Kernel Methods and Flight-to-Safety in Financial Asset Returns

Ryan McCorvie[*]

UC Berekeley – Stat 241a

November 1st, 2016

Contents

1 Flight-to-Safety 1

2 Factor Models and Gaussian Processes 2

3 Model Inference 4

4 Methods 5

5 Results 7

6 Analysis and Conclusions 9

References 10

1 Flight-to-Safety

During periods of high market uncertainty, a common phenomenon is for investors to investors liquidate positions in more risky assets in favor of positions in "safe" assets (such as US Treasury bonds). The upshot is that many asset classes may simultaneously experience price shocks, though their movements may not typically be correlated. The dimensions of risk and safety may take many forms, and phenomenon sometimes called flight-to-safety or flight-to-liquidity. For a dynamic equilibrium economic model which gives rise to flight to safety behavior, see [Vay04]. For an analysis of how flight-to-safety may be a a transmission mechanism for macroeconomic shocks, see [BGG94].

The time series in figure 1 exhibits two episodes of flight-to-safety dynamics. In late 1998, the market reacted to the sudden insolvency of the Long-term Capital Management hedge fund. In late 2001, the unwinding of the dot-com bubble led investors to seek safety. In both episodes US treasury bond cumulative returns sharply spiked, while equity returns sharply declined. Through each episode, VIX measures of market uncertain rose to atypically high levels.

[*]mccorvie@berkeley.edu

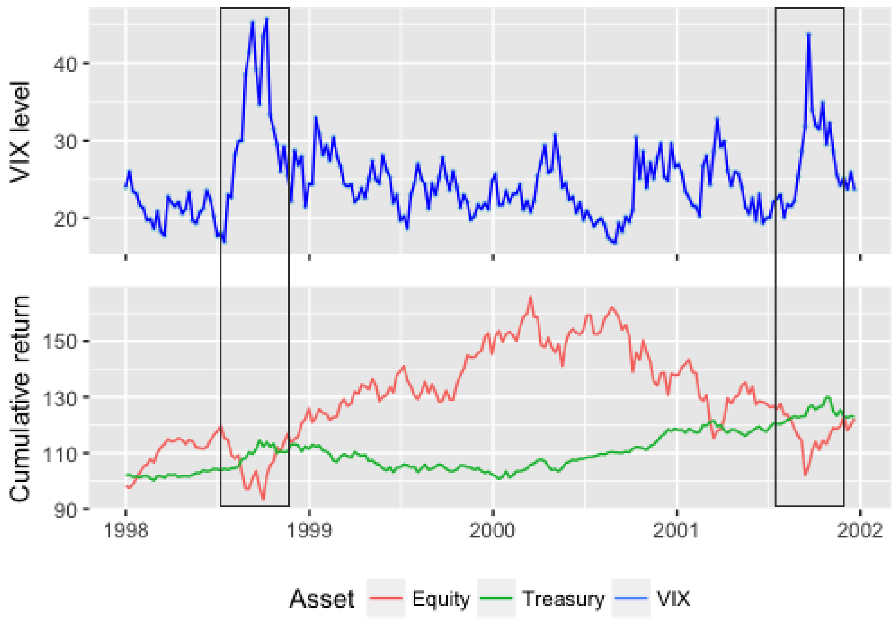

Figure 1: Episodes of flight-to-safety

Flight-to-safety behavior is distinguished from typical asset price movements by two qualities First, it happens only occasionally and conditionally, in the face of news or innovations. At these times, market participants to suddenly reevaluate the risk of their portfolios. Second, flight-to-safety can simultaneously affect many asset classes, leading to simultaneous declines in assets which may not usually exhibit much correlation. This makes flight-to-safety an especially important consideration for risk management, since many common risk metrics (such as value-at-risk) are based on typical market correlations. Thus, the phenomenon of flight-to-safety may result in unexpected losses in portfolios which are thought to be well-diversified.

