Flights to Safety*

Lieven Baele[1] Geert Bekaert[2] Koen Inghelbrecht[3]
Min Wei[4]

June 2014

Abstract

Using only daily data on bond and stock returns, we identify and characterize flight to safety (FTS) episodes for 23 countries. On average, FTS days comprise less than 3% of the sample, and bond returns exceed equity returns by 2.5 to 4%. The majority of FTS events are country-specific not global. FTS episodes coincide with increases in the VIX and the Ted spread, decreases in consumer sentiment indicators and appreciations of the Yen, Swiss franc, and US dollar. The financial, basic materials and industrial industries under-perform in FTS episodes, but the telecom industry outperforms. Money market instruments, corporate bonds, and commodity prices (with the exception of metals, including gold) face abnormal negative returns in FTS episodes. Hedge funds, especially those belonging to the "event-driven" styles, display negative FTS betas, after controlling for standard risk factors. Liquidity deteriorates on FTS days both in the bond and equity markets. Both economic growth and inflation decline right after and up to a year following a FTS spell.

JEL Classification: G11, G12, G14, E43, E44

Keywords: Flight-to-Safety, Flight-to-Quality, Stock-Bond Return Correlation, Liquidity, Hedge Funds

* The authors greatly benefited from discussions with seminar participants at Antwerp University, the National Bank of Belgium, the University of the Basque Country, Tilburg University, the Spring 2014 Inquire Europe-UK conference in Vienna, and the 2014 Bank of Canada/Banco de España Workshop on "International Financial Markets", and in particular with Hans Dewachter, Frank de Jong, Joost Driessen, Eric Ghysels, Campbell Harvey, Philippe Mueller, Laurens Swinkels and Allan Timmermann. Baele, Bekaert, and Inghelbrecht gratefully acknowledge financial support from the National Bank of Belgium and Inquire Europe. Bekaert acknowledges financial support from Netspar. The opinions expressed are those of the authors and do not necessarily reflect views of the National Bank of Belgium or the Federal Reserve System.

[1] Finance Department, CentER, Tilburg University. E-mail: Lieven.Baele@uvt.nl

[2] Graduate School of Business, Columbia University, and NBER. E-mail: gb241@columbia.edu

[3] Department Financial Economics, Ghent University. E-mail: Koen.Inghelbrecht@UGent.be

[4] Federal Reserve Board of Governors, Division of Monetary Affairs, Washington DC. Email: min.wei@frb.gov

1 Introduction

In periods of market stress, the financial press interprets extreme and inverse market movements in the bond and equity markets often as “flights to safety” or “flights to quality.” In particular, between August 2004 and June 2012, a period marred by a global financial crisis, the Financial Times referred 805 times to “Flight(s)-to-Quality” and 533 times to “Flight(s)-to-Safety.”

There is an active theoretical academic literature studying such phenomena. In Vayanos (2004)’s model, risk averse investment managers fear redemptions during high volatility periods and therefore an increase in volatility may lead to a “flight-to-liquidity.” At the same time, their risk aversion also increases, leading to a “flight-to-safety,” meaning that they require higher risk premiums, which in turn drives down the prices of risky assets (a flight to quality). In Caballero and Krishnamurthy (2008), Knightian uncertainty may lead agents to shed risky assets in favor of uncontingent and safe claims when aggregate liquidity is low thereby provoking a flight to quality or safety. Brunnermeier and Pedersen (2009) study a model in which speculators, who provide market liquidity, have margin requirements increasing in volatility. They show how margin requirements can help cause a liquidity spiral following a bad shock, where liquidity deteriorates in all markets, but also a flight to quality, which they define as a sharp drop in liquidity provision for the high margin, more volatile assets. Representative agent models can also generate “flights-to-safety.” In the consumption based asset pricing literature (e.g. Barsky (1989); Bekaert et al. (2009)) a flight to safety is typically defined as the joint occurrence of higher economic uncertainty (viewed as exogenous) with lower equity prices (through a cash flow and/or risk premium effect) and low real rates (through a precautionary savings effect).

These articles seem to treat flights to quality, safety and/or liquidity as Justice Potter treated porn: we know it when we see it. However, to be able to test and refute a diverse set of theoretical models, an empirical characterization of flight to safety episodes would appear essential. The goal of our paper is to define, detect and characterize flight-to-safety episodes for 23 countries. In doing so, we use only daily data on the prototypical risky asset (a well-diversified equity index) and the prototypical safe and liquid asset (the benchmark Treasury bond). Beber et al. (2009) use the Euro-area government bond market to show that in times of market stress, investors demand liquidity rather than credit quality. Longstaff (2004), focusing on the US Treasury market, shows that the liquidity premium in Treasury bonds can represent up to 15% of their value. In other words, flights to safety may be as much or more about flights to liquidity than about flights to quality. It is therefore important to focus on a liquid bond benchmark in our work. To define a flight to safety, referred to as FTS henceforth, we use the simple criteria that it happens during periods of market stress (high equity market volatility), entails a large and positive bond return, a large and negative equity return, and negative high-frequency correlations between bond and stock returns. Note that stock and bond returns are likely positively correlated outside the flights-to-safety periods as both represent high duration assets. Negative aggregate demand shocks may also entail negative stock-bond return correlations but will only be identified as FTS when accompanied by substantial market stress.

We use a plethora of econometric techniques, detailed in Sections 2.2 and 2.3, to identify flight-to-safety episodes from these features. In Section 2.4, we then analyze the identified flight to safety episodes in 23 countries in more detail. We find that FTS episodes comprise less than 3% of the sample on average, and bond returns exceed equity returns by about 2.5 to 4% on FTS days. Only a minority of FTS events can be characterized as global (less than 25% for most countries). About 89% of FTS days lasts 3 days or less, but a small fraction lasts longer than 10 days. In section 2.5, we show that FTS episodes coincide with increases in the VIX and the TED spread, decreases in consumer sentiment indicators in the US, Germany and the OECD and appreciations of the so-called "safe-haven" currencies – the yen, the Swiss franc, and the US dollar. In section 3, we characterize the dynamic cross-correlations between flights to safety and the financial and economic environment. We compute flight to safety betas for equity and bond portfolios and for commodity futures contracts, controlling for systematic exposures to the broad equity and bond markets. The financial, basic materials and industrial industries under-perform in FTS episodes, whereas the telecom industry outperforms. Large cap stocks outperform small cap stocks. For the bond market, we find that both money market instruments and corporate bonds face abnormal negative returns during FTS episodes. Most commodity prices decrease sharply during FTS episodes, whereas precious metal and gold prices measured in dollars increase slightly. Turning to the macro-economy, both economic growth and inflation decline right after and up to a year following a FTS spell. As an application of our methodology, we examine in section 4 whether hedge funds "hedge" FTS-events, with the disappointing finding that nearly all hedge fund styles, the event-driven ones in particular, have negative FTS betas.

