot the risk-adjusted returns of the QMJ factor against market excess returns and show that the quality factor presents a mild positive convexity, which suggests that the QMJ factor benefits from flight-to-quality during financial and economic crises. The BAB factor, which may be a proxy for funding liquidity, presents a negative and significant relation with the first principal component. Note that the way in which Frazzini and Pedersen (2014) construct the BAB factor implies that low or negative returns of this factor are times of poor funding liquidity or high borrowing constraints. The adjusted *R*-square of the four-factor model is 53.4%, and the inclusion of the excess market return practically has no effect on the adjusted *R*-square.[6]

[6] Given the strong effect of the DEF premium, we analyze the explanatory power of the credit risk premium of Asvanunt and Richardson (2016), which can also be downloaded from the AQR data library. The credit risk premium is the long-term corporate bond total return minus empirical-duration-matched long-term government bond total return. Using the available until December 2014, this variable presents a negative and significant relation with the first principal component even if we control for the DEF premium. When we run the model with the four significant state variables, the credit risk premium remains statistically different from zero only with OLS standard errors, but not with HAC standard errors.

Panel B of Table 3 shows the results regarding the second principal component. This second factor is much harder to explain than the first component. The relative smooth behavior of the second principal component may be an explanation of this finding. If anything, the HML, CMA, MOM, the market-wide illiquidity factor, and the market VRP with a negative sign show a very weak statistical significance. Overall, the adjusted *R*-square is low and equal to 6.1%. To be precise, across alternative combinations of factors, the MOM factor is the only variable statistically different from zero, even with respect to HAC standard errors. Thus, there is a positive and significant relation between the second principal component of expected returns and the MOM factor.

7. The Cross-Sectional Variability of (Lower Bound) Expected Risk Premia

Having documented the time-series determinants of the first two principal components of lower bounds, we now turn to study the variability of expected returns in cross-section. Our approach is to perform a traditional two-pass cross-sectional regression of Fama and MacBeth (1973) with monthly data, and the estimated (lower bound) ERP of the 20 risk-neutral variance-sorted portfolios as the left-hand side variable. We define the explanatory aggregate factors in the following subsections.[7] We also employ the rigorous econometric methodology of Kan, Robotti, and Shanken (2013) (KRS hereafter), who derive the asymptotic distribution of the cross-sectional *R*-square as a measure of model ability to price the cross-section of average returns. Moreover, they provide standard errors of risk premium estimators adjusted for errors-in-variable and model misspecification.

[7] Kadan and Tang (2016) also perform a cross-sectional analysis with stock characteristics as explanatory variables.

7.1 Principal Components

Our first cross-sectional test performs the following cross-sectional regression:

$$E_t\left(R^e_{pt+1}\right) = \lambda_0 + \lambda_1\beta_{p,f1} + \lambda_2\beta_{p,f2} + e_{pt} \;\; ; \; p = 1,\ldots,20 \;\; , \tag{23}$$

where $E_t\left(R^e_{pt+1}\right)$ is the (lower bound) expected risk premium of portfolio *p*, and the two betas for each portfolio are estimated using rolling time-series regressions of the observed returns of each portfolio on the two principal components of lower bounds (f_1 and f_2), using the past 59 months and the current month:[8]

$$R^e_{pt+1} = \alpha_p + \beta_{p,f1} f_{1t+1} + \beta_{p,f2} f_{2t+1} + \varepsilon_{pt+1} \;\; . \tag{24}$$

Panel A of Table 4 contains the results using the (lower bound) expected risk premium of each portfolio as test assets. Below the risk premium estimators, we report the *p*-values associated with the traditional Fama and MacBeth (1973) standard error in parentheses and, in brackets, the *p*-values of the standard errors adjusted for errors-in-variable and the potential misspecification of the model due to KRS (2013). We provide two *R*-square statistics as measures of goodness of model fit. The first number is the standard cross-sectional *R*-square given by the following expression:

$$R^2 = 1 - \frac{Var_N\left(\bar{\hat{e}}_p\right)}{Var_N\left(\overline{E\left(R^e_p\right)}\right)} , \tag{25}$$

where $Var_N\left(\cdot\right)$ indicates cross-sectional variance across all *N* (20) portfolios, $\bar{\hat{e}}_p$ is the average estimated pricing error of portfolio *p*, and $\overline{E\left(R^e_p\right)}$ is the average of the (lower

[8] In all cross-sectional regressions reported in the next sections, we estimate betas using the same rolling window with 60 months data of realized returns on the risk factors.

bound) expected risk premia of portfolio *p*. The second cross-sectional statistic is the *R*-squared suggested by KRS (2013) and below, in parenthesis, we provide the *p*-value for the null that the estimated *R*-square equals zero.

According to the classic standard errors, the two risk premia associated with each principal component are strongly statistically significant. This is also true for the first principal component even if we adjust the standard error. However, in the case of the second principal component, the adjusted standard error presents an associated *p*-value, which is high and equal to 0.42. As expected, the sign of the first risk premium is negative given the counter-cyclical pattern of the first principal component. This first factor increases in times of high marginal utility, which suggest a negative risk premium. This is exactly what we find in the cross-sectional results. On the other hand, the risk premium associated with the second principal component is positive and much lower than the first one. These betas explain approximately 91% of the cross-sectional variability of expected returns. Figure 7.1 visualizes the strong cross-sectional fit reflected in the high *R*-square. The most problematic portfolio is P20 with a high pricing error of 1.01%. The high variability of this portfolio explains that the cross-sectional *R*-square due to KRS (2013) is lower and equal to 52%. This statistic weights the portfolio pricing errors by the inverse of the variance-covariance matrix of returns and thus it assigns a much larger weight to portfolio 20. In any case, the *p*-value is low and equal to 0.06.

Panel B of Table 4 reports the results of the cross-sectional regression with average realized excess returns as the dependent variables on the betas of the principal components extracted from the (lower bound) expected returns. None of the estimated risk premium coefficients is statistically different from zero and the classic cross-sectional *R*-square is 24%. Finally, Panel C shows the results of the cross-section of average realized excess returns on the betas of principal components obtained from the variance-covariance matrix of realized excess returns. Note that this is the usual empirical procedure of testing asset pricing models. As before, the risk premia are not statistically different from zero, and the classic *R*-square is still low and equal to 10.5%. The poor adjustment of the model can be visualized in Figure 7.2.

7.2 Out-of-Sample Information Discount Factors

The SDF cross-sectional specification is given by

$$E\left(R_{pt+1}^{e}\right)=-\frac{Cov_t\left(\hat{M}_{t+1},R_{pt+1}^{e}\right)}{E_t\left(\hat{M}_{t+1}\right)}=\left(-\frac{Var_t\left(\hat{M}_{t+1}\right)}{E_t\left(\hat{M}_{t+1}\right)}\right)\frac{Cov_t\left(\hat{M}_{t+1},R_{pt+1}^{e}\right)}{Var_t\left(\hat{M}_{t+1}\right)}=\lambda_{\hat{M}}\beta_{p\hat{M}} \quad (26)$$

where $\hat{M}$ is the estimated (out-of-sample) ISDF of Ghosh, Julliard, and Taylor (2016 a, b). Note that the estimated ISDF depends on the previously selected set of test assets. This opens the door to a potential lack of robustness in the empirical results. However, theoretically, the covariance between $\hat{M}$ (marginal utility) and the return of any risky asset must be negative thus obtaining a positive expected excess return. Therefore, the risk premium ($\lambda_{\hat{M}}$) must be negative to have economic sense.[9]

We estimate three alternative series of out-of-sample ISDF using the methodology explained in Section 5 and three sets of assets: the 25 FF portfolios sorted by size and book-to-market, the 35 FF portfolios composed of the 25 portfolios plus 10 momentum portfolios, and 52 FF portfolios containing the 32 portfolios by size, book-to-market and profitability, the 10 portfolios by momentum, and the 10 portfolios by investment aggressiveness.

[9] Ghosh, Julliard, and Taylor (2016 b) test their model using four alternative sets of ISDF estimates from 1963 to 2010. In all four cases, the estimated risk premium is negative and statistically significant with cross-sectional *R*-squares always higher than 60%.

As before, the test assets for the cross-sectional regression are the 20 portfolios sorted by risk-neutral variance. We use (lower bound) of ERP in Panel A, and average realized excess returns in Panel B, as dependent variables. Panel A.1 reports results regarding the first estimated ISDF from returns on the 25 FF portfolios. The cross-sectional *R*-square is high and equal to 74%. As expected, the risk premium is negative but the *p*-value associated with the KRS standard error is 0.12. In other words, the associate risk premium is estimated with too much noise to be confident about the statistical significance of the coefficient. Figure 8.1 shows the cross-sectional fit of the model. It seems that the non-significant negative risk premium is due to the extremely high pricing error of portfolio P20, which is equal to 2.0%. Otherwise, the fit seems to be very reasonable. The very high pricing error of portfolio P20 explains the low *R*-square of KRS (2013). González-Urteaga and Rubio (2016) show that the DEF premium is a key factor explaining the cross-sectional variability of the volatility risk premia. They also show that this result reflect a very different behavior of the underlying components of their sample portfolios with respect to credit risk that generates a significant dispersion of the volatility swap pricing of their portfolios. In our case, portfolio P20 has a high and negative return beta relative to the DEF premium. This suggests that the underlying components of this portfolio have a high credit risk relative to the rest of the portfolios used in our sample, and investors are willing to pay a high variance swap price to hedge default risk. To conclude, ISDF beta cannot explain the behavior of the extreme high risk-neutral variance portfolio.

Panel B.1 of Table 5 and Figure 8.2 show the results with an ISDF constructed with the same set of portfolios but average realized excess returns as the dependent variable of the model. The fit is extremely poor, and the model cannot explain at all the cross-sectional variability of average returns.

Panels A.2 and B.2 of Table 5 present a similar evidence when we add the momentum portfolios to the estimation of the ISDF. The *p*-value associated with the KRS *R*-square is 0.09 in the case of (lower bound) ERP, which suggests stronger results. However, the risk premium becomes positive although not statistically different from zero with respect to the KRS standard errors. Finally, Panels A.3 and B.3 employ the 52 portfolios in the estimation of the ISDF. When we employ expected returns as test assets, the risk premium becomes high, positive and statistically different from zero. Moreover, the classic and KRS cross-sectional *R*-squares are 81% and 30%, respectively. The surprising positive sign of the risk premium invalidates the economic interpretation of the estimated coefficient. Therefore, the cross-sectional results are very sensitive to the assets involved in the estimation of the ISDF. In fact, the correlation between the first two ISDF and the market risk premium are -0.25, and -0.05, respectively. This is the economic sensible sign of the correlation between any SDF and the market. However, the correlation between the ISDF estimated with 52 portfolios and the market is positive and equal to 0.15. As in all other cases, the risk premium using realized returns is not statistically different from zero.