2 Factor Models and Gaussian Processes

A common approach for analyzing returns across a broad portfolio assets is to attribute those returns to a small number of factors. Let r_i be a random variable representing the asset returns for the ith asset. In their basic form, a factor model relates the returns r_i is related to a vector of factor returns $(f_1, f_2, \ldots, f_d)$ by a linear equation

$$(1) \qquad r_i = \sum_{k=1}^{d} \beta_{i,j} f_j + \epsilon_i \qquad \text{or} \qquad \boldsymbol{r} = \boldsymbol{B}\boldsymbol{f} + \boldsymbol{\epsilon}$$

where the $\beta_{i,j}$ are constants. Here the ϵ_i is a random variable which represents the residual returns for the ith asset. The factors may be latent, and determined statistically, or they may be separately observable macroeconomic or fundamental data. By taking many observations of the returns, its possible to fit the factor loadings $\beta_{i,j}$. See [CR82, CK88] for theoretical and empirical analysis of factor models for US equity returns.

A taxonomy of commonly used types of factors is given in [Con95]. None of these captures the notion of flight-to-safety very well. For example, statistical or latent factors result in a correlation structure which is independent of market uncertainty levels. On the other hand macroeconomic factors may not reflect the nonlinear dependence of flight-to-safety behavior as uncertainty crosses a threshold. The paper [ACV15] proposes a modification to avoid the problems of traditional factor models.

$$(2) \qquad r_i = \alpha_i + \beta_i \phi(v) + \epsilon_i \qquad \text{or} \qquad \boldsymbol{r} = \boldsymbol{\alpha} + \boldsymbol{\beta} \phi(v) + \boldsymbol{\epsilon}$$

Here v is the VIX, a macroeconomic variable which captures the level of overall market uncertainty, and ϕ is an unknown nonlinear function, which is determined during model calibration. Using a variety of econometric regressions, [ACV15] show that regressions to a non-linear function $\phi(v)$ have a significant factor loadings $\beta \neq 0$, whereas linear regression against v alone do not. The function ϕ has increases for higher levels of v, and the coefficients β_i are negative for risky assets such as equity market returns and positive for safe assets such as US Treasuries. These qualities suggest that $\phi(v)$ is capturing a flight-to-safety response.

In this paper, we use kernel methods to find the factor loadings α, β and the functional form of ϕ. These techniques allow for much greater deal of flexibility in modeling ϕ, which is limited to low-degree polynomials and B-splines in [ACV15]. In contrast, kernel methods allow for non-parametric modeling of infinite dimensional spaces, and also allow for precise control of the regularity properties of ϕ. However, a weakness of the techniques in this paper is they do not allow for analysis of the significance of the factor loadings. See the conclusion for further discussion of the tradeoffs.

A *Gaussian process* (X_t) for $t \in T$ for some parameter space T is any process where, for any $n \in \mathbb{N}, t_1, t_2, \ldots, t_n \in T$, the random vector $(z_{t_1}, z_{t_2}, \ldots, z_{t_n})$ has a (multivariate) normal distribution. These distributions are completely characterized by the functions $\mu(t) = \mathrm{E}\, z_t$ and $k(t,s) = \mathrm{Cov}(z_t, z_s)$. A necessary and sufficient condition on k is that the matrix given by $\boldsymbol{K}_{i,j} = k(t_i, t_j)$ be positive semidefinite for every $\{t_1, \ldots, t_n\} \subset T$. See [Kal06, SK07] for more background.

In the context of machine learning and statistical inference, a Gaussian process provides a mechanism for assigning prior probabilities to a class of functions. A wide variety of processes are encompassed by Gaussian processes, including Brownian motion, Langevin processes, Wiener processes, Kalman filters, and others. A large survey of kernel functions and the resulting properties of the function space in terms of symmetry, continuity and differentiability can be found in [Abr97]. While the function space may be infinite dimension, model inference remains tractable since we only need consider the finite dimensional joint Gaussian distribution of $\phi(v_t)$ at the observed data points. See

[Mac03, Bis06] for information.