There are, of course, a number of empirical papers that bear some indirect relation to what we attempt to accomplish. Baele et al. (2010) show that a dynamic factor model with standard fundamental factors fails to provide a satisfactory fit for stock and bond return comovements. The ability of the model to capture episodes of negative stock-bond return correlations only improves when stock-bond illiquidity factors (potentially capturing "flight-to-liquidity") and the VIX (potentially capturing "flight-to-safety") are included. Connolly et al. (2005) and Bansal et al. (2010) show that higher stock market uncertainty is associated with lower correlations between stock and bond returns and higher bond returns. Goyenko and Sarkissian (2012) define a flight to liquidity and/or quality using illiquidity in short-term (non-benchmark) US Treasuries and show that it affects future stock returns around the globe. Baur and Lucey (2009) define a flight to quality as a period in which stock and bond returns decrease in a falling stock market and differentiate it from contagion, where asset markets move in the same direction. They define the 1997 Asian crisis and the 1998 Russian crisis as flight to safety episodes. The recent financial crisis also sparked a literature on indicators of financial instability and systemic risk which are indirectly related to our flight to safety indicator. The majority of those articles use data from the financial sector only (see e.g. Acharya et al. (2012); Adrian and Brunnermeier (2011); Allen et al. (2012); Brownlees and Engle (2011)), but Hollo et al. (2012) use a wider set of stress indicators and we revisit their methodology in Section 2.2.2.

2 Identifying Flight-to-Safety Episodes

2.1 Data and Overview

Our dataset consists of daily stock and 10-year government bond returns for 23 countries over the period January 1980 till January 2012. Our sample includes two countries from North-America (US, Canada), 18 European countries (Austria, Belgium, Czech Republic, Denmark, France, Finland, Germany, Greece, Ireland, Italy, Netherlands, Norway, Poland, Portugal, Spain, Sweden, Switzerland, UK), as well as Australia, Japan, and New-Zealand. We use Datastream International's total market indices to calculate daily total returns denominated in local currency, and their 10-year benchmark bond indices to calculate government bond returns. For countries in the euro zone, we use returns denominated in their original (pre-1999) currencies (rather than in synthetic euros), with German government bonds serving as the benchmark. For the other European countries, local government bonds serve as benchmark bonds. More details as well as the summary statistics can be found in the online Appendix.

2.2 Measures of Flights to Safety

Our goal is to use only these daily bond and stock return data to identify a flight-to-safety episode. That is, ultimately we seek to create a $\{0, 1\}$-*FTS dummy variable* that identifies whether on a particular day a FTS took place. Given the theoretical literature, the symptoms of a flight to safety are rather easy to describe: market stress (high equity and perhaps bond return volatility), simultaneous high bond and low equity returns, and a low (negative) correlation between bond and equity returns. We use 4 different methodologies to calculate *FTS measures* or *probabilities*, numbers in the interval $[0, 1]$ that reflect the likelihood of a FTS occurring that day. The first methodology directly turns the incidence of (a subset of) the symptoms into a $\{0, 1\}$-FTS dummy, with 1 indicating an FTS episode and 0 otherwise. The second methodology delivers a continuous signal in $[0, 1]$ that is converted into a FTS probability. The last two use regime switching models to identify the probability of a flight to safety. These probabilities can be converted into FTS dummies using certain classification rules. In the following sub-sections, we detail these various approaches in calculating FTS probabilities, whereas section 2.3 discusses how to aggregate the 4 different FTS probabilities into one aggregate FTS dummy.

2.2.1 A Flight-to-Safety Threshold Model

Our simplest measure identifies a flight-to-safety event as a day with both an (extreme) negative stock return and an (extreme) positive bond return. The flight-to-safety measure FTS for country i at time t is calculated as:

$$FTS_{i,t} = I\left\{r^b_{i,t} > z_{i,b}\right\} \times I\left\{r^s_{i,t} < z_{i,s}\right\} \tag{2.1}$$

where I is the indicator function and $r^b_{i,t}$ and $r^s_{i,t}$ the time-t returns for country i in its benchmark government bond and equity market, respectively. We allow for country-specific thresholds $z_{i,b}$ and $z_{i,s}$. Because flights-to-safety are typically associated with large drops (increases) in equity (bond) prices, we use thresholds to model $z_{i,b}$ and $z_{i,s}$:

$$z_{i,b} = \kappa \times \sigma_{i,b,t} \qquad z_{i,s} = -\kappa \times \sigma_{i,s,t} \tag{2.2}$$

where $\sigma_{i,b,t}$ and $\sigma_{i,s,t}$ are the country-specific, time-varying volatilities for bond and stock returns at time t, respectively, and κ is the threshold parameter. Consequently, equity (bond) returns must be κ standard deviations below (above) zero before we identify a day to be a FTS day.

We allow for low-frequency changes in equity and bond market volatilities and covariances, and model them using a simple kernel method. Given any date t_0 in a sample $t = 1, .., T$, the kernel method calculates stock and bond return variances at the normalized date $\tau = t_0/T \in (0, 1)$ as:

$$\sigma_{i,\tau}^2 = \sum_{t=1}^{T} K_h\left(t/T - \tau\right) r_{i,t}^2, \qquad i = s, b$$

where $K_h(z) = K(z/h)/h$ is the kernel with bandwidth $h > 0$. The kernel determines how the different observations are weighted. We use a two-sided Gaussian kernel with bandwidths of 250 days (expressed as a fraction of the total sample size T):

$$K(z) = \frac{1}{\sqrt{2\pi}} exp\left(\frac{z^2}{2}\right)$$

Thus, the bandwidth can be viewed as the standard deviation of the distribution, and determines how much weight is given to returns either in the distant past or future. For instance, for a bandwidth of 250 days, it takes ±320 days to cover 90% of the probability mass[1]. We use a two-sided symmetric kernel rather than a one-sided and/or non-symmetric kernel because, in general, the bias from two-sided symmetric kernels is lower than for one-sided filters (see e.g. Ang and Kristensen (2012)). The time-varying covariances are calculated similarly.

Once the matrix of volatilities and covariances is determined for each country, the incidence of FTS under the threshold model depends on the magnitude of the threshold parameter κ, with the number of FTS days decreasing rapidly from about $\frac{1}{4}$ of the sample for $\kappa = 0$ to generally less than 0.5% for $\kappa = 2$. To benchmark these numbers we conduct a simulation experiment. Imagine that bond and stock returns are normally distributed with their standard deviations and correlations equal to the estimates described above; the means are set at their full-sample means. That is, we draw from a bivariate normal distribution with country- and time-specific second moments. In such a world, we would expect flights to safety to be rarer than in the real world with fat tails and negative skewness.