7.3 A Five-Factor Multi-Beta Pricing Model

In Section 6 of this paper, we show that the return generating process underlying the first principal component of (lower bound) ERP may be written as

$$f_{1t} = \alpha + \beta_1 R_{mt}^e + \beta_2 QMJ_t + \beta_3 BAB_t + \beta_4 DEF_t + \beta_5 VRP_{mt} + \varepsilon_t \ , \qquad (27)$$

where R_m^e is the excess market portfolio returns, and the rest of the factors have been defined previously. This specification explains 53.5% of the variability of the first principal component (see Table 3). Thus, when analyzing the cross-sectional variability of expected returns, it seems natural to assume an ICAPM 5-factor model given by

$$E_t\left(R^e_{pt+1}\right) = \lambda_0 + \lambda_1\beta_{p,excm} + \lambda_2\beta_{p,qmj} + \lambda_3\beta_{p,bab} + \lambda_4\beta_{p,def} + \lambda_5\beta_{p,vrp} + e_{pt} \; ; \; p = 1,\ldots,20 \quad (28)$$

Panel A of Table 6 shows the results employing expected returns as test assets. The performance of the model is striking. The cross-sectional *R*-square is 98.2%. The corrected KRS *R*-square is slightly lower but equal to a high 83.6% and it is statistically different from zero. Figure 9.1 displays the clear strong fit between ERP across portfolios and the corresponding fitted values. Again the portfolio P20 presents the highest pricing error but it is equal to 0.47% which is a much lower error than in the previous cases of principal components and ISDFs.

The market risk premium is positive and strongly significant. As we report later, this contrasts with the negative market risk premium obtained when we use average realized returns. The risk premia associated with the QMJ, DEF, and market VRP are statistically different from zero with the right negative sign. This holds for both, the classic and KRS standard errors. Note that all these three factors have a negative correlation with the excess market return. This implies that they have positive values in high marginal utility times, which explains the negative and significant risk premium associated with these factors. The risk premium of the BAB factor is also negative, but it is statistically different from zero only when we employ the classic standard errors. The *p*-value becomes 0.34 for the KRS standard errors. It turns out that the betas of high (lower bound) expected returns, like portfolios P18 to P20, are highly negative with respect to the BAB factor, and positive with respect to portfolios P1 and P2. It turns out that this also holds for the QMJ factor and the default premium. The performance of portfolios P18 to P20 becomes worse when QMJ, BAB, and the default premium increase. This may explain the overall BAB negative risk premium, although it is being estimated with much more noise.

The pricing of the DEF premium and the market VRP deserves a more detailed comment. González-Urteaga and Rubio (2017) show that these two factors are priced economically and statistically different in the volatility and return segments of the market. On average, common factors in both segments explain 90% of the variability of volatility risk premium portfolios, but only 65% of the variability of equity return portfolios. Indeed, the market VRP is priced significantly in the volatility segment but not in the equity portfolios. In this paper, we find that both factors, DEF premium and market VRP, are priced in the cross-section of expected returns. Note that the test assets are expected returns extracted from the volatility segment of the market. In this sense, this new evidence is consistent with the findings of González-Urteaga and Rubio (2017).[10]

Panel B of Table 6 shows the empirical results regarding average realized returns. The nice cross-sectional fit of expected returns strongly contrasts with the much poorer fit of the model when we employ average realized returns. Figure 9.2 illustrates this weak cross-sectional fit. None of the risk premia is statistically different from zero, and the traditional cross-sectional R-square is 35.9%. The intercept of the model is highly positive and equal to 9.96% on annual basis. This is not an economically sensible magnitude for the zero-beta return, and it is much higher than the reasonable intercept in Panel A, which equals 3.0% on annual basis.

[10] See also the evidence of Barras and Malkhozov (2016) who rejects the null hypothesis that the conditional market VRP measured in the equity and option markets are equal.

The consistent significant behavior of the model both in the time-series and in the cross-section makes this five-factor model to be a strong candidate for explaining expected returns.

8. The Expected Risk Premia under the Equality Condition: A Robustness Analysis

Martin (2017) shows that the expected risk premium of any asset *j* is given by

$$E_t^P\left(R_{jt+1}\right)-R_{ft}=\frac{1}{R_{ft}}Var_t^Q\left(R_{jt+1}\right)-Cov_t^P\left(M_{t+1}R_{jt+1},R_{jt+1}\right). \qquad (29)$$

Given that the SDF is not observable, we cannot calculate the covariance in the right hand side of the expression unless we impose a particular asset pricing model.[11] This is why we work with the lower bounds of ERP. However, the procedure in Ghosh, Jullliard, and Taylor (2016 a, b) estimates a non-parametric (model free) ISDF. This ISDF is what we use in the estimation of exact expected returns following equation (29).[12]

In this case, we employ daily data of 60 portfolios, including the 25 FF portfolios by size and investment aggressiveness, the 25 FF portfolios by book-to-market and profitability and the 10 portfolios sorted by momentum.[13] We use a rolling window of previous 30 years of daily data to estimate each $\hat{\theta}_T$. These parameters remain constant to compute daily values for the ISDF in the next year. Then, the window rolls one year to generate the complete daily series of ISDF. The covariance in the right hand side of expression (29) is estimated monthly using daily data of the ISDF and realized returns within the month.

[11] Recall that Martin and Wagner (2016) ignore the covariance term when they work with individual stock returns. In fact, their results support the view that the covariance term does not play a fundamental role in the analysis.

[12] Of course, unavoidably the exact expected returns estimates will depend on the set of assets employed in the estimation of the ISDF.

[13] Note that we include portfolios based on all factors of the FF five-factor model plus momentum.

Figure 10.1 compares the average values of the lower bound and the exact expected risk premia across all 20 risk-neutral variance-sorted portfolios, while Figure 10.2 shows the volatility of these two sets. The pattern is identical across portfolios, although the exact average expected risk premia (and the volatility) tend to be lower than the corresponding lower bounds.

Table 7 reproduces the factor structure of expected returns but with the exact condition. The percentage explained by the first principal component is much lower than in the case of lower bounds. This is the case independently of the methodology employed to extract principal components. Thus, the first principal component explains 84.4% of the variability of exact expected excess returns relative to the 96.2% for the case of lower bounds. We need even more than the first five principal components of exact expected returns to explain the 99.4% captured by the first two principal components of lower bounds under the traditional estimation procedure. In any case, as before, these percentages are much higher than in the case of realized returns.

Figure 11.1 and 11.2 show the time-varying behavior of the first two principal components using lower bound and the exact estimation of expected returns. In both cases, the first principal component is strongly counter-cyclical with a very similar temporal pattern. The behavior of the second principal component also follows a similar temporal pattern under both procedures. The largest differences between the bounds and the exact expected returns for both principal components occur at the end of 2001, 2002 and (especially) 2008, years that correspond to periods of financial distress.

We now turn to the cross-sectional determinants of exact expected excess returns. Panel A of Table 8 shows the results regarding the principal components. The results are very similar to the ones reported in Table 4, although the risk premium associated with the first principal component is estimated with slightly less precision. Panel B contains the results for the risk premium associated with the ISDF obtained from the 25 FF portfolios as in Table 5. The estimator of the risk premium is also negative but much lower in absolute value than in Table 5. Moreover, we estimate the coefficient with much less precision than with lower bounds. Note that we obtain the exact expected excess returns with a daily ISDF estimated with a much ample set of portfolios than in Panel A.1 of Table 5. However, the betas in the cross-section are estimated with respect to the ISDF obtained from the 25 FF portfolios. Panel C of Table 8 shows that the four variables explaining a large percentage of the temporal behavior of the first principal component of lower bounds, QMJ, BAB, DEF premium, and market VRP, are also significant determinants of the first principal component of exact expected returns. As before, all four risk premia are negative. The intercept and the market risk premium are 3.7% and 8.2% on annual basis, respectively. Both coefficients are statistically different from zero and show reasonable magnitudes for both the zero-beta rate and the market risk premium. The variability explained by the five-factor model is practically the same as with lower bounds. Once again, this result suggests that the five-factor model is a powerful and reasonable candidate to explain the cross-sectional behavior of expected returns.

9. Conclusions

After several decades of intense research using realized past returns to study the behavior of stock returns, we still know very little about the factor structure and cross-sectional variability of expected returns. Merton (1980) already shows how difficult estimating the mean market return is, and he argues that only extending the sample over time we are able to approximate means adequately. Sampling at higher frequencies does not help with precise estimation of mean returns.

This paper covers partially this gap using a combination of option pricing results from Martin (2013, 2017), and insights of the recent SDF literature due to Ghosh, Julliard, and Taylor (2016 a, b). We estimate (lower bound) ERP, and exact ERP for 20 risk-neutral variance-sorted portfolios using option prices. We find that the factor structure of ERP can be summarized with the two first principal components. The first principal component explains 96.6% (84.4%) of the variability of (lower bound) ERP (exact ERP). This first principal component presents a reasonable and strong counter-cyclical behavior. The second principal explains an additional 2.8% (9.7%) of the variability of (lower bound) ERP (exact ERP). These percentages are clearly higher than the percentages found when employing realized returns.

When we use (lower bound) ERP, we conclude that both the time-series and cross-sectional variability of expected returns are mainly explained by similar state variables. These aggregate variables are the differences between high and low quality portfolios, the differences between leveraged and deleveraged betas (funding liquidity or the tightness of borrowing constraints), the default premium, and the market variance risk premium. All risk premia are negative and statistically significant. The market risk premium is positive and statistically different from zero. These results do not seem to depend on the inclusion of the covariance term in the expression for expected excess returns. Overall, our results suggest that expected returns are time-varying in a strong counter-cyclical way and vary much more than what is usually accepted. The robust identification of the set of factors that significantly explain a very large percentage of their variability is an important step to understand the behavior of expected returns.

References

Asness, C., A. Frazzini, and L. Pedersen (2014), Quality Minus Junk, *Working Paper, AQR Capital Management and New York University, Stern School of Business*.

Asvanunt, A., and S. Richardson (2016), The Credit Risk Premium, *Available at SSRN: http://dx.doi.org/10.2139/ssrn.2563482SSRN*.

Bakshi, G., Cao, C., Chen, Z., 1997. Empirical Performance of Alternative Option Pricing Models. *Journal of Finance* 52, 2003–2049.

Bakshi, G., F. Chabi-Yo, and X- Gao (2016), A Recovery That We Can Trust? Deducing and Testing the Restrictions of the Recovery Theorem, *Working Paper, University of Maryland*.

Barras, L., and A. Malkhozov (2016), Does Variance Risk Have Two Prices? Evidence from the Equity and Option Markets, *Journal of Financial Economics* 121, 79-92.

Black, F., and M. Scholes (1973), The Pricing of Options and Corporate Liabilities, *Journal of Political Economy* 81, 637-654.

Borovicka, J., L. Hansen, and J. Scheinkman (2016), Misspecified Recovery, *Journal of Finance* 71, 2493-2544.

Breeden, D., and R. Litzenberger (1978), Prices of State Contingent Claims Implicit in Option Prices, *Journal of Business* 51, 621-651.

Britten-Jones, M., and A. Neuberger (2002), Option Prices, Implied Price Processes, and Stochastic Volatility, *Journal of Finance* 55, 839-866.

Carhart, M. (1997), On Persistence in Mutual Fund Performance, *Journal of Finance* 52, 57-82.

Carr, P., and L. Wu (2009), Variance Risk Premia, *Review of Financial Studies* 22, 1311-1341.

Cochrane, J. (2007), Financial Markets and the Real Economy, In *Handbook of the Equity Premium*, edited by Rajnish Mehra, Elsevier, 237-325.

Cochrane, J. (2011), Discount Rates: American Finance Association Presidential Address, *Journal of Finance* 66, 1047-1108.