We take the kernel to have the following form, which is a popular choice for many machine learning applications [Bis06]

$$k(u, v) = \theta_0 \, e^{-\theta_1 (u-v)^2} + \theta_2 \, uv + \theta_3 \tag{3}$$

In this study, we take the parameters to be $\theta_0 = 1, \theta_1 = 0.005, \theta_2 = 0$ and $\theta_3 = 20$. The Gaussian process exhibits continuity, smoothness, and a reasonable amount of "stiffness" over the range of values observed for the VIX. See the conclusions section for more discussion of this kernel function.

3 Model Inference

Let $r_{t,i}$ be the returns of the ith asset at time t. Let $z_t = \phi(v_t)$ to be a random Gaussian process with $\mu = 0$ and $k(u, v)$ given by (3). Let v_t be the macroeconomic variable which captures the market uncertainty level at time t

$$p(\boldsymbol{z} \mid \boldsymbol{v}) \sim \mathcal{N}(0, \boldsymbol{K}) \qquad \text{where} \qquad K_{i,j} = k(v_i, v_j) \tag{4}$$

For each $t \in \{1, \ldots, T\}$.

$$p(\boldsymbol{r}_t \mid z_t) \sim \mathcal{N}(\boldsymbol{\alpha} + \boldsymbol{\beta} \, z_t, \boldsymbol{\Psi}) \tag{5}$$

where $\boldsymbol{\alpha}, \boldsymbol{\beta}$ are the factor loadings and $\boldsymbol{\Psi} = \mathrm{diag}(\psi_1, \ldots, \psi_d)$ is a diagonal matrix representing the volatility of the residuals, and none of these parameters depend on time t.

We summarize some of the features of the model

1. The dependence across returns comes entirely exposure to the nonlinear transformation $z_t = \phi(v_t)$.
2. For fixed t, the returns $r_{t,i}$ are conditionally independent across all assets given z_t
3. For fixed i, the residual error values $\epsilon_{t,i} = r_{t,i} - \alpha_i - \beta_i \, z_t$ are independent and identically distributed according to $\mathcal{N}(0, \psi_i)$ for some variance ψ_i

Equations (4) and (5) allow us to find the conditional posterior distribution for z, which is Gaussian

$$p(\boldsymbol{z} \mid \boldsymbol{r}) = \frac{P(\boldsymbol{r} \mid \boldsymbol{z}) P(\boldsymbol{z})}{P(\boldsymbol{r})} \sim \mathcal{N}\left(\boldsymbol{\Lambda} \gamma, \boldsymbol{\Lambda}\right) \tag{6}$$

where

$$\begin{aligned} \gamma_t &= \sum_{i=1}^{d} \frac{\beta_i}{\psi_i} (r_{t,i} - \alpha_i) \\ \boldsymbol{\Lambda} &= \boldsymbol{K} - s\boldsymbol{K} \left(\boldsymbol{I} + s\boldsymbol{K}\right)^{-1} \boldsymbol{K} \qquad \text{with} \qquad s = \sum_{i=1}^{d} \frac{\beta_i^2}{\psi_i} \end{aligned} \tag{7}$$

Note that if $\boldsymbol{K}$ is invertible, then the second equation is the same as $\boldsymbol{\Lambda}^{-1} = \boldsymbol{K}^{-1} + s\boldsymbol{I}$.

The statistical inference problem is to determine the values of $\phi, \boldsymbol{\alpha}, \boldsymbol{\beta}$ and $\boldsymbol{\Psi}$ from the observed values v_t and $r_{t,i}$. To find the model parameters, we use the EM algorithm to iteratively converge to the maximum likelihood parameters. Given the latent values z_t, then the complete log-likelihood is given by

$$l_c(\boldsymbol{r}, \boldsymbol{z}) = -\frac{T}{2}\sum_i \log|\psi_i| - \frac{1}{2}\sum_{t,i}\frac{(r_{t,i} - \alpha_i - \beta_i z_t)^2}{\psi_i} + \dots \tag{8}$$