In Figure 1, we plot the percentage of FTS events both in the actual data and in the simulated data, as a function of κ. This percentage of FTS events is first computed for each country and then averaged across countries. For a low κ, the simulated data from the bivariate normal distribution generates more FTS events than the actual data. The curves intersect for $\kappa = 1.398$. For our subsequent work we select κ to be 1.5. At this threshold parameter, 0.79% of all days are FTS days, compared to 0.72% in the bivariate normal world. To get a sense of what happens

[1] To ensure that the weights sum to one in a finite sample, we divide by their sum.

on such extreme days, we also compute the average difference between bond and equity returns on flight to safety days. This return impact, averaged over the various countries, is also graphed in Figure 1 (see right hand side for the units). It increases from 1.20% for $\kappa = 0$ to 3.12% for $\kappa = 1$ to more than 5.6% for $\kappa = 2$. For $\kappa = 1.5$, the return impact is 4.26%.

2.2.2 Ordinal FTS Index

While the threshold methodology only uses stock and bond returns, our second methodology employs 6 variables that are correlated with a FTS occurrence either positively (+) or negatively (-) and for which we can define natural boundary values beyond which they can be viewed as exhibiting "mild FTS-symptoms":

1. The difference between the bond and stock returns (+; FTS symptom if ≥ 0)
2. The difference between the bond and stock returns, relative to its long-term moving average (+; FTS symptom if ≥ 0)
3. The short-term stock-bond return correlation (-; FTS symptom if ≤ 0)
4. The difference between the short and long-term stock-bond return correlation (-; FTS symptom if ≤ 0)
5. The short-term equity return volatility (+; FTS symptom if more than one standard deviation above its unconditional value, that is, larger than double the unconditional standard deviation)
6. The difference between the short and long-term equity return volatility (+; FTS symptom if ≥ 0)

Most of these variables are self explanatory. We measure short- and long-term variables using the same kernel method as described in Section 2.2.1 with a bandwidth of 5 and 250 days, respectively. For a bandwidth of 5 days, about 90% of the probability mass is allocated to observations ± 6 days away from the current observation. The long-term estimates are designed to capture low-frequency variations in asset returns and comovement due to changes in the macro-economic environment.

To come up with a FTS dummy based on this information, we proceed in two steps. First, we create a composite "ordinal" index defined on the [0,1] interval, based on the 6 variables. Second, we transform this index into a FTS probability, our actual FTS measure, incorporating information about whether the "weak" symptoms are satisfied or not.

To create the composite index, we combine observations on the 6 FTS-sensitive variables using the "ordinal" approach developed in Hollo et al. (2012), who propose a composite measure of systemic stress in the financial system. As a first step, we rank the observations in ascending (descending) order according to the variables that increase (decrease) with the likelihood of a FTS, such as the difference between bond and stock returns, both in itself and relative to its 250-day moving average, short-term equity market volatility, and the difference between short and long-term equity market volatility (the short-term stock-bond correlation, the difference between the short- and long-term stock bond correlations). Next, for each of the 6 variables, we replace each observation by its ranking number normalized by the total number of observations, so that values close to one (zero) are associated with a larger (lower) likelihood of FTS. For instance, a value of 0.95 at time t_0 for, say, short-term equity return volatility means that only 5 percent of observations over the full sample have a short-term equity volatility that is larger or equal than the time t_0 value. Finally, for each point in time, we take the average of the normalized ordinal numbers across the 6 FTS variables[2].

The ordinal approach yields numbers for each variable that can be interpreted as a cumulative density function probability, but it does not tell us necessarily the probability of a flight to safety. For example, numbers very close to 1 such as 0.99 and 0.98 strongly suggest the occurrence of a FTS, but whether a number of say 0.80 represents a FTS or not is not immediately clear. Despite the imperfect correlation between the different variables, the maximum ordinal numbers for the composite index are quite close to 1 for all 23 countries, varying between 0.9775 and 0.9996. To transform these ordinal numbers into a FTS ordinal probability, we first collect the ordinal numbers of the days that satisfy all the "mild" FTS symptoms defined above. We view the minimum of this set of ordinal index values as a threshold. All observations with an ordinal number below this threshold get a FTS ordinal measure value equal to zero. It would appear unlikely that such days can be characterized as flights to safety. For observations with an ordinal number above the threshold, we set the FTS ordinal measure equal to one minus the percentage of "false positives", calculated as the percentage of observations with an ordinal number above the observed ordinal number that do not match our FTS criteria. The number of false positives will be substantial for observations with relatively low ordinal numbers (but still above the minimum threshold) but close to

[2]We also considered taking into account the correlation between the various variables as suggested by Hollo et al. (2012), where higher time series correlations between the stress-sensitive variables increase the stress indicator's value. However, our inference regarding FTS episodes was not materially affected by this change.

zero for observations with ordinal numbers close to 1.

The left panel of Figure 2 plots the original FTS ordinal index values and corresponding threshold levels for the US, Germany, and the UK; the right panel shows the derived FTS ordinal measures. We view this measure as an estimate of the probability that a particular day was a FTS; a standard classification rule therefore suggests a FTS event when that probability is larger than 0.5. Values with a probability larger than 50% are depicted in black, values below 50% in light gray. The percentage of days that have an ordinal index value above the threshold ranges from 6% of the total sample for Germany to 9% for the UK. Of those observations, about 65% have a FTS probability larger than 50% in the UK, compared to about 75% in the US. In Germany, this proportion even exceeds 98%.

We further characterize FTS incidence with the ordinal measure in Table 1. The threshold levels show a tight range across countries with a minimum of 0.65 and a maximum of 0.80. The mean is 0.72. The percentage of sample observations above the threshold equals 10.5% with an interquartile range of 9.3%-11.4%. The raw ordinal index values seem to display consistent behavior across countries. Our measure is also influenced by the number of false positives above the threshold value. Therefore, the third column shows the percentage of observations above the threshold that have a FTS ordinal probability larger than 50%. The mean is 52.9% and the interquartile range is 39.1%-64.9%. Germany proved to be an outlier with a detection rate of 98.7% while the minimum value of 18.59% is observed for the Czech Republic. The final column assesses how rare FTS episodes according to this measure are. The percentage of observations with a FTS ordinal probability larger than 50% as a percentage of the total sample is 5.2% on average, with an interquartile range of 4.6%-6.3%.[3] The range is quite tight across countries (the minimum is 2.7%, the maximum is 7.9%).