Cochrane, J. (2016), Macro-Finance, *Working Paper, Hoover Institution, Stanford University*.

Connor, G., and R. Korajczyk (1988), Risk and Return in Equilibrium APT: Application of a New Test Methodology, *Journal of Financial Economics* 21, 255-290.

Driessen, J., Maenhout, P., Vilkov, G., 2009. The price of correlation risk: evidence from equity options. *Journal of Finance* 64, 1377–1406.

Fama, E., and J. MacBeth (1973), Risk, Return and Equilibrium: Empirical Tests, *Journal of Political Economy* 81, 607-636.

Fama, E., and K. French (2015), A Five-Factor Asset Pricing Model, *Journal of Financial Economics* 116, 1-22.

Frazzini, A., and L. Pedersen (2014), Betting against Beta, *Journal of Financial Economics* 111, 1-25.

Ghosh, A., C. Julliard, and P. Taylor (2016 a), What is the Consumption-CAPM Missing? An Information-Theoretic Framework for the Analysis of Asset Pricing Models, forthcoming in the *Review of Financial Studies*.

Ghosh, A., C. Julliard, and P. Taylor (2016 b), An Information-Theoretic Asset Pricing Model, *Working Paper*, *London School of Economics*.

González-Urteaga, A., and G. Rubio (2016), The Cross-Sectional Variation of Volatility Risk Premia, *Journal of Financial Economics* 119, 353-370.

González-Urteaga, A., and G. Rubio (2017), The Joint Cross-Sectional Variation of Equity Returns and Volatilities, *Journal of Banking and Finance* 75, 27-34.

Harvey, C., Y. Liu, and H. Zhu (2016), …. And the Cross-Section of Expected Returns, *Review of Financial Studies* 29, 5-68.

Jackwerth, J., and M. Menner (2016), Does the Ross Recovery Theorem Works Empirically? *Working Paper, University of Konstanz*.

Jiang, G., and Y. Tian (2005), The Model Free Implied Volatility and its Information Content, *Review of Financial Studies* 18, 1305-1342.

Kadan, O., and X. Tang (2016), A Bound on Expected Stock Returns, *Working Paper, Washington University in St. Louis*.

Kan, R., C. Robotti, and J. Shanken (2013), Pricing Model Performance and the Two-Pass Cross-Sectional Regression Methodology, *Journal of Finance* 68, 2617-2649.

Martin, I. (2013), Simple Variance Swaps, *Working Paper, London School of Economics*.

Martin, I. (2017), What is the Expected Return on the Market? *Quarterly Journal of Economics* 132, 367-433.

Martin, I., and C. Wagner (2016), What is the Expected Return on a Stock? *Working Paper, London School of Economics*.

Merton, R. (1980), On Estimating the Expected Return on the Market: An Exploratory Investigation, *Journal of Financial Economics* 8, 323-361.

Novy-Marx, R. (2013), The Other Side of Value: The Gross Profitability Premium, *Journal of Financial Economics* 108, 1-28.

Pastor, L., and R. Stambaugh (2003), Liquidity Risk and Expected Stock Returns. *Journal of Political Economy* 111, 642-685.

Ross, S. (2015), The Recovery Theorem, *Journal of Finance* 70, 615-648.

Table 1. Descriptive Statistics of (Lower Bound) Expected Risk Premia and Realized Returns for 20 Portfolios Sorted by Risk-Neutral Variance: January 1996-July 2015.

	Panel A: (Lower Bound) Expected Risk Premium			Panel B: Realized Returns		
	Average	Volatility	Market ERP Beta	Average	Volatility	Market Beta
P1	0.0056	0.0037	0.6375	0.0143	0.0337	0.4722
P2	0.0069	0.0046	0.8047	0.0137	0.0376	0.5393
P3	0.0078	0.0053	0.9457	0.0131	0.0399	0.5844
P4	0.0086	0.0059	1.0674	0.0128	0.0404	0.6206
P5	0.0093	0.0063	1.1452	0.0107	0.0418	0.6681
P6	0.0100	0.0067	1.2268	0.0134	0.0459	0.8187
P7	0.0107	0.0071	1.3171	0.0126	0.0468	0.8274
P8	0.0113	0.0076	1.4053	0.0145	0.0450	0.7627
P9	0.0120	0.0081	1.5008	0.0118	0.0549	0.9955
P10	0.0128	0.0085	1.5802	0.0135	0.0550	0.9717
P11	0.0136	0.0091	1.6810	0.0137	0.0582	1.0429
P12	0.0145	0.0097	1.7956	0.0114	0.0580	0.9932
P13	0.0157	0.0106	1.9782	0.0131	0.0618	1.1528
P14	0.0170	0.0117	2.1817	0.0139	0.0630	1.0879
P15	0.0185	0.0128	2.3955	0.0125	0.0679	1.1576
P16	0.0205	0.0143	2.6670	0.0101	0.0675	1.1855
P17	0.0236	0.0171	3.1532	0.0069	0.0753	1.2998
P18	0.0281	0.0211	3.8136	0.0134	0.0872	1.5392
P19	0.0355	0.0268	4.7015	0.0143	0.0959	1.6151
P20	0.0587	0.0470	7.9929	0.0146	0.1315	2.0688
MARKET	0.0053	0.0049	1.0000	0.0060	0.0449	1.0000

This table presents the descriptive statistics of 20 portfolios sorted by risk-neutral variance and the market, represented by the Standard & Poor 100 Index. The first two columns show the mean and the volatility of the (lower bound) expected risk premia. The third column is the sensitivity of the expected risk premium of the 20 portfolios to the expected market risk premium. The last three columns contain the mean, the volatility and the market beta of realized returns.

Table 2. The Factor Structure of (Lower Bound) Expected Risk Premia for 20 Portfolios Sorted by Risk-Neutral Variance. Percentage Explained by the First Five Principal Components: January 1996-July 2015.

	Panel A: Principal Components		Panel B: Principal Components Connor & Korajczyk	
	(Lower Bound) Expected Risk Premia	Realized Excess Returns	(Lower Bound) Expected Risk Premia	Realized Excess Returns
Factor (PC) 1	96.62	66.74	98.66	67.11
Factor (PC) 2	99.38	76.75	99.77	77.22
Factor (PC) 3	99.71	80.50	99.90	80.90
Factor (PC) 4	99.87	83.47	99.95	83.85
Factor (PC) 5	99.93	85.73	99.97	86.05

The two columns of Panel A show the percentage of the variability of the 20 x 20 variance-covariance matrix of (lower bound) expected and realized excess returns of our 20 portfolios, respectively, explained by the first five principal components. The two columns of Panel B report the same percentages of the variability but of the *T x T* covariance matrix, in this case, computed following the procedure of Connor and Korajczyk (1988).

Table 3. Determinants of the Factor Structure of (Lower Bound) Expected Risk Premia. January 1996-July 2015.

Panel A: Determinants of the First Principal Component

Const.	Market	SMB	HML	RMW	CMA	MOM	QMJ	BAB	P&S	DEF	VRP	Adj R^2
0.027	-0.081	0.035	-0.130	0.096	0.124							
(20.25)	(-2.34)	(0.78)	(-2.25)	(1.57)	(1.57)							0.074
[10.14]	[-1.50]	[0.85]	[-1.24]	[1.58]	[1.25]							
0.027	-0.103	0.057	-0.194	0.128	0.167	-0.090						
(21.01)	(-3.00)	(1.29)	(-3.29)	(2.13)	(2.14)	(-3.62)						0.120
[10.30]	[-1.83]	[1.44]	[-1.96]	[1.96]	[1.88]	[-2.19]						
0.026							0.158					
(20.63)							(3.91)					0.058
[10.54]							[1.78]					
0.028								-0.102				
(21.44)								(-3.33)				0.041
[10.31]								[-2.21]				
0.027									0.070			
(21.38)									(3.71)			0.052
[10.67]									[3.33]			
-0.012										1.589		
(-3.92)										(13.62)		0.441
[-1.72]										[5.32]		
0.029											0.010	
(20.83)											(3.88)	0.057
[9.61]											[2.90]	
-0.005	-0.020	0.018	0.021	-0.002	0.057	-0.008	0.150	-0.129	0.014	1.375	0.004	
(-1.70)	(-0.66)	(0.54)	(0.44)	(-0.02)	(0.98)	(-0.39)	(2.07)	(-4.43)	(0.98)	(11.54)	(1.96)	0.533
[-0.79]	[-0.64]	[0.50]	[0.42]	[-0.02]	[0.83]	[-0.37]	[2.61]	[-3.56]	[0.79]	[4.90]	[1.26]	
-0.006	-0.034						0.131	-0.118		1.410	0.004	
(-1.99)	(-1.20)						(3.25)	(-4.85)		(12.81)	(2.16)	0.535
[-0.95]	[-1.13]						[2.51]	[-4.13]		[5.70]	[1.40]	
-0.006							0.158	-0.113		1.425	0.005	
(-2.19)							(4.73)	(-4.71)		(13.03)	(2.73)	0.534
[-1.02]							[2.99]	[-4.21]		[5.71]	[1.73]	

Panel B: Determinants of the Second Principal Component

Const.	Market	SMB	HML	RMW	CMA	MOM	QMJ	BAB	P&S	DEF	VRP	Adj R^2
0.003	-0.003	-0.017	-0.050	-0.001	0.043							
(8.52)	(-0.36)	(-1.54)	(-3.46)	(-0.04)	(2.15)							0.050
[5.49]	[-0.28]	[-1.78]	[-2.37]	[-0.05]	[1.82]							
0003	-0.001	-0.020	-0.043	-0.005	0.038	0.011						
(8.38)	(-0.06)	(-1.76)	(-2.81)	(-0.29)	(1.88)	(1.67)						0.058
[5.44]	[-0.05]	[-1.93]	[-1.99]	[-0.30]	[1.73]	[2.04]						
0.003							0.015					
(8.44)							(1.41)					0.004
[5.89]							[1.37]					
0.003								-0.008				
(8.74)								(-1.04)				0.000
[5.76]								[-0.80]				
0.003									0.009			
(8.78)									(1.79)			0.009
[6.17]									[1.48]			
0.003										-0.003		
(2.88)										(-0.08)		-0.004
[1.22]										[-0.03]		
0.003											-0.001	
(7.38)											(-0.85)	-0.001
[4.54]											[-0.60]	
0.003	-0.003	-0.015	-0.036	-0.003	0.036	0.012	0.011	-0.012	0.006	-0.019	-0.001	
(2.70)	(-0.24)	(-1.31)	(-2.08)	(-0.14)	(1.72)	(1.66)	(0.45)	(-1.13)	(1.26)	(-0.45)	(-1.71)	0.061
[1.26]	[-0.26]	[-1.26]	[-1.27]	[-0.11]	[1.28]	[2.04]	[0.46]	[-0.94]	[1.04]	[-0.18]	[-1.19]	
0.002						0.017			0.009		-0.001	
(6.83)						(2.77)			(1.80)		(-1.73)	0.039
[4.01]						[2.37]			[1.56]		[-1.12]	

This table shows the estimated coefficients of time-series regressions of each of the two first principal components on alternative state variables. The first six variables are the intercept and the Fama and French (2015) factors, MOM is the Momentum Factor of Carhart (1997), QMJ is the quality minus junk Factor of Asness, Frazzini, and Pedersen (2014), BAB is the betting-against-beta factor of Frazzini and Pedersen (2014), P&S is the Pastor and Stambaugh (2003) illiquidity factor, DEF is the default premium, and VRP is the market variance risk premium defined as the logarithm of the realized variance divided by the risk-neutral variance. OLS *t*-statistics are reported in parenthesis and *t*-statistics based on HAC standard errors in brackets.