where the additional terms don't depend on the model parameters. To perform the M-step, take expectations of l_c over some probability distribution for $\boldsymbol{z}$ and maximize $\langle l_c \rangle$ with respect to α_i and β_i to get

$$\begin{pmatrix} 1 & \frac{1}{T}\sum_t \langle z_t \rangle \\ \frac{1}{T}\sum_t \langle z_t \rangle & \frac{1}{T}\sum_t \langle z_t^2 \rangle \end{pmatrix} \begin{pmatrix} \alpha_i \\ \beta_i \end{pmatrix} = \begin{pmatrix} \frac{1}{T}\sum_t r_{t,i} \\ \frac{1}{T}\sum_t \langle z_t \rangle r_{t,i} \end{pmatrix} \tag{9}$$

Note that if $\boldsymbol{z}$ is deterministic, these are the normal equations for linear regression. Maximizing $\langle l_c \rangle$ with respect to ψ_i gives

$$\begin{aligned} \psi_i &= \left\langle \frac{1}{T}\sum_t (r_{t,i} - \alpha_i - \beta_i z_t)^2 \right\rangle \\ &= \frac{1}{T}\sum_t (r_{t,i} - \alpha_i)^2 - 2\beta_i (r_{t,i} - \alpha_i)\langle z_t \rangle + \beta_i^2 \left\langle z_t^2 \right\rangle \end{aligned} \tag{10}$$

Thus, (9) and (10) show that $\langle z_t \rangle$ and $\left\langle z_t^2 \right\rangle$ are sufficient statistics.

We can find these from the posterior distribution (7), which gives

$$\langle z_t \rangle = \mathbb{E}[z_t \mid \boldsymbol{r}] \qquad \left\langle z_t^2 \right\rangle = \mathbb{E}[z_t^2 \mid \boldsymbol{r}] = \mathrm{Var}[z_t \mid \boldsymbol{r}] + \mathbb{E}[z_t \mid r]^2 \tag{11}$$

These equations comprise the E-step.

4 Methods

As our proxy for market uncertainty, we take the VIX Volatility Index, which is calculated by the Chicago Board Options Exchange. The index is widely viewed as a "fear index" which gauges general market uncertainty [Wha00]. This value of this index comes from the price of equity variance swaps on the S&P 500 index, but it may be considered an average of the option implied volatility. See figure 2 for a plot of the VIX levels over the analysis period. Note that the history is punctuated by several episodes where VIX sharply spikes upward for a short period.

We collected returns data from a large number of asset classes. Our primary data sets comes from the Center for the Research in to Securities Prices (CSSP). First, we incorporated their value-weighted equity market total return series as a measure for stock market returns. This series, which is similar to the S&P 500 or the Wilshire 5000, and includes a wide variety of US equity shares