2.2.3 A Univariate Regime-Switching FTS Model

Define $y_{i,t} = r^b_{i,t} - r^s_{i,t}$, with $r^s_{i,t}$ the stock return for country i and $r^b_{i,t}$ the return on the benchmark government bond for that country. We model $y_{i,t}$ as a three-state regime-switching (RS) model. Two regimes are necessary to capture the low and high volatility regimes that are typically identified in RS models for equity returns (see Ang and Bekaert (2002a) and Perez-Quiros and Timmermann (2001)). The third regime then functions as the FTS regime. The regime variable follows a Markov

[3]Note that for the US, Germany, and UK, the number in the last column is of course equal to the product of the numbers in columns 2 and 3, but this is no longer true once we show cross-country statistics.

Chain with constant transition probabilities. Let the time-t regime be indexed by v.

$$y_{i,t} = \mu_{i,v} + \sigma_{i,v}\epsilon_{i,t} \tag{2.3}$$

with $\epsilon_{i,t} \sim N(0,1)$. The means and volatilities can take on 3 values. Of course, in a FTS, $y_{i,t}$ should be high. To identify regime 3 as the flight-to-safety regime, we therefore restrict its mean to be positive and higher than the means in the other two regimes, i.e. $\mu_{i,3} > 0, \mu_{i,3} > \mu_{i,1}, \mu_{i,3} > \mu_{i,2}$. The transition probability matrix, Φ_i, is 3×3, where each probability p^i_{kj} represents $P\left[S_{i,t} = k \middle| S_{i,t-1} = j\right]$, with $k, j \in \{1,2,3\}$:

$$\Phi_i = \begin{pmatrix} p^i_{11} & p^i_{21} & (1 - p^i_{11} - p^i_{21}) \\ p^i_{12} & p^i_{22} & (1 - p^i_{12} - p^i_{22}) \\ (1 - p^i_{23} - p^i_{33}) & p^i_{23} & p^i_{33} \end{pmatrix} \tag{2.4}$$

Panel A of Table 2 reports the estimation results. The first column reports detailed estimation results for the US, followed by the average estimate and interquartile range across all 23 countries. Regime 1 is characterized by low volatility, and a significantly negative bond-stock return difference for all countries. This is in line with the expectation that equities outperform bonds in tranquil times. Regime 2 corresponds to the intermediate volatility regime, and also features a mostly negative bond-stock return difference, yet typically of a smaller magnitude than in regime 1 and often not statistically significant. Annualized volatility is about twice as high in regime 2 than in regime 1 (19.5% versus 9.7%).

The volatility in regime 3, the FTS regime, is on average more than 47%, which is more than 2.35 (4.5) times higher than in regime 2 (1). Looking at the interquartile range, the bottom volatility quartile of the FTS regime is nearly twice as high as the top volatility quartile of regime 2. The bond-stock return difference in the FTS regime is about 0.25% on average, significantly different from zero at the 5% (10%) level in 11 (16) of the 23 countries with an interquartile range of [0.198%; 0.271%]. While this is a relatively small number, the effect is substantially higher on days that the FTS regime transitions to the "on" state (1.09% on average, with an interquartile range of 0.73%-1.33%).

To identify the regimes and to characterize their persistence, we use the smoothed regime probabilities, which simply represent regime probabilities conditional on information from the full sample (see Kim (1994); Hamilton (1994)). In the model, the agents are assumed to observe the regime while the econometrician does not. The smoothed probabilities reflect the best estimate an econometrician can make regarding the probability of a particular regime at a particular point in time using full sample information. Good regime classification in a two-regime model would require smoothed probabilities close to one or zero (see e.g. Ang and Bekaert (2002b)). The FTS regime is the least persistent regime (with an average probability of staying in the FTS regime of 94.7% versus 98.1% for regime 1 and 96.7% for regime 2). To classify a day as a FTS-event, we require the smoothed probability of the FTS regime to be larger than 0.5, even though there are three regimes[4]. The average FTS spell lasts 26.4 days. The large interquartile range (35.2 versus 17.2 days) reflects the substantial cross-sectional dispersion in the average FTS regime durations across countries. There is an average of 26 FTS spells in the sample. This number is somewhat hard to interpret as the sample period varies between 23 years and less than 13 years across different countries. Yet, most of the spells occur in the second half of the sample, and it is therefore useful to compare this number across different models.

2.2.4 A Bivariate Regime-Switching FTS Model

The univariate RS FTS model uses minimal information to identify FTS episodes, namely days of relatively high differences between bond and stock returns. It therefore can be viewed as a regime switching model counterpart to the threshold model. In the bivariate RS model, we try to incorporate more identifying information as in the ordinal index but using an RS model for bond and stock returns. Essentially, we attempt to build in the "FTS symptoms" of positive (negative) bond (stock) returns, a negative correlation between the two returns, and the presence of market stress, especially in the equity market. We estimate the following bivariate model for stock and bond returns in each country (we remove the country subscript i for ease of notation):

$$
\begin{aligned}
r_{s,t} &= \alpha_0 + \alpha_1 J_{s,t}^{lh} + \alpha_2 J_{s,t}^{hl} + \alpha_3 \left(J_t^{FTS} + v S_t^{FTS}\right) + \varepsilon_{s,t}, && (2.5)\\
&\quad \varepsilon_{s,t} \sim N\left(0, h_s\left(S_t^s\right)\right) && (2.6)\\
r_{b,t} &= \beta_0 + \beta_1 J_{b,t}^{lh} + \beta_2 J_{b,t}^{hl} + \beta_3 \left(J_t^{FTS} + v S_t^{FTS}\right) + \\
&\quad \left(\beta_4 + \beta_5 S_t^{FTS}\right) r_{s,t} + \varepsilon_{b,t},\ \varepsilon_{b,t} \sim N\left(0, \theta_{t-1} h_b\left(S_t^b\right)\right) && (2.7)
\end{aligned}
$$

[4]The percentage of FTS days would increase on average with about 1 percent of daily observations if we were to use 1/3 rather than 1/2 as a classification rule. Testing whether a third regime is necessary is complicated because of the presence of nuisance parameters under the null (see e.g. Davies (1987)), and is therefore omitted.