Table 4. The Cross-Section of Portfolios sorted by Risk-Neutral Variance and the Factor Structure of (Lower Bound) Expected Risk Premia. January 1996-July 2015.

Panel A: (Lower Bound) Expected Risk Premia and Principal Components from the Variance-Covariance Matrix of (Lower Bound) Expected Risk Premia

λ_0	λ_{f_1}	λ_{f_2}	*R*-square
0.0020	-0.0162	0.0006	0.909
(0.046)	(0.000)	(0.002)	0.523
[0.000]	[0.003]	[0.417]	(0.06)

Panel B: Average Realized Excess Returns and Principal Components from the Variance-Covariance Matrix of (Lower Bound) Expected Risk Premia

λ_0	λ_{f_1}	λ_{f_2}	*R*-square
0.0069	-0.0043	0.0009	0.243
(0.012)	(0.322)	(0.235)	0.010
[0.001]	[0.249]	[0.365]	(0.96)

Panel C: Average Realized Excess Returns and Principal Components from the Variance-Covariance Matrix of Realized Excess Returns

λ_0	λ_{f_1}	λ_{f_2}	*R*-square
0.0097	-0.0020	-0.0007	0.105
(0.013)	(0.733)	(0.861)	0.011
[0.004]	[0.725]	[0.875]	(0.93)

This table reports the risk premia estimates from the two-pass cross-sectional regression of returns on rolling betas associated with the two first principal components. Panel A shows the results using (lower bound) expected risk premia, and Panels B and C employ excess realized returns. Principal components come from the variance-covariance matrix of (lower bound) expected risk premia in Panels A and B and from the variance-covariance matrix of realized excess returns in Panel C. We report the *p*-value in parentheses and the Kan, Robotti, and Shanken (2013) adjusted *p*-value in brackets. The cross-sectional *R*-square reported in the first line is 1 minus the cross-sectional variance of average pricing errors divided by the cross-sectional variance of the average dependent variable. Kan, Robotti, and Shanken (2013) propose the cross-sectional *R*-square reported in the second line, and the corresponding (asymptotically valid) *p*-value is given in parentheses.

Table 5. The Cross-Section of Portfolios sorted by Risk-Neutral Variance and the Information Stochastic Discount Factor. January 1996-July 2015.

Panel A: (Lower Bound) Expected Risk Premia				Panel B: Average Realized Excess Returns			
Panel A1: ISDF Estimated from 25 Size and Book-to-Market Portfolios				Panel B1: ISDF Estimated from 25 Size and Book-to-Market Portfolios			
Model 1	λ_0	λ_{ISDF}	*R*-square	Model 1	λ_0	λ_{ISDF}	*R*-square
	0.0089	-0.0510	0.741		0.0068	-0.0497	-0.087
	(0.000)	(0.187)	0.020		(0.032)	(0.518)	0.017
	[0.000]	[0.116]	(0.57)		[0.003]	[0.560]	(0.69)
Panel A2: ISDF Estimated from 25 Size and Book-to-Market + 10 Momentum Portfolios				Panel B2: ISDF Estimated from 25 Size and Book-to-Market + 10 Momentum Portfolios			
Model 2	λ_0	λ_{ISDF}	*R*-square	Model 2	λ_0	λ_{ISDF}	*R*-square
	0.0089	0.1847	0.655		0.0090	0.0341	-0.115
	(0.000)	(0.010)	0.072		(0.005)	(0.832)	0.016
	[0.000]	[0.229]	(0.09)		[0.000]	[0.830]	(0.66)
Panel A3: ISDF Estimated from 32 Size, Book-to-Market and Profitability + 10 Momentum + 10 Investment Portfolios				Panel B3: ISDF Estimated from 32 Size, Book-to-Market and Profitability + 10 Momentum + 10 Investment Portfolios			
Model 3	λ_0	λ_{ISDF}	*R*-square	Model 3	λ_0	λ_{ISDF}	*R*-square
	0.0092	0.5368	0.808		0.0075	0.0117	-0.090
	(0.000)	(0.000)	0.301		(0.013)	(0.926)	0.061
	[0.000]	[0.003]	(0.02)		[0.000]	[0.935]	(0.43)

This table reports the risk premia estimates from the two-pass cross-sectional regression of returns on rolling betas associated with the Information Stochastic Discount Factor (ISDF) of Ghosh, Julliard, and Taylor (2016 a, b). Panel A shows the results using (lower bound) expected risk premia, and Panel B employs excess realized returns. The set of portfolios used for the estimation of the ISDF is indicated at the top of each subpanel from A1 to B3. We report the *p*-value in parentheses and the Kan, Robotti, and Shanken (2013) adjusted *p*-value in brackets. The cross-sectional *R*-square reported in the first line is 1 minus the cross-sectional variance of average pricing errors divided by the cross-sectional variance of the average dependent variable. Kan, Robotti, and Shanken (2013) propose the cross-sectional *R*-square reported in the second line, and the corresponding (asymptotically valid) *p*-value is given in parentheses.

Table 6. The Cross-Section of Portfolios sorted by Risk-Neutral Variance and a Multi-Factor Asset Pricing Model. January 1996-July 2015.

Panel A: (Lower Bound) Expected Risk Premia

λ_0	λ_m	λ_{qmj}	λ_{bab}	λ_{def}	λ_{vrp}	R-square
0.0025	0.0129	-0.0137	-0.0035	-0.0022	-0.1234	0.982
(0.002)	(0.000)	(0.000)	(0.011)	(0.000)	(0.000)	0.836
[0.229]	[0.000]	[0.000]	[0.338]	[0.000]	[0.000]	(0.00)

Panel B: Average Realized Excess Returns

0	λ_m	λ_{qmj}	λ_{bab}	λ_{def}	λ_{vrp}	R-square
0.0083	-0.0003	0.0011	-9.0e-5	-0.0007	-0.0750	0.359
(0.070)	(0.955)	(0.706)	(0.986)	(0.583)	(0.425)	0.063
[0.085]	[0.957]	[0.762]	[0.991]	[0.740]	[0.630]	(0.99)

This table reports the risk premia estimates from the two-pass cross-sectional regression of returns on rolling betas associated with the five aggregate risk factors: *m* represents the market excess return, *qmj* is the quality minus junk Factor of Asness, Frazzini, and Pedersen (2014), *bab* denotes the betting-against-beta factor of Frazzini and Pedersen (2014), *def* is the default premium, and *vrp* is the market variance risk premium defined as the logarithm of the realized variance divided by the risk-neutral variance. Panel A shows the results using (lower bound) expected risk premia, and Panel B employs excess realized returns. We report the *p*-value in parentheses and the Kan, Robotti, and Shanken (2013) adjusted *p*-value in brackets. The cross-sectional *R*-square reported in the first line is 1 minus the cross-sectional variance of average pricing errors divided by the cross-sectional variance of the average dependent variable. Kan, Robotti, and Shanken (2013) propose the cross-sectional *R*-square reported in the second line, and the corresponding (asymptotically valid) *p*-value is given in parentheses.

Table 7. The Factor Structure of Expected Returns under the Exact Condition for 20 Portfolios Sorted by Risk-Neutral Variance. Percentage Explained by the First Five Principal Components: January 1996-July 2015.

	Panel A: Principal Components		Panel B: Principal Components Connor & Korajczyk	
	Exact Expected Risk Premia	Realized Excess Returns	Exact Expected Risk Premia	Realized Excess Returns
Factor (PC) 1	84.36	66.74	94.39	67.11
Factor (PC) 2	94.10	76.75	98.03	77.22
Factor (PC) 3	96.20	80.50	98.73	80.90
Factor (PC) 4	97.19	83.47	99.06	83.85
Factor (PC) 5	98.05	85.73	99.35	86.05

The two columns of Panel A show the percentage of the variability of the 20 x 20 variance-covariance matrix of exact expected and realized excess returns of our 20 portfolios, respectively, explained by the first five principal components. The two columns of Panel B report the same percentages of the variability but of the *T x T* covariance matrix, in this case, computed following the procedure of Connor and Korajczyk (1988).

Table 8. The Cross-Section of Exact Expected Risk Premia. January 1996-July 2015.

Panel A: Principal Components	λ_0	λ_{f_1}	λ_{f_2}				*R*-square
	0.0048	-0.0007	0.0006				0.847
	(0.000)	(0.012)	(0.000)				0.374
	[0.000]	[0.149]	[0.095]				(0.00)
Panel B: ISDF	λ_0	λ_{ISDF}					*R*-square
	0.0063	-0.0021					0.798
	(0.000)	(0.917)					0.020
	[0.000]	[0.801]					(0.22)
Panel C: Multi-factor	λ_0	λ_m	λ_{qmj}	λ_{bab}	λ_{def}	λ_{vrp}	*R*-square
	0.0031	0.0068	-0.0075	-0.0032	-0.0008	-0.0335	0.983
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.027)	0.836
	[0.000]	[0.000]	[0.000]	[0.073]	[0.002]	[0.130]	(0.00)

This table reports the risk premia estimates from the two-pass cross-sectional regression of exact expected excess returns on rolling betas associated with the two principal components (Panel A), the ISDF estimated with the 25 FF portfolios (Panel B), and with five aggregate risk factors (Panel C). *m* represents the market excess return, *qmj* is the quality minus junk Factor of Asness, Frazzini, and Pedersen (2014), *bab* denotes the betting-against-beta factor of Frazzini and Pedersen (2014), *def* is the default premium, and *vrp* is the market variance risk premium defined as the logarithm of the realized variance divided by the risk-neutral variance. We report the *p*-value in parentheses and the Kan, Robotti, and Shanken (2013) adjusted *p*-value in brackets. The cross-sectional *R*-square reported in the first line is 1 minus the cross-sectional variance of average pricing errors divided by the cross-sectional variance of the average dependent variable. Kan, Robotti, and Shanken (2013) propose the cross-sectional *R*-square reported in the second line, and the corresponding (asymptotically valid) *p*-value is given in parentheses.

Figure 1. Option- and Dividend Yield-based Standardized Expected Market Risk Premium. January 1996-July 2015.

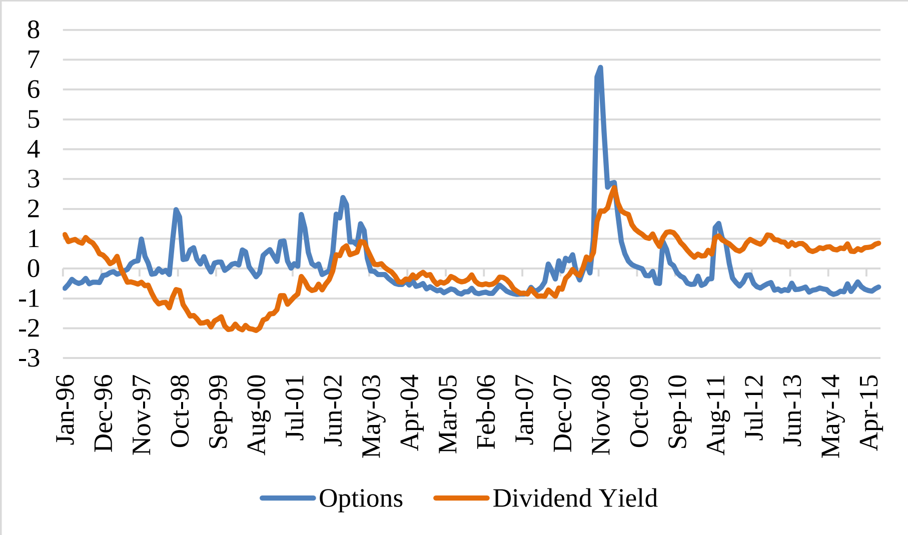


Figure 2.1. Mean and Volatility of (Lower Bound) Expected Risk Premia of 20 Portfolios Sorted by Risk-Neutral Variance. Monthly Data, January 1996-July 2015.