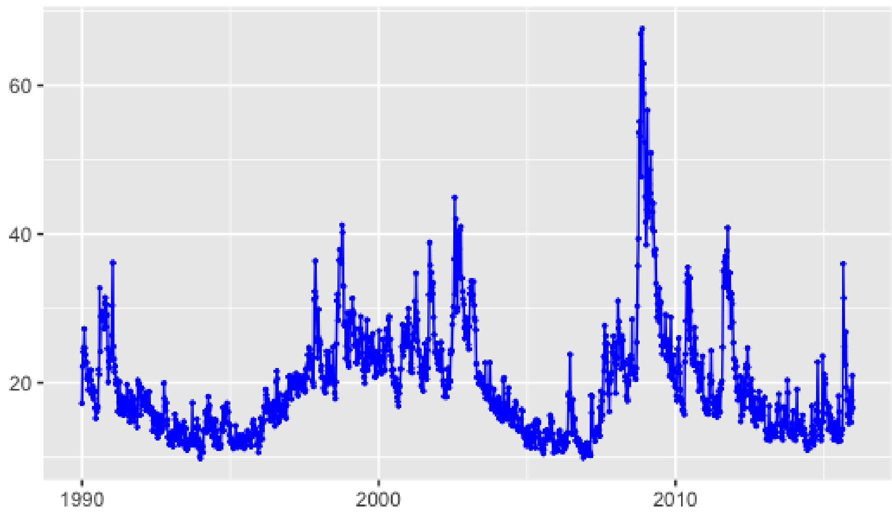

Figure 2: VIX time series

from the NYSE, Nasdaq and other markets. Its construction incorporates total distributions, such as dividends, and accounts for new issues and delistings. For US Treasury bond returns, we take the CSRP constant-maturity total returns series. These are constructed by fitting a yield curve to all outstanding Treasury bond prices, and interpolating to a constant-offset date.

For corporate bond returns, we use the Bank of America Merill Lynch corporate bond total return indexes. For low quality bonds we take the US High Yield Master II index and for investment grade bonds we take the US Corporate Master index. For precious metals, we use the Goldman Sachs Commodities Index (GSCI) precious metals total returns index. Finally, to get more granularity on equity market returns, we take the price indexes for ten S&P 500 industry sector portfolios. The companies are sorted into portfolios based on their Global Industry Classification Standard (GICS) code. While these are prices, not total returns, the returns from these prices should nevertheless give an approximate indication of the flight-to-safety behavior.

We restrict the data sets to the dates between January 1, 1990 and December 31, 2015. This start date coincides with the inception of the current version of the VIX index. In order to avoid problems related to missing or non-aligned data, we calculate cumulative weekly returns for each series, assigning zero returns to days with missing data. This should also avoid problems stemming from different times series being calculated at different times during the day. The VIX on a particular date is regressed against the returns for the subsequent week, starting on that day, and ending on the date the next VIX level is observed. In all, our model is calibrated using 1356 observation dates over 22 asset return categories.

Model calibration was performed on a MacBook Air personal computer with a 2.2 GHz Intel Core i7 processor. The model calibration software was scripted using the R programming language. The underlying BLAS linear algebra routines incorporated the vecLib library, which is a part of Apple's Accelerate framework. Model calibration typically took less than 30 seconds.

5 Results

The primary results of the kernel regression are presented in table 1. Note the signs of the regression coeffients β broadly correspond to notions of which assets are considered safe and risky. Thus, equity indices and high yield corporate bond show a negative loading to the flight-to-safety factor, whereas US Treasuries and investment grade bonds show a positive loading. One exception are stock prices in the energy sector show a positive loading, albeit with low r^2. Also, precious metals, which are often thought of as a safe storehouse, have a negative loading, with small r^2.

Table 1: Factor model parameters

Asset class	β	α	r^2 (%)
Equity-ValueWeight	-0.777	0.003	0.16
SP-Energy	0.584	0.001	0.10
SP-Materials	-0.892	0.002	0.15
SP-Industrials	-1.022	0.002	0.10
SP-ConsumerDisc	-0.221	0.002	0.08
SP-ConsumerStap	-0.424	0.002	0.11
SP-Healthcare	-0.793	0.002	0.17
SP-Financials	-2.348	0.003	0.43
SP-Technology	-1.442	0.003	0.20
SP-Telecoms	-0.351	0.001	0.09
SP-Utilities	-0.135	0.001	0.08
Corp-HighYield	-1.577	0.003	2.64
UST-1y	0.086	0.001	5.59
UST-2y	0.233	0.001	9.32
UST-5y	1.048	0.000	2.65
UST-7y	1.792	0.000	4.14
UST-10y	1.943	0.000	3.40
UST-20y	3.033	-0.001	3.88
UST-30Y	3.982	-0.002	4.38
Corp-InvGrade	0.758	0.001	1.07
GSCI-Metals	-0.297	0.000	0.08

In general, the r^2 of the regression is fairly low, less than 2%. This may reflect the fact that flight-to-safety episodes are somewhat uncommon. Overall, the model provides the most explanatory power for treasury bonds. The regression coefficients α are fairly close to 0, reflecting the the fact that asset returns themselves tend to have mean close to 0.