The variance of the stock return shock follows a two-state regime-switching model with latent regime variable S_t^s. The variance of the bond return shock has two components, one due to a spillover from the equity market, and a bond-specific part. The latter follows a two-state regime-switching square-root model with latent regime variable S_t^b; θ_{t-1} is the lagged bond yield.[5] The "jump" terms $J_{s,t}^{lh}$ and $J_{s,t}^{hl}$ are equal to 1 when the equity return shock variance switches regimes (from low to high or high to low), and zero otherwise. We expect α_1 to be negative and α_2 to be positive. $J_{b,t}^{lh}$ and $J_{b,t}^{hl}$ are defined in a similar way but depend on the bond return shock variance. Without the jump terms, regime switching models such as the one described above often identify negative means in the high volatility regime. However, we would expect that there is a negative return when the regime jumps from low to high volatility but that the higher volatility regime features expected returns higher not lower than the low volatility regime. The jump terms have this implication with $\alpha_1 < 0$ and $\alpha_2 > 0$. There is a mostly unexpected negative (positive) return when the regime switches from the low (high) volatility to the high (low) volatility regime. Within the high volatility regime, there is some expectation that a positive jump will occur driving the mean higher than in the low volatility regime where there is a chance of a jump to a high volatility regime. This intuition was first explored and analyzed in Mayfield (2004).

The structure so far describes a fairly standard regime switching model for bond and stock returns, but would not allow us to identify flights to safety. Our identification for the flight to safety regime uses information on the means of bonds versus equities, on equity return volatility and on the correlation between bond and stock returns. Let S_t^{FTS} be a latent regime variable that equals 1 on FTS days and zero otherwise. We impose $\alpha_3 < 0$ (stock markets drop during FTS episodes), $\beta_3 > 0$ (bond prices increase during FTS), and $\beta_5 < 0$ (the covariance between stocks and bonds decreases during FTS episodes). It is conceivable that a flight to safety lasts a while, but it is unlikely that the returns will continue to be as extreme as on the first day. Therefore we introduce the J_t^{FTS} variable, which is 1 on the first day of a FTS-regime and zero otherwise, and a scaling parameter v, imposed to be positive. On the first day, the total effect is maximal at $(1+v)\,\alpha_3$ and $(1+v)\beta_3$, while on subsequent FTS days the negative (positive) flight-to-safety effect on equity (bond) returns is allowed to decline to $v\alpha_3$ ($v\beta_3$). We expect but do not impose v to be substantially below 1.

We assume S_t^b and S_t^{FTS} to be independent Markov chain processes. S_t^s is as-

[5] By making the bond return shock variance a function of the (lagged) interest rate level, we avoid the result that the high volatility regime is only observed in the first years of sample, as the early 1980s was a period of high interest rates.

sumed to be independent of S_t^b, but we assume that the equity volatility regime is always in the high volatility state, given that we experience a FTS episode:

$$\Pr\left(S_t^s = 1 |\, S_{t-1}^s, S_t^{FTS} = 1\right) = 1 \tag{2.8}$$

Panel B of Table 2 summarizes the estimation results. The jump terms have the expected signs for the equity market (and are mostly significant) but for bond returns, the results are more mixed. We clearly identify a high and low volatility regime for both the bond and the stock market, with volatilities typically about twice as high in the high volatility regime. In terms of the parameters governing the FTS regime, we find that α_3 is -7.863% in the US, and -5.22% on average, with a substantial interquartile range ([-7.42%, -1.63%]). Not surprisingly, the υ-scaling parameter is mostly rather small (interquartile range of [0.014,0.047]), indicating that a FTS mostly only induces one day of heavy losses[6]. For bond returns, β_3 is 0.81% on average, but it is also often drawn to the lower boundary of zero. Finally, we do find that β_5 is statistically significantly negative, indicating that a FTS induces a negative covariance between bond and stock returns (or at least one lower than the covariance in non-FTS regimes) above and beyond what is induced by the jump terms. As reflected by the average and interquartile values for β_4, the average stock-bond correlation in 'normal' times is relatively close to zero in our sample, but positive on average.

To identify a FTS day, we use the standard classification rule that the smoothed FTS regime probability be larger than 0.5. We do find that the bivariate model predicts FTS spells to last substantially longer than in the univariate model, with an average of 89.9 days in the US and 86.8 days on average in all countries and a substantial interquartile range of [58-101] days. On average, the number of FTS spells from the bivariate RS model is lower than that from the univariate RS model, although it is higher for the US (24 compared to 18).

2.3 Aggregate FTS Incidence

At this point, we have transformed data on bond and stock returns and simple information about the "symptoms" of a FTS into 4 noisy measures regarding the presence of a FTS day. These measure are between 0 and 1 and can be interpreted as a measure of the probability of observing a FTS event. The literature on classification suggests that the optimal rule (in the sense that it minimizes misclassification) is to

[6] The average value for ν (2.079) is higher than the value for the top quartile (0.047) because a small number of countries have a very high value of ν.

classify the population based on the relative probability. Given that there are two states (FTS or not), a probability of a flight to safety higher than 0.5 would lead to the conclusion that there is a flight to safety. We use this rule to create 4 alternative FTS dummies. Table 3 (right hand side columns) reports the average number of days classified as FTS according to each of the 4 approaches. In general, the threshold and ordinal method yield a relatively low incidence of FTS days, whereas the regime-switching approaches deliver relatively persistent FTS regimes and classify more days as FTS events. For most countries, the proportion of time spent in a FTS-episode increases monotonically moving from the threshold indicator (0.84% on average) to the ordinal indicator (4%), then to the univariate RS model (9.76%) and finally the bivariate RS model (14.91%). Within each method, the interquartile ranges are relatively tight, ranging from 0.71%-0.93% for the threshold indicator to 2.5%-5.6% for the ordinal indicator to 7.9%-12% and 10.9%-18.9% for the univariate and bivariate RS models, respectively.

To aggregate the information in the 4 measures into one FTS measure, our main methodology relies on the extant literature on regime classification based on qualitative variables (see e.g. Gilbert (1968))[7]. We view the 4 methods as yielding a Bernouilli draw on the FTS with the probability estimated at each point of time. It recognizes that if 3 of the 4 variables indicate a flight to safety, we should be rather confident a flight to safety indeed occurred. We extract the joint probability that at least 3 out of our 4 indicators identify a FTS on a particular day from a multivariate Bernoulli distribution using the method proposed by Teugels (1990) (see Appendix A for technical details). This computation requires not only the probabilities of the 4 Bernoulli random variables at each point in time but also their covariances. It goes without saying that inference based on the 4 different measures is likely to be positively correlated. Sample correlations between the 4 FTS dummies vary roughly between 20% and 65%. In these day by day computations, we use full sample estimates of the covariances between the different FTS dummies (the underlying Bernoulli variables), which we estimate using the usual 50% classification rule as explained above. We then set the joint FTS dummy equal to one when that joint probability is larger than 50%, and zero otherwise.