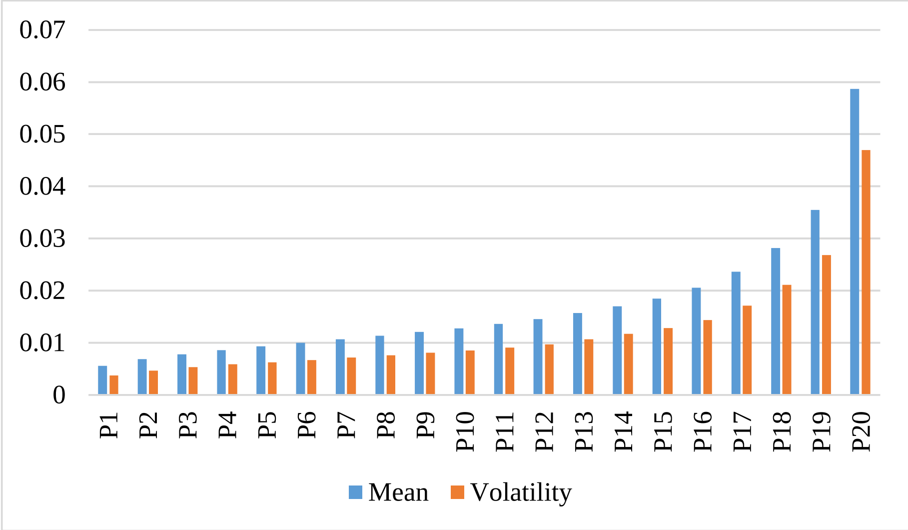


Figure 2.2. Mean and Volatility of Realized Returns of 20 Portfolios Sorted by Risk-Neutral Variance. Monthly Data, January 1996-July 2015.

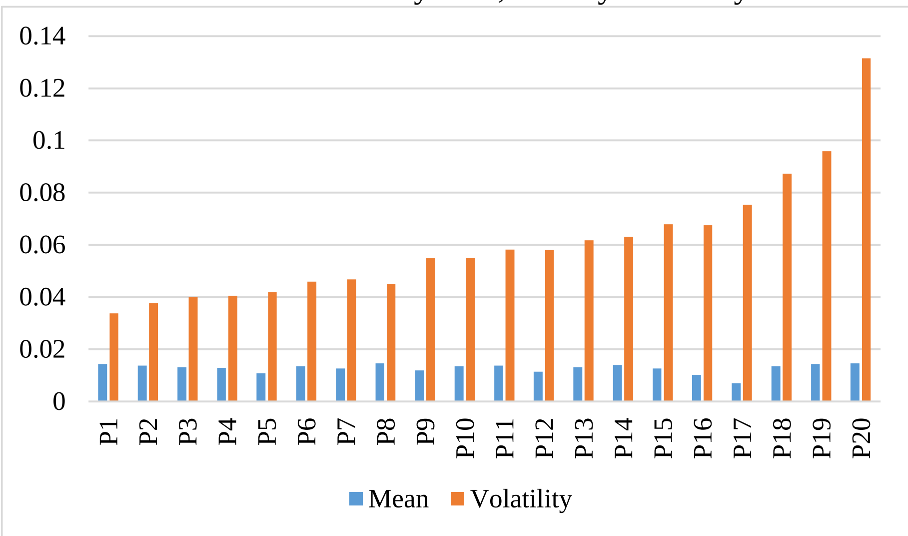


Figure 2.3. Mean and Beta of (Lower Bound) Expected Risk Premia of 20 Portfolios Sorted by Risk-Neutral Variance. Monthly Data, January 1996-July 2015.

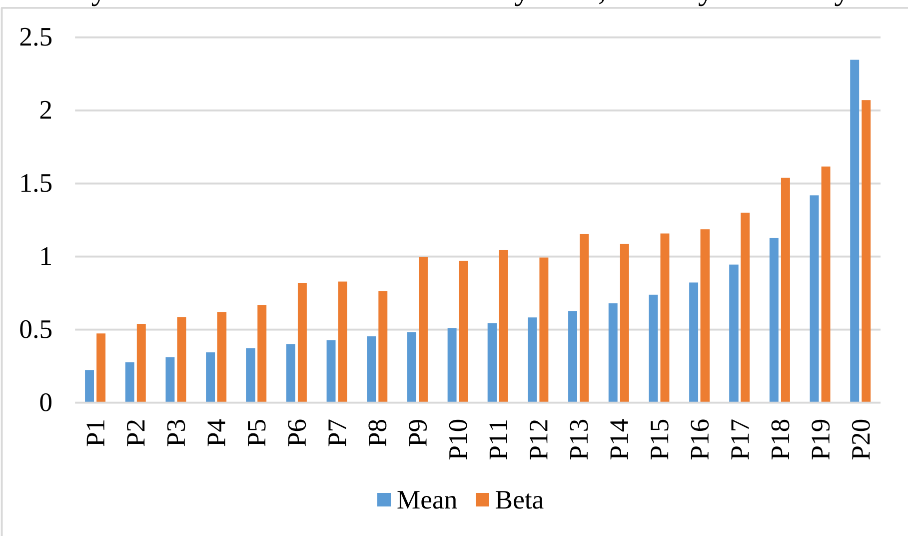


Figure 3. (Lower Bound) Expected Risk Premia for some Representative Portfolios. Monthly Data, January 1996-July 2015.

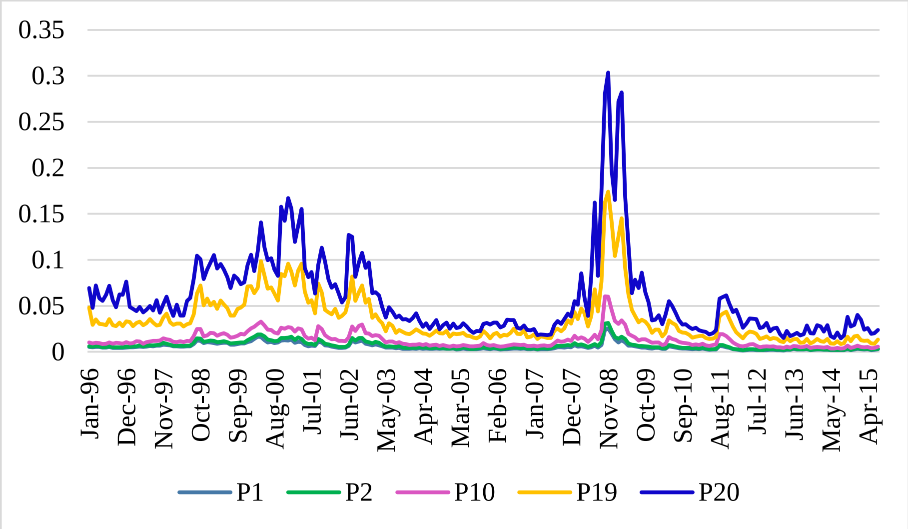


Figure 4. Deltas for 20 Portfolios Sorted by Risk-Neutral Variance. Full sample period January 1996-July 2015.

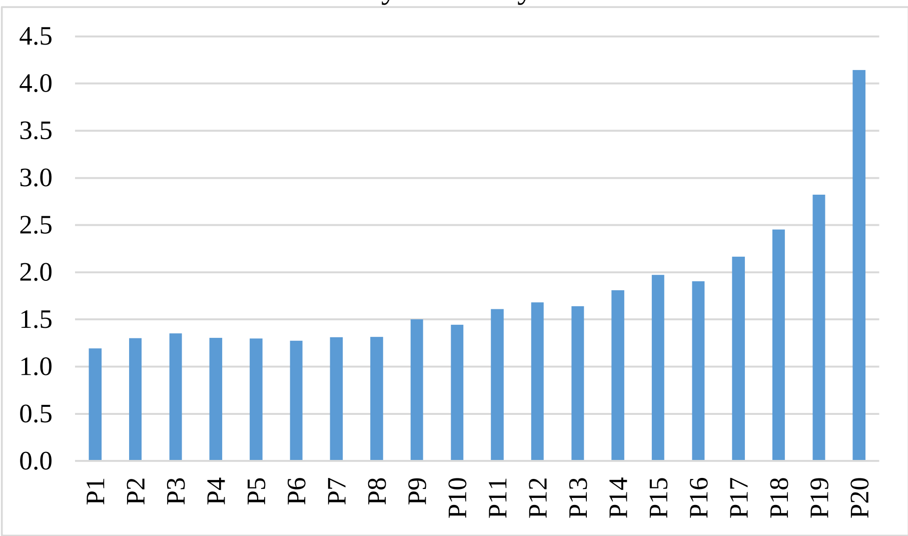


Figure 5. Information Stochastic Discount Factor Estimated with 25 Fama & French Size-Book to Market Portfolios. January 1996-July 2015.

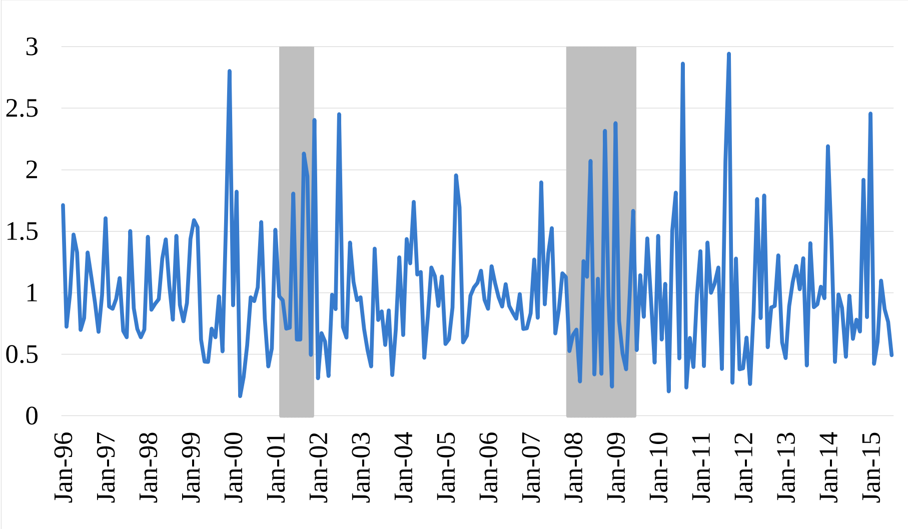


Figure 6.1. Principal Components from the Variance-Covariance Matrix of (Lower Bound) Expected Excess Returns. January 1996-2015.