The non-linear response function ϕ for the flight-to-safety factor is shown in figure 3. This plot gives the 95th percentile range for $z_t = \phi(v_t)$ at each observed value v_t for the VIX. While the right portion of the graph exhibits more wild oscillations, the bulk of the observations are for smaller values of the VIX,

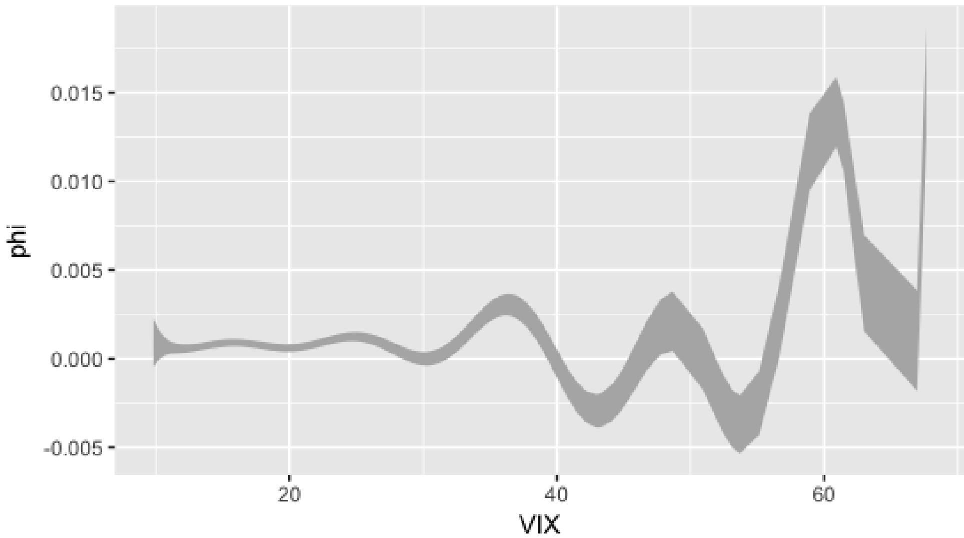

Figure 3: The response function ϕ vs VIX

near the left axis. Note that according to table 2, 90% of the observations of the VIX are less than 29. This corresponds to the portion where the loading is essentially flat and near zero. Thus, in normal times, the flight-to-safety factor has almost no influence on asset returns. Another 7% of observations correspond to the hump between 30 and 40, and in this range the main effects of the flight-to-safety factor occur. Judging by the spikes in figure 2, these periods correspond to periods of greater uncertainty. Only a small fraction of data corresponds to high levels of the VIX. However, in these regions the factor is generally positive, especially for levels around 60, which were only observed during the 2008 financial crisis.

Table 2: VIX quantiles

Quantile	50%	90%	97%	99%	99.5%
VIX level	18.12	29.12	37.68	47.15	55.47

Turning again to the β values, its worthwhile to compare the non-linear factor loadings to more traditional linear factors. In table 3 we compare the CAPM market factor loading to the flight-to-safety hedge ratio. Since the CAPM loading corresponds to a factor model where the single factor is the value weighted equity market index return, and coefficient is also commonly called beta, but it shouldn't be confused with our model β's. The flight-to-safety hedge ratio is simply the ratio of model β's to the flight-to-safety factor, and represents the number of units of the market portfolio required to offset the exposure to the flight-to-safety factor.

In general, the CAPM beta's track the flight-to-safety hedge ratios, so a market neutral portfolio roughly corresponds to a portfolio which would be hedged against flight-to-safety shocks. However, there are some notable ex-

Table 3: Market beta vs. flight-to-safety beta

Sector	CAPM beta	FTS hedge ratio
Energy	0.812	-0.752
Materials	1.020	1.148
Industrials	1.040	1.315
ConsumerDisc	1.065	0.284
ConsumerStap	0.520	0.546
Healthcare	0.728	1.021
Financials	1.276	3.022
Technology	1.299	1.856
Telecoms	0.792	0.452
Utilities	0.567	2.020

ceptions. Financial stocks seem to have significantly more sensitivity to flight-to-safety shocks vs. normal market movements, where consumer discretionary stocks have significant lower sensitivity. Perhaps this is because flight-to-safety episodes are primarily related to financial market uncertainty, and are somewhat disconnected from the real economy.