Given this aggregation method, we record the proportion of time spent in a FTS episode in Table 3 (left column). The average proportion is 2.54% (interquartile range of 1.29%-3.55%). The incidence of FTS days using the aggregate measure is

[7] As an alternative, we also considered a naive aggregator which simply averages the probabilities at each point in time. When that average is above 0.5, we conclude there is a flight to safety, and set the average FTS dummy equal to 1. Both measures largely select the same periods as FTS episodes, and the dummy variables are highly correlated at 85.2%.

therefore somewhere in between the threshold and ordinal measures, but yields far fewer FTS days than the regime switching models do. Figure 3 shows the results for the US. The top Panel of Figure 3 plots the joint FTS probability. The bottom panel plots the corresponding FTS dummy, which equals one when the joint probability is larger than 50%, and zero otherwise. The FTS measure yields very few FTS instances before 1995 (the October 1987 crash, and a short period reflecting the 1990 recession in the US being the exceptions), but FTS are much more frequent in the second half of the sample.

2.4 Characterizing FTS Episodes

With the identification of FTS events in place, we now analyze several features of these events, including the return impact on FTS days, the persistence of FTS events, the relative contribution of the various methodologies, and the global or local nature of FTS events.

Return impact

Figure 4 summarizes average returns on equities and bonds as well as the return impact (bond minus equity return) before, during, and after FTS events. The horizontal axis records 7 points on a time line, namely 5 to 1 days before a FTS, 1 day before a FTS, the first day of a FTS, the rest of a FTS spell except for the last day, the last day of a FTS spell, and finally, 1 to 5 days after a spell. Returns on the vertical axis are in percent and the dashed lines connect the average return across countries. The vertical bars represent the interquartile ranges across countries.

Outside FTS spells, equity and bond returns are close to zero and so is the return impact with inter-quartile ranges that are very tight. Interestingly, one day before and one day after a FTS spell, equity returns are solidly positive and bond returns negative, leading to strong positive return impact just before and after a FTS spell. While this seems puzzling at first, it is entirely driven by FTS events identified by the ordinal method. As we show below, because the ordinal FTS measure's persistence is between that of the threshold and RS models, it has a strong influence on our joint FTS inference. Moreover, we find that this result is entirely due to days that happen "in between" FTS events; that is, on the day before a FTS that is not preceded by another FTS event in the previous 5 days, both the equity and bond return impacts are in fact close to zero. A plausible hypothesis is that these events reflect reversals during stressful times. In particular, crisis periods tend to be somewhat persistent (as identified by our RS models) but within such periods, flashes of good news can lead to short-term reversals with positive equity and negative bond returns. Anyone having watched financial markets during the 2008-2009 financial crises will remember such market behavior.[8]

During a FTS spell, we see very significantly negative (positive) equity (bond) returns, with the returns being slightly larger in absolute magnitude both in the beginning and towards the end of a spell. The return impact is 3.16% on the first day of a FTS spell with an interquartile range of 2.51-3.87%, 3.45% on the last day of a FTS spell with an interquartile range of 2.80-4.07%, and 2.49% on days in between with an interquartile range of 1.44-2.86%. On average, the return impact during the sample is 0.014%, so that the FTS spells generate first-day return impacts roughly 2.4 standard deviations above the average difference between equity and bond returns, where we used the ensemble standard deviation to do this computation.

The persistence of FTS

Figure 5 plots the cross-country average number of FTS spells that take exactly 1, 2, or 3 days, or whose length is in the intervals [4-9], [10-49], [50-99], or more than 99 days, based on each of the 4 individual measures as well as the joint measure. Spells are usually very short-lived under the threshold model and never longer than 3 days. Spells are often much longer under the RS models. This is not surprising, as the identification of regimes in regime switching models often relies heavily on second moments (volatilities, correlations) that tend to be highly persistent. The persistence of spells identified using the joint measure seems most similar to that using the ordinal method. If we express the distribution of the duration of FTS spells under the joint measure in fractions, we find that 55.7% of the spells last 1 day, 22.9% last two days, 10.3% last three days, 9.14% last between 4 and 9 days, and only 1.9% last longer than 10 days.

Contribution of different models

The different persistence across the various methodologies suggests that they may contribute differently to the joint measure. Given the nature of the aggregation methodology described in Section 2.3, quantifying the contribution of each method to the joint measure is non-trivial. An indirect way to do so is to set an individual FTS measure to 0, and then to recompute the aggregate measure as in Section 2.3.

[8]This behavior provides a challenge for dynamic models of stock and bond returns, as it is not captured well by RS models. The Bad Environment Good Environment model in Bekaert and Engstrom (2009) can potentially capture such behavior.

If the measure is absolutely essential for FTS identification, the resulting "restricted" FTS measure should not overlap much with our actual measure. When the individual measure is not that important and is not essential for FTS identification, the overlap should be very substantial. In Table 4, we report the percentage overlap between the "restricted" FTS measures and the joint FTS measure, which is the percentage of FTS days under the original FTS measure that are also classified as FTS based on the "restricted" measure. Setting any of the last three FTS measures to zero dramatically reduces the incidence of FTS, but setting the threshold indicator to zero has only a marginal effect. The threshold model seems to have an marginal effect for FTS identification especially in the US, UK and Germany. However, the threshold methodology still contributes to the joint measure, as the average overlap across countries is only 83% with a minimum of zero and an interquartile range of [76%-96%]. When any of the other three measures is set to zero, the incidence of FTS declines much more dramatically, with only around 20% of the FTS days based on the joint measure continuing to be classified as FTS days. This fraction does not differ dramatically across the three measures, suggesting they are all roughly equally important in the identification of FTS events.

Propagation and global crises

Figure 6 plots the percentage of countries experiencing a FTS at each point in time. The FTS dummies clearly select well known global crises as global FTS events, including the October 1987 crash, the 1997 Asian crisis, the Russian crisis and LTCM debacle in 1998, the Lehman Brothers collapse and several spells during the European sovereign debt crisis. Defining a global FTS as one where at least two thirds of our countries experience a FTS, there are a total of 39 global FTS-days. In Table 5, we report the proportion of FTS spells that are global in nature across countries. The cross-country average of local FTS spells that are global in nature amounts to 23.0%, with an interquartile range of 14.0%-22.9%. Large developed countries such as the US, the UK and Germany (reported separately) feature a relatively low proportion of global spells, suggesting they are more subject to idiosyncratic flights to safety. While the interquartile ranges are relatively tight, a number of small countries, such as Norway, the Czech Republic and Poland have a very high proportion of global FTS episodes.