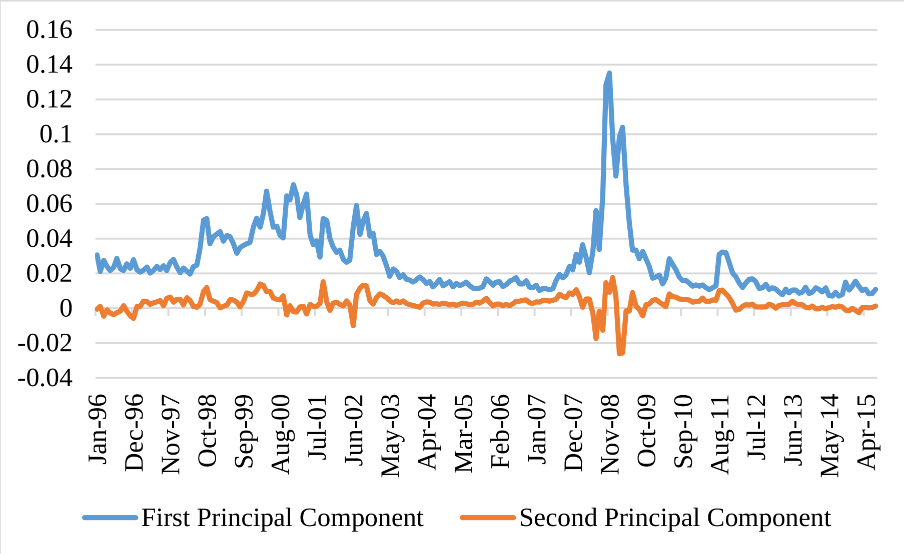


Figure 6.2. Factor Loadings on Principal Components. January 1996-July 2015.

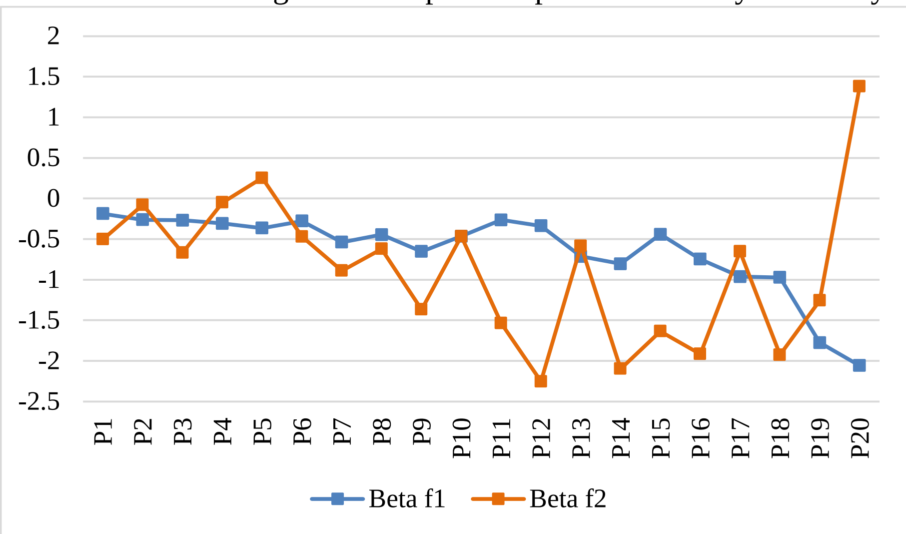


Figure 7.1. The Cross-Section of (Lower Bound) Expected Excess Returns. Two Principal Components Model. January 1996-July 2015.

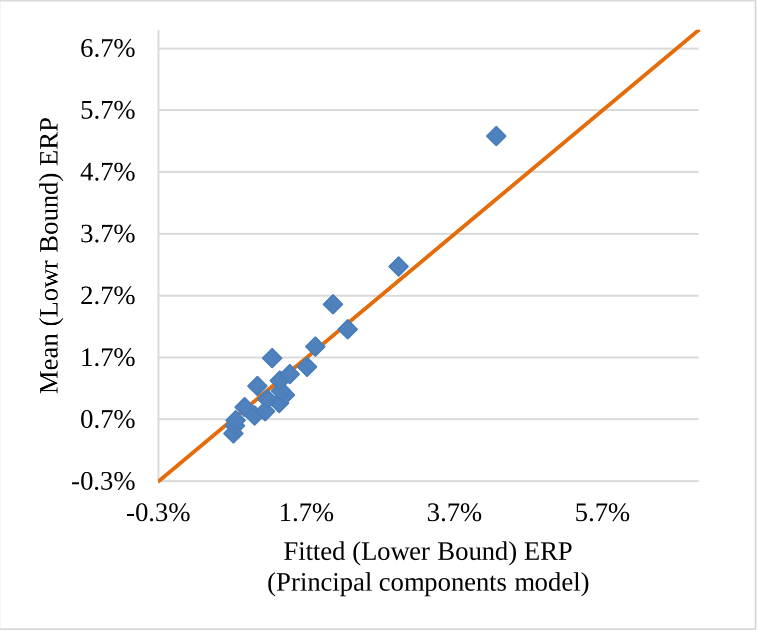


Figure 7.2. The Cross-Section of Average Realized Excess Returns. Two Principal Components. January 1996-July 2015.

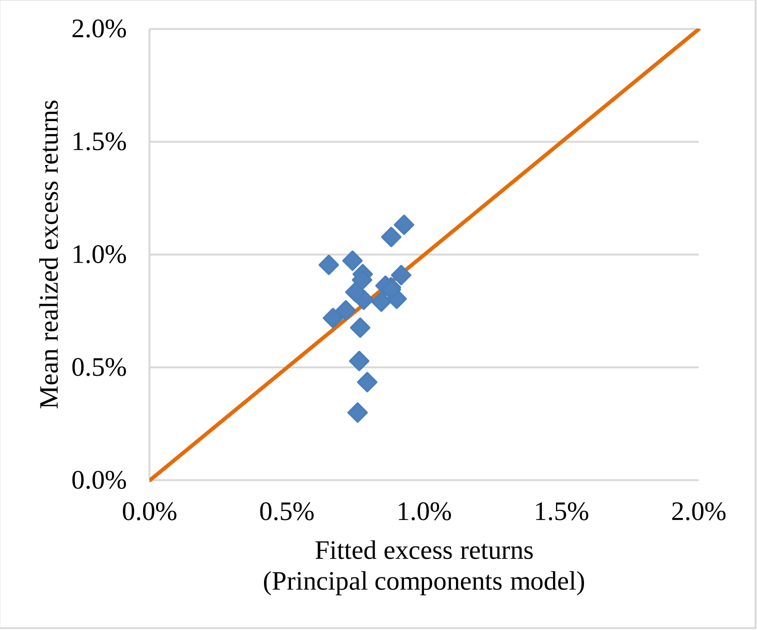


Figure 8.1. The Cross-Section of (Lower Bound) Expected Excess Returns. Information Stochastic Discount Factor Model. ISDF Estimation with 25 Fama and French Portfolios by Size and Book-to-Market. January 1996-May 2015.

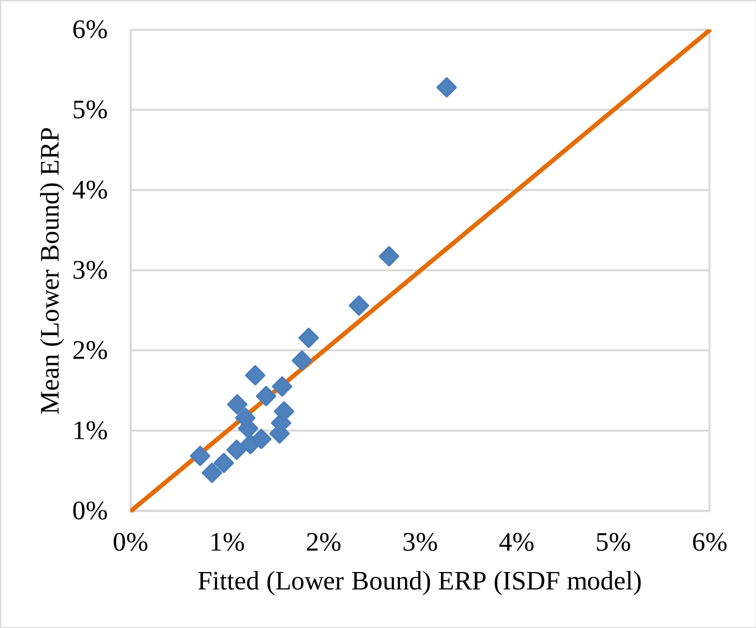


Figure 8.2. The Cross-Section of Average Realized Excess Returns. Information Stochastic Discount Factor Model. ISDF Estimation with 25 Fama and French Portfolios by Size and Book-to-Market. January 1996-May 2015.

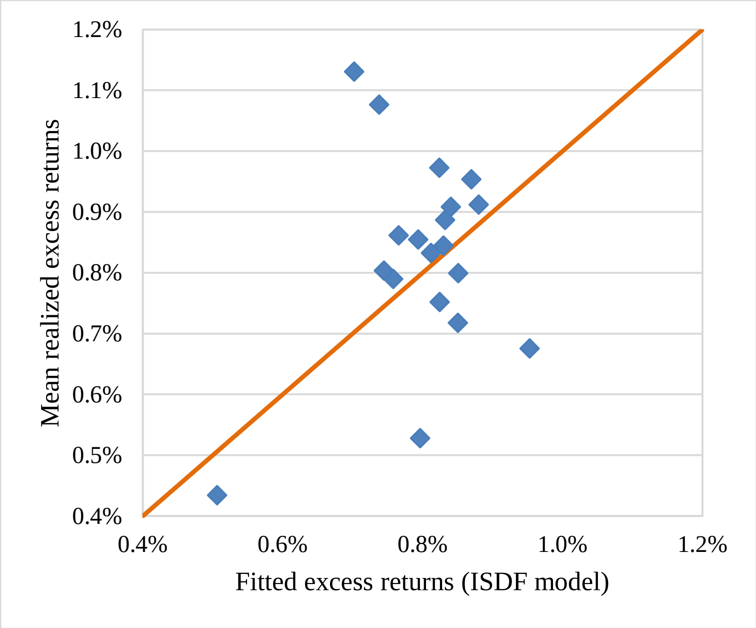


Figure 9.1. The Cross-Section of (Lower Bound) Expected Excess Returns. Multi-Factor Model. January 1996-July 2015.

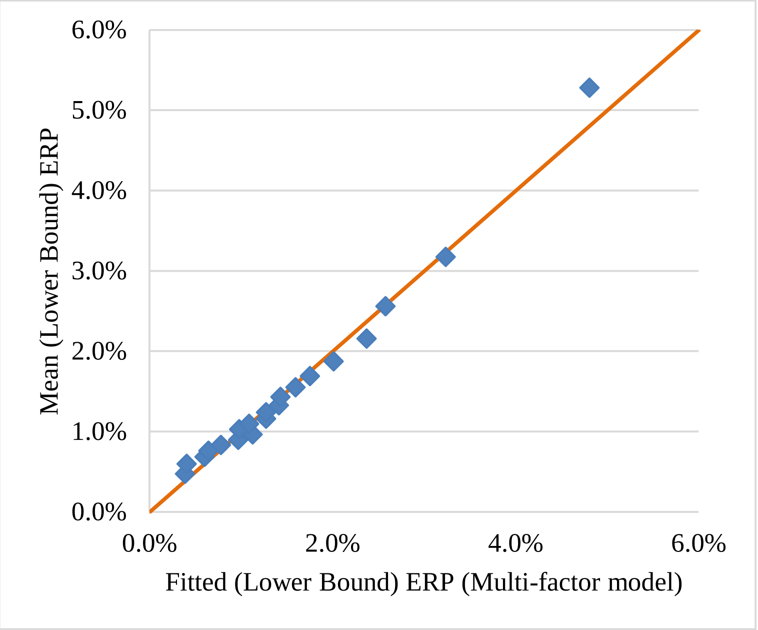


Figure 9.2. The Cross-Section of Average Realized Excess Returns. Multi-Factor Model. January 1996-July 2015.