6 Analysis and Conclusions

The results of this study confirm and support the results of [ACV15]. Though the present study uses different methods, asset classes, and returns periods, we find a broadly similar nonlinear function $\phi(v_t)$ and broadly similar factor loadings for the model (2). The current approach does not give p-values to measure the significance, so its difficult to compare the approaches on purely statistical grounds. This is a direction for future study.

One important open question is how to best chose the kernel function (3). The author experimented with several variations before settling on this particular form. This selection resulted in a continuous, smooth, and relatively stiff response function ϕ, which still retained flexibility in its overall shape. Choosing a more flexible class of functions allows for higher likelihood, but it reduces the ability to interpret the loadings, and may be susceptible to overfitting. Overall, there is a bias/variance trade off in selecting the class of Gaussian processes. A direction for future research is to select the kernel and its hyperparameters using more statistical approaches.

A major potential application for the flight-to-safety factor is to improve risk management of asset portfolios. Knowing potential exposures during flight-to-safety situations allows is valuable for risk managers. Thus by managing total flight-to-safety exposure can complement traditional approaches using linear correlations, such as value-at-risk and factor models. The flight-to-safety factor can also provide a statistical basis for the construction of stress-test scenarios. Finally the methods of the current study can be extended to find nonlinear dependence on other variables. A nonlinear regression on Carhart's 4 factors may lead to better forecasting of asset returns [Car97].

The present study finds evidence that there is a relationship between asset returns during periods of unusual market uncertainty, which is different from the linear correlation observed normally. Thus, the present study provides a quantitative framework for analyzing flight-to-safety behavior, using kernel regression methods.

References

[Abr97] Petter Abrahamsen. A review of gaussian random fields and correlation functions. Technical Report 917, Norwegian Computing Center, 1997.

[ACV15] Tobias Adrian, Richard K Crump, and Erik Vogt. Nonlinearity and flight to safety in the risk-return trade-off for stocks and bonds. Technical Report FEDNSR723, FRB of New York Working Paper, 2015.

[BGG94] Ben Bernanke, Mark Gertler, and Simon Gilchrist. The financial accelerator and the flight to quality. Technical report, National Bureau of Economic Research, 1994.

[Bis06] Christopher M. Bishop. *Pattern Recognition and Machine Learning*. Springer, 2006.

[Car97] Mark M Carhart. On persistence in mutual fund performance. *The Journal of finance*, 52(1):57–82, 1997 1997.

[CK88] Gregory Connor and Robert A. Korajczyk. Risk and return in equilibrium apt. *Journal of Financial Economics*, pages 255–289, 1988.

[Con95] Gregory Connor. The three types of factor models: a comparison o ftheir explanatory power. *Financial Analysis Journal*, May / June 1995.

[CR82] Gary Chamberlain and Michael Rothschild. Arbitrage, factor structure, and mean-variance analysis on large asset markets. Technical Report 996, National Bureau of Economic Research, 1982.

[Kal06] Olav Kallenberg. *Foundations of Modern Probability*. Springer Science & Business Media, 2006.

[Mac03] David J.C. MacKay. *Information theory, inference and learning algorithms*. Cambridge University Press, 2003.

[SK07] Cosma Shalizi and Alex Kontorovich. *Almost None of the Theory of Stochastic Processes*. 2007.

[Vay04] Dimitri Vayanos. Flight to quality, flight to liquidity, and the pricing of risk. Technical report, National Bureau of Economic Research, 2004.

[Wha00] Robert E Whaley. The investor fear gauge. *The Journal of Portfolio Management*, 26(3):12–17, 2000.