Table 6 investigates how FTS events propagate across countries. In particular, do global FTS episodes tend to originate from some countries more than from others? To answer this question, we examine three statistics for each country. The first set of columns simply reports the (empirical) probability that country j is in a

FTS state given that country i is in a FTS, averaged over all countries $j \neq i$. This statistic therefore captures how prevalent FTS events are in other countries, given that they occur in country i. According to this statistic, the UK is the top originator of FTS events, with 44.3% of countries experiencing a FTS when the UK does. However, because we do not differentiate between idiosyncratic and global FTS's in this calculation, it is possible that this statistic simply measures the propensity of a country to be embroiled in a global FTS. Therefore, the second set of columns excludes global FTS events from the computation. It is remarkable that the ranking across countries does not change very much. Finally, to capture better the idea of FTS originating in one country and propagating to another country, the third set of columns focuses on the two days before global FTS events and computes the percentage of those days that are FTS days for each country, with the goal of identifying the countries from which those crises originated. Interestingly, the three classification schemes yield a very similar set of countries as the main FTS originators. We show in bold countries that have percentage statistics two standard errors above the country average, where the standard error is computed as the cross-sectional standard deviation divided by the square root of the number of countries (23). At most 9 countries appear in bold for any of the three statistics, and no less than 7 countries (the UK, Germany, Sweden, the US, France, Canada and Austria) appear near the top in all three columns. In contrast, Japan, New Zealand and the emerging European countries are clearly not the most important originator countries, which can also be seen from the regional averages reported at the bottom of the table. The regional analysis suggests that the North-American region is the strongest originator followed by Developed Europe.

2.5 Alternative FTS Measures

Our FTS measures require minimal data inputs and provide a daily reading of flight to safety episodes. Of course, there are other financial indicators that may allow identification of a flight to safety episode. We therefore investigate the comovement between our FTS dummies and four types of alternative stress indicators. The first set comprises implied volatility indices on major indices: the US S&P500 (VIX), generally viewed as a fear index, the UK FTSE100 (VFTS), the German DAX (VDAX), and the Japanese Nikkei 225 (VXJ). We regress daily changes in those indices on our FTS dummies. Second, we investigate a series of sentiment/confidence indicators. The sentiment variables include the Baker and Wurgler (2006) sentiment indicator (purged of business cycle fluctuations), the Michigan consumer sentiment index (which measure sentiment in the US), the Ifo Business Climate indicator (which measures sentiment in Germany), and the country-specific OECD consumer confidence indicators (seasonally-adjusted). Because these sentiment variables are only available on a monthly basis, we regress their monthly changes on the fraction of FTS days within the month (expressed in decimals). Next, we regress percentage changes in the value of three safe haven currency values (i.e. the Swiss Franc, the Japanese Yen, and the US Dollar) on the FTS dummy using daily data. Note that the currencies are expressed in domestic currency units per unit of the safe currency, so that positive changes indicate an appreciation of the safe-haven currency. For this exercise, we leave out the three safe currency countries. Finally, we regress the TED spread (both in levels and in changes) on our FTS dummies.[9] The TED spread is the difference between the three month LIBOR rate for a particular currency (country) and the corresponding three-month T-bill interest rate. While it directly reflects default risk in the banking sector, the TED spread is more generally viewed as an indicator of the perceived credit risk in the economy and tends to spike in times of crises.

Table 7 contains the results. We show slope parameter estimates for the US, Germany and the UK, as well as the average, standard deviation and top/bottom quartile parameter estimates across all 23 countries. The last column shows the number of countries for which the parameter estimates are significant, using White (1980) heteroskedasticity-consistent standard errors.

The VIX increases by 3.30% on average when the US experiences a FTS. The effect of FTS on the US VIX is significant at the 10 (5) percent level in 22 (19) of the 23 countries. When country-specific implied volatilities (VIX for US, Canada; VFTS for the UK; VDAX for the other European countries; VJX for Japan, Australia and New Zealand) are used, however, the FTS effect increases in magnitude and becomes significant at the 5 percent level in all countries.

There is clear evidence of a significant decline in consumer and business sentiment during FTS episodes. The Baker-Wurgler sentiment indicator and the Michigan consumer sentiment index decrease significantly when there is a FTS in the US. The Michigan index also reacts significantly to flight to safety instances in Germany and the UK, although these countries witness only a limited number of global flights to safety (see Table 5). There are another 7 countries whose FTS episodes have significant effects on the Michigan index, but only 5 on the Baker-Wurgler index. The Ifo business climate indicator declines significantly in times of FTS for all but two countries. This is somewhat surprising as this indicator measures the German

[9] We only have TED spread data for 19 countries (not for Ireland, Austria, the Czech Republic and Poland).

business climate. A FTS negatively affects OECD consumer confidence in 19 countries, as measured by the country-specific OECD indicator of consumer sentiment. Thus, the Ifo business climate and OECD leading indicators seem linked to FTS events across the globe.

There is also strong evidence of a flight to safe haven currencies in times of a FTS. On average, during a FTS day, the Swiss Franc appreciates by 0.40%, the Japanese Yen by 0.72%, and the US Dollar by 0.35%. The appreciation of the Yen is significant following a FTS in all 22 countries, compared to in 19 countries for both the Swiss Franc and US dollar.

For the US, the TED spread is on average 27 basis points higher in times of FTS, but the cross-country average is essentially zero. The TED spread increases (daily) by 3 basis points when the US experiences a FTS. Across all countries, during a FTS day, the TED spread increases on average by 2 basis points, but this effect is only significant for 5 out of 19 countries.

3 FTS and the Economic and Financial Environment

In this section, we examine the comovement of FTS spells with a large number of financial and economic variables. Our goal is to document comovements rather than to look for causality. Methodologically, the framework is therefore a simple linear regression at the country (i) level, as follows:

$$r_{t,i} = \alpha + \beta FTS_{t,i} + \gamma Ctrl_{t,i} + e_{t,i} \tag{3.1}$$

Here, $r_{t,i}$ is mostly a return on a financial asset, but may also represent a yield or macro-economic variable. When daily data are available, FTS is the FTS dummy; when the data are measured over a time interval such as a month, FTS represents the fraction of days over the interval that are identified as FTS. The variable $Ctrl$ is a vector of control variables that may differ across regressions. For equity portfolios, for instance, it includes global and local stock market returns. The standard errors in the regressions are mostly heteroskedasticity-consistent and adjusted for serial correlation when overlapping dates are used (as in the regressions with macroeconomic variables reported in Section 3.5). We also considered an alternative specification where we split up the FTS dummy into a "first day" dummy and a "remaining days" dummy. These regressions overwhelmingly indicate that the two FTS betas are not significantly different from one another, and if they are, there is no systematic evidence in favor of the effects being stronger on the first day or the rest of the FTS days. We therefore relegate these results to an unpublished appendix. Unless otherwise mentioned, the format of our tables is identical across different classes of variables: we show the estimates for the US, Germany, and the UK, as well as the average, the standard deviation, and the top/bottom quartile estimates across all 23 countries.