Figure 10.1. Averages of Expected Risk Premia of 20 Portfolios Sorted by Risk-Neutral Variance. Monthly Data, January 1996-July 2015.

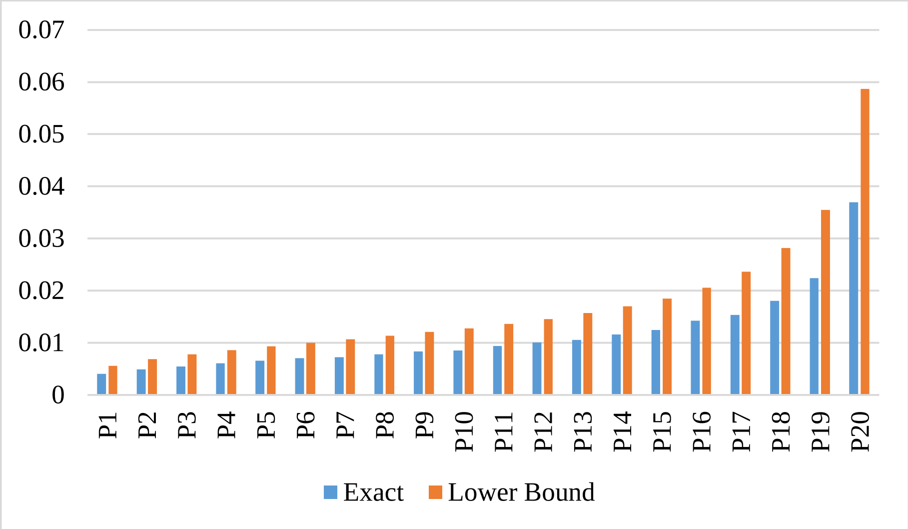


Figure 10.2. Volatilities of Expected Risk Premia of 20 Portfolios Sorted by Risk-Neutral Variance. Monthly Data, January 1996-July 2015.

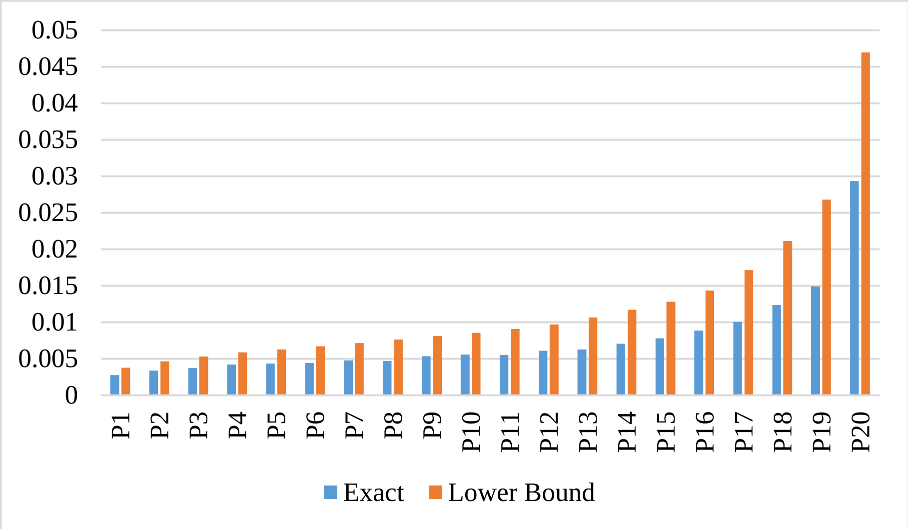


Figure 11.1. First Principal Component from the Variance-Covariance Matrix of Expected Returns. January 1996-July 2015.

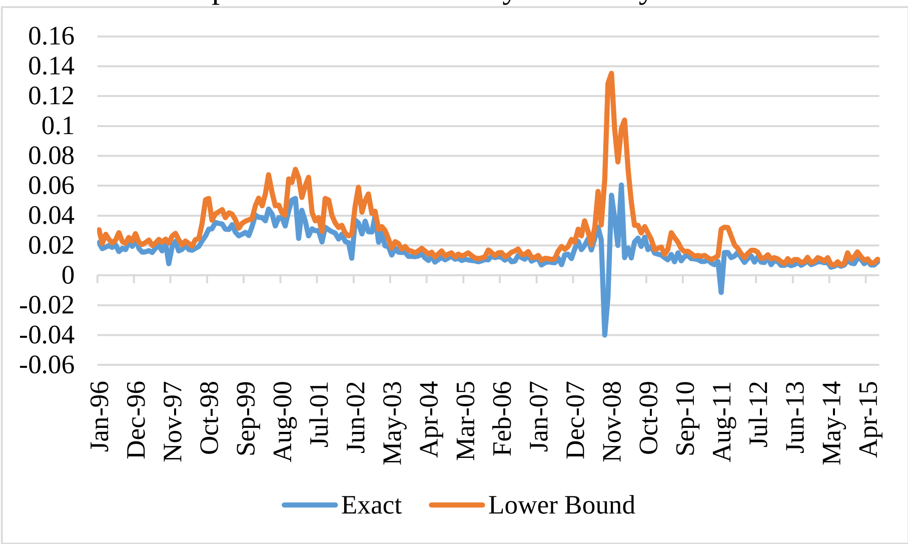
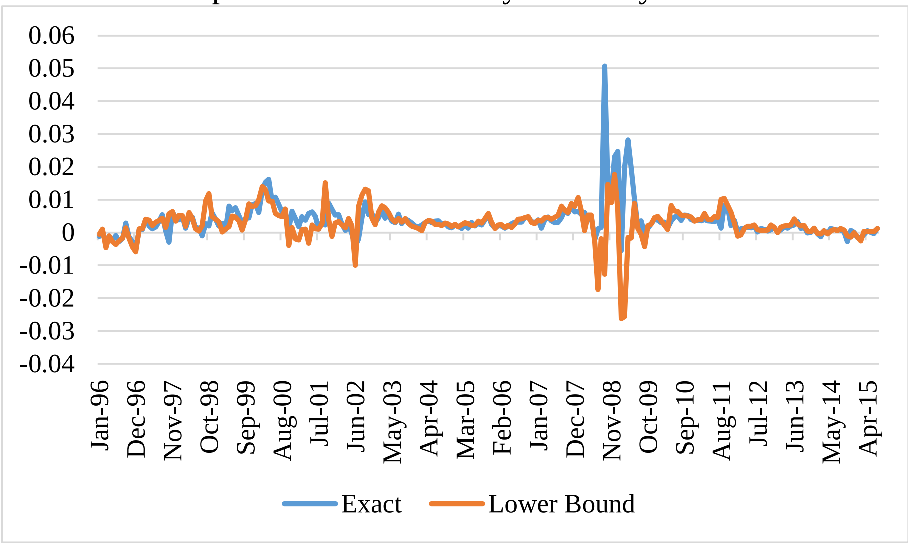


Figure 11.2. Second Principal Component from the Variance-Covariance Matrix of Expected Returns. January 1996-July 2015.



Appendix. The Behavior and Factor Structure of Expected Returns Using the Model-Free Implied Variance of Jiang and Tian (2005)

Martin (2017) shows how to precisely relate the market's expected risk premium to its risk-neutral variance. The resulting volatility index is called SVIX. The square of this index measures risk-neutral variance. On other hand, the square of VIX measures risk-neutral entropy. Previous literature does not report a detailed analysis of the empirical differences generated by both procedures. This appendix covers partially this gap.

We next approximate risk-neutral variances following the insight of Britten-Jones and Neuberger (2000), who derived the model-free option implied variance under diffusion assumptions. Jiang and Tian (2005) extend their paper to show that their method is also valid in a jump-diffusion framework and, therefore, their methodology is a model-free procedure. We obtain the model-free implied variance denoted $MFIV_{t,t+\tau}^{j}$ by the following integral over a continuum of strikes:

$$MFIV_{jt,t+\tau} = 2R_{ft,t+\tau}\int_{0}^{\infty}\frac{C_{jt,t+\tau}\left(K\right)-\max\left(S_{jt}-K/R_{ft,t+\tau},0\right)}{K^{2}}dK \ , \qquad \text{(A.1)}$$

where $C_{jt,t+\tau}\left(K\right)$ is the price at time t of a τ-maturity call option on either an asset or an index j with strike K, and S_{jt} is the spot price of asset j at time t minus the present value of all expected future dividends to be paid before the option maturity. Specific implementation of equation (A.1) follows the approach of Jiang and Tian (2005), and the details are described in González-Urteaga and Rubio (2016).

The difference between both measures is that the risk-neutral variance derived by Martin (2017) is the price of a portfolio of out-the-money options equally-weighted by strike. Expression (A.1) is not the theoretical right measure to calculate risk-neutral variance. It represents the price of a portfolio of out-of-the-money options weighed by the inverse of the square of their strikes. To see this note that we can write expression (A.1) as

$$MFIV_{jt,t+\tau} = 2R_{ft,t+\tau}\left[\int_{0}^{F_{jt,t+\tau}} \frac{1}{K^2} P_{jt,t+\tau}(K)\,dK + \int_{F_{jt,t+\tau}}^{\infty} \frac{1}{K^2} C_{jt,t+\tau}(K)\,dK\right], \quad (A.2)$$

which resembles equation (11) but with a different weighting scheme.

From the estimation of equation (A.1) for each asset in the sample, we follow the same procedure of Section 4. We construct 20 risk-neutral variance-sorted portfolios. As before, P1 contains the assets with the lowest risk-neutral entropy, and P20 includes the stocks with the highest risk-neutral entropy. We update the components every month. We also obtain the corresponding realized returns of the 20 portfolios. Table A.1 contains the descriptive statistics. Note that, in all cases, the patterns of average, volatility, and betas are the same as in Table 1. However, the range of all descriptive statistics is higher than under Martin's procedure. P1 presents lower average statistics and P20 displays higher numbers in all cases. The average (lower bound) expected market excess return is 0.33% or 3.96% on annual basis.

In Table A.2, we report the factor structure obtained from (A.1). The results are very similar. The first two principal components explain practically the same percentage under both procedures. In this case, the first principal component captures a slightly lower variability than before. Table A.3 contains the temporal determinants of the first two principal components. In Panel A of Table A.3, we show that the same four factors, QMJ, BAB, DEF and the market VRP, plus the market explain 60% of the variability of the first principal component. Once again, the results are very similar under both methodologies. The main difference comes from the second principal component. If under the correct procedure, the MOM factor was practically the only factor significantly explaining the time-varying behaviour of the second principal component, we now find that market-wide illiquidity and the market VRP explain the second component. In any case, as before, we are able to capture only 6.6% of the variability of the second component. Finally, in Table A.4, we show that the cross-sectional variability of the approximate (lower bound) expected excess returns extracted from the model-free implied variance is explained by the same set of factors. The traditional *R*-square is 97.7% instead of 98.2% for the results obtained under expression (11). The difference comes from the corrected KRS *R*-square. Under the correct procedure, this is 83.6% but in Table A.4 is lower and equal to 60% approximately. The reason is that the pricing error of portfolio P20 is 0.47% under expression (11) but 0.61% under expression (A.1).

We can therefore conclude that for most empirical results regarding the behaviour of expected returns, both procedures generate similar results. Of course, the only valid theoretical approach if we want to estimate risk-neutral variance is provided by expression (11).

Table A.1. Descriptive Statistics of Approximate (Lower Bound) Expected Risk Premia and Realized Returns for 20 Portfolios Sorted by Model-Free Implied Variance (Jiang & Tian): January 1996-July 2015.

	(Lower Bound) Expected Risk Premium			Realized Returns		
	Average	Volatility	Market ERP Beta	Average	Volatility	Market Beta
P1	0.0024	0.0026	0.5892	0.0117	0.0262	0.3103
P2	0.0034	0.0034	0.7778	0.0131	0.0321	0.4150
P3	0.0040	0.0040	0.9195	0.0131	0.0339	0.4712
P4	0.0046	0.0046	1.0849	0.0134	0.0373	0.5925
P5	0.0051	0.0051	1.2310	0.0144	0.0386	0.6274
P6	0.0057	0.0056	1.3863	0.0135	0.0400	0.6803
P7	0.0062	0.0061	1.5310	0.0129	0.0446	0.8099
P8	0.0068	0.0065	1.6796	0.0145	0.0469	0.8063
P9	0.0073	0.0070	1.8114	0.0137	0.0490	0.9109
P10	0.0080	0.0075	1.9624	00133	0.0505	0.9251
P11	0.0086	0.0081	2.1398	0.0129	0.0548	0.9848
P12	0.0094	0.0090	2.3919	0.0090	0.0566	1.0410
P13	0.0103	0.0098	2.6306	0.0127	0.0569	0.9790
P14	0.0115	0.0112	2.9947	0.0112	0.0636	1.1718
P15	0.0130	0.0131	3.5025	0.0148	0.0650	1.2093
P16	0.0149	0.0156	4.1424	0.0107	0.0687	1.2992
P17	0.0176	0.0190	5.0712	0.0096	0.0755	1.3344
P18	0.0220	0.0238	6.3440	0.0107	0.0886	1.6184
P19	0.0295	0.0329	8.8641	0.0088	0.0958	1.6240
P20	0.0682	0.0868	20.752	0.0167	0.1608	2.6661
MARKET	0.0033	0.0029	1.0000	0.0066	0.0449	1.0000

This table presents the descriptive statistics of 20 portfolios sorted by model-free implied variance and the market, represented by the Standard & Poor 100 Index. The first two columns show the mean and the volatility of the approximate (lower bound) expected risk premia. The third column is the sensitivity of the expected risk premium of the 20 portfolios to the expected market risk premium. The last three columns contain the mean, the volatility and the market beta of realized returns.

Table A.2. The Factor Structure of Approximate (Lower Bound) Expected Risk Premia for 20 Portfolios Sorted by Model-Free Implied Variance (Jiang & Tian). Percentage Explained by the First Five Principal Components: January 1996-July 2015.

	Panel A: Principal Components		Panel B: Principal Components Connor & Korajczyk	
	(Lower Bound) Expected Risk Premia	Realized Excess Returns	(Lower Bound) Expected Risk Premia	Realized Excess Returns
Factor (PC) 1	94.89	72.41	96.64	72.42
Factor (PC) 2	99.40	82.58	99.62	82.87
Factor (PC) 3	99.90	86.71	99.94	87.09
Factor (PC) 4	99.95	88.89	99.97	89.25
Factor (PC) 5	99.97	90.50	99.98	90.79

The two columns of Panel A show the percentage of the variability of the 20 x 20 variance-covariance matrix of approximate (lower bound) expected and realized excess returns of our 20 portfolios, respectively, explained by the first five principal components. The two columns of Panel B report the same percentages of the variability but of the *T x T* covariance matrix, in this case, computed following the procedure of Connor and Korajczyk (1988).

Table A.3. Determinants of the Factor Structure of the Approximate (Lower Bound) Expected Risk Premia (Jiang & Tian). January 1996-July 2015.

Panel A: Determinants of the First Principal Component

Const.	Market	SMB	HML	RMW	CMA	MOM	QMJ	BAB	P&S	DEF	VRP	Adj R^2
0.033	-0.256	0.097	-0.304	0.253	0.281							
(13.39)	(-4.18)	(1.19)	(-2.83)	(2.11)	(1.90)							0.154
[7.25]	[-2.18]	[1.40]	[-1.78]	[2.51]	[1.73]							
0.033	-0.295	0.123	-0.396	0.297	0.348	-0.131						
(13.80)	(-4.77)	(1.53)	(-3.58)	(2.49)	(2.35)	(-2.87)						0.180
[7.30]	[-2.35]	[1.84]	[-2.39]	[2.89]	[2.39]	[-1.81]						
0.031							0.401					
(12.93)							(5.26)					0.102
[7.51]							[2.20]					
0.034								-0.162				
(13.50)								(-2.72)				0.027
[6.52]								[-1.32]				
0.033									0.149			
(13.59)									(4.12)			0.064
[7.13]									[2.40]			
-0.037										2.855		
(-6.07)										(12.10)		0.383
[-3.02]										[4.84]		
0.041											0.033	
(16.43)											(7.63)	0.196
[6.25]											[2.76]	
-0.020	-0.041	0.046	0.069	-0.132	0.077	-0.006	0.400	-0.232	0.021	2.400	0.022	
(-3.59)	(-0.74)	(0.78)	(0.75)	(-0.92)	(0.71)	(-0.17)	(2.86)	(-4.53)	(0.84)	(11.41)	(6.20)	0.601
[-2.57]	[-0.79]	[1.19]	[0.55]	[-1.12]	[0.79]	[-0.21]	[3.68]	[-2.47]	[0.93]	[6.40]	[3.29]	
-0.019	-0.085						0.254	-0.220		2.423	0.023	
(-3.75)	(-1.68)						(3.53)	(-5.09)		(12.36)	(6.50)	0.602
[-2.78]	[-1.70]						[3.43]	[-2.76]		[7.25]	[3.10]	
-0.021							0.321	-0.209		2.461	0.025	
(-4.03)							(5.36)	(-4.88)		(12.60)	(7.51)	0.599
[-2.88]							[3.59]	[-2.68]		[7.08]	[3.21]	

Panel B: Determinants of the Second Principal Component

Const.	Market	SMB	HML	RMW	CMA	MOM	QMJ	BAB	P&S	DEF	VRP	Adj R^2
0.004	-0.033	-0.006	-0.029	-0.011	0.040							
(7.18)	(-2.34)	(-0.35)	(-1.18)	(-0.42)	(1.18)							0.024
[4.59]	[-1.91]	[-0.38]	[-1.12]	[-0.47]	[1.35]							
0.004	-0.030	-0.008	-0.023	-0.015	0.035	0.009						
(7.07)	(-2.09)	(-0.44)	(-0.88)	(-0.53)	(1.03)	(0.86)						0.023
[4.41]	[-1.77]	[-0.46]	[-0.93]	[-0.57]	[1.37]	[0.63]						
0.004							0.025					
(6.94)							(1.48)					0.005
[4.43]							[1.08]					
0.004								-0.010				
(7.21)								(-0.79)				-0.002
[4.14]								[-0.49]				
0.004									0.021			
(7.35)									(2.69)			0.026
[4.78]									[3.65]			
0.002										0.057		
(1.48)										(0.90)		-0.001
[0.94]										[0.49]		
0.005											0.004	
(8.27)											(3.72)	0.052
[5.43]											[2.25]	
0.005	-0.021	-0.014	-0.022	0.021	0.039	0.014	-0.037	-0.021	0.012	0.010	0.003	
(2.52)	(-1.16)	(-0.73)	(-0.72)	(0.45)	(1.10)	(1.17)	(-0.80)	(-1.24)	(1.49)	(0.14)	(2.33)	0.054
[1.90]	[-0.81]	[-0.77]	[-0.65]	[0.44]	[1.13]	[1.08]	[-0.72]	[-0.99]	[2.07]	[0.10]	[1.18]	
0.005									0.016		0.003	
(8.19)									(2.10)		(3.30)	0.066
[5.45]									[2.90]		[1.94]	

This table shows the estimated coefficients of time-series regressions of each of the two first principal components on alternative state variables. The first six variables are the intercept and the Fama and French (2015) factors, MOM is the Momentum Factor of Carhart (1997), QMJ is the quality minus junk Factor of Asness, Frazzini, and Pedersen (2014), BAB is the betting-against-beta factor of Frazzini and Pedersen (2014), P&S is the Pastor and Stambaugh (2003) illiquidity factor, DEF is the default premium, and VRP is the market variance risk premium. This is defined as the logarithm of the realized variance divided by the model-free implied variance. OLS *t*-statistics are reported in parenthesis and *t*-statistics based on HAC standard errors in brackets.

Table A.4. The Cross-Section of Portfolios sorted by Model-Free Implied Variance and a Multi-Factor Asset Pricing Model (Jiang & Tian). January 1996-July 2015.

Panel A: (Lower Bound) Expected Risk Premia

λ_0	λ_m	λ_{qmj}	λ_{bab}	λ_{def}	λ_{vrp}	*R*-square
0.0046	0.0080	-0.0179	-0.0087	-0.0019	-0.1150	0.977
(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.002)	0.597
[0.000]	[0.011]	[0.000]	[0.070]	[0.021]	[0.217]	(0.00)

Panel B: Average Realized Excess Returns

λ_0	λ_m	λ_{qmj}	λ_{bab}	λ_{def}	λ_{vrp}	*R*-square
0.0121	-0.0041	0.0022	-0.0013	-0.0014	-0.0637	0.337
(0.000)	(0.371)	(0.546)	(0.830)	(0.356)	(0.507)	0.273
[0.000]	[0.409]	[0.637]	[0.875]	[0581]	[0.765]	(0.92)

This table reports the risk premia estimates from the two-pass cross-sectional regression of returns on rolling betas associated with the five aggregate risk factors: *m* represents the market excess return, *qmj* is the quality minus junk Factor of Asness, Frazzini, and Pedersen (2014), *bab* denotes the betting-against-beta factor of Frazzini and Pedersen (2014), *def* is the default premium, and *vrp* is the market variance risk premium defined as the logarithm of the realized variance divided by the model-free implied variance. Panel A shows the results using (lower bound) expected risk premia, and Panel B employs excess realized returns. We report the *p*-value in parentheses and the Kan, Robotti, and Shanken (2013) adjusted *p*-value in brackets. The cross-sectional *R*-square reported in the first line is 1 minus the cross-sectional variance of average pricing errors divided by the cross-sectional variance of the average dependent variable. Kan, Robotti, and Shanken (2013) propose the cross-sectional *R*-square reported in the second line, and the corresponding (asymptotically valid) *p*-value is given in parentheses.