Before we begin, we want to provide one illustrative example of the importance of FTS. It is to be expected that bond and stock returns, the two major asset classes, are positively correlated as they both represent long duration assets. Over our sample period, which starts fairly late in 1980, this correlation is nonetheless negative for 19 out of 23 countries. It is conceivable that this negative correlation is mainly caused by the relatively high incidence of FTS in the last 30 years. If such a "FTS-heavy" era is not likely to occur again in the near future, investors may want to re-assess the computation of the bond-stock return correlation. To assess the importance of FTS events for this important statistic, we eliminated FTS events in each country from the sample and recomputed the stock-bond return correlation. The stock-bond return correlation is -2.1% on average in "normal" periods with an interquartile range of [-7.4%, 4.2%], and -9.12% overall (interquartile range of [-13.1%,-5.3%]). The average correlation in FTS periods is in fact -42.86% and the absolute difference between those correlations in normal and in FTS times is on average 41.7%, with a relative tight interquartile range ([33.2%, 54.8%]). Thus, FTS events indeed render the bond-equity return correlation substantially more negative.

3.1 FTS and Equity Portfolios

To assess the FTS "beta" of different equity portfolios, we regress their daily returns on the FTS dummy and two controls for "standard" systematic risks, the world market return and the local stock market return, both measured in local currency units. As a consequence, the FTS beta must be interpreted as the abnormal return earned during FTS episodes, controlling for normal beta risks. Importantly, it does not indicate which portfolios perform best or worst during FTS spells, as portfolios with positive (negative) FTS betas may have also high (low) market betas, making them perform overall relatively poorly (well) during a FTS spell. We also estimated a specification with interaction terms between the FTS dummy and the benchmark returns, but this specification often runs into multi-collinearity problems and the results are therefore omitted. We also considered a specification where in addition we control for the global bond market return, constructed as the GDP-weighted average of the government bond returns across all 23 countries. This specification does not meaningfully alter our FTS betas and we relegate the results to the Online appendix.

Table 8 reports the FTS betas for 10 local industry portfolios (using the Datastream industry classification), 5 local style portfolios from MSCI (large caps, mid caps, small caps, value and growth), and two additional style portfolios: a SMB portfolio (i.e. the return on the small cap portfolio minus the return on the large cap portfolio) and a HML portfolio (i.e. the return on the value portfolio minus the return on the growth portfolio).

Among the industry portfolios, financials, basic materials and industrials generally show significant underperformance during a FTS, controlling for their "normal" betas. The inter-quartile ranges are negative for these industries and the FTS betas are statistically significant in many countries. In contrast, the only robustly "defensive" industry is telecom, the price of which on average increases 30.4 basis points on FTS days, after controlling for its normal beta. Other industries show significant but country-specific results. For instance, the technology sector outperforms in the US, but underperforms significantly in the UK. Consumer goods, health care, and consumer services outperform in the US but perform similarly to the broad market in Germany. When we investigate the global and local betas of the various industries, which are reported in the online appendix, we observe a weak negative correlation between the factor exposures and the FTS betas for the US. In other words, the performance of the various industries during a FTS is even more diverse than their betas would suggest. However, this pattern does not generalize across countries.

Among the style portfolios, large cap portfolios have mostly positive FTS betas, whereas small cap portfolios have negative FTS betas. Value portfolios tend to have negative FTS betas and growth portfolios positive ones, but the betas are small and the results are statistically weaker than for the size portfolios. This is naturally confirmed when we look at spread portfolios, where the SMB portfolio has an average FTS beta of -47 basis points (significant in 15 out of 23 countries), but the HML portfolio only has a FTS beta of -14 basis points (significant in only 6 countries). Perhaps the size results can be interpreted as a flight to quality into larger, well-known companies or a flight to liquidity into widely traded firms.

3.2 FTS and Bond Portfolios

In Table 9, we focus on how FTS events affect the bond markets. Panel A reports how bond yields and spreads react during FTS episodes. Because interest rates are highly persistent and appear to be on a downward trend over the sample period, a regression of yields on an FTS dummy may just record the lower interest rates prevailing in the FTS-heavy later part of the sample. We therefore measure yields and spreads relative to their moving averages over the most recent 150 days. We construct the level, slope and curvature factors from 3-month T-bill rates and 5- and 10-year bond yields in the usual fashion (see the Table notes for details). The corporate bond indices are only available for the US, Japan, Canada, Australia and the Eurozone as a whole; we therefore use the Euro-zone corporate bond index for the European countries and the Australian corporate bond index for New Zealand.

On average, the nominal government bond yield curve shifts down, flattens and becomes less hump-shaped in times of FTS (our curvature factor is decreasing in the degree of curvature). Nominal government bond yields decline significantly in all but some southern European countries (e.g. Greece, Portugal and Italy), which see significant increases in their government bond yields. This is consistent with a FTS from those countries towards safer countries (like Germany and the US). Central banks seem to respond to FTS episodes, as the targeted interest rate declines considerably in most countries. Turning to corporate spreads, we see mixed results for the spreads between yields on AAA-rated corporate bond and those on 10-year government bonds: most developed countries (e.g. US, UK, Germany) observe a significant widening of those spreads, likely reflecting both higher credit risk premiums and higher liquidity premiums during a FTS. In contrast, certain non-core European countries (e.g. Belgium, Italy, Spain, Greece, Portugal) and New Zealand see those spreads narrowing, likely reflecting the fact that local investors prefer highly-rated regional corporate bonds above local government bonds in times of FTS. Finally, we find a significant increase in the BBB-AAA spread for all but 3 countries.

In unreported results, we also examine inflation-indexed government bond yields from seven countries for which such data is available: US, UK, Japan, Canada, Sweden, Australia, and France. For the majority of the countries, nominal government bond yields decline by much more than real yields do.[10] This indicates a decrease in inflation expectations or inflation risk premiums in such times (see Section 3.5 for a thorough discussion on the comovement between FTS episodes and the macroeconomy) in addition to a drop in the real yield. For Canada, however, the real yield curve on average shifts up while the nominal yield curve shifts down during a FTS episode, whereas for Japan the declines in real yields are larger than those in nominal yields, although neither decline is significant.

Panel B of Table 9 reports the FTS betas for daily returns on various bond

[10]When we compare the reaction of both nominal and real bond yields to FTS, we restrict the sample for the nominal bond yields to the (slightly) shorter period that real bond yields are available.

portfolios. We follow a similar procedure as for equity returns and control for the exposure to the long-term benchmark bond portfolio in each regression. For corporate bond returns, we also control for the local stock market return. The bond portfolios include JP Morgan Libor-based cash indices with maturities of 1, 2, 3, 6 and 12 months, benchmark Datastream government bond indices with maturities of 2, 5, 7, 10, 20 and 30 years, and Bank of America/Merrill-Lynch corporate bond indices for AAA, AA, A and BBB rating groups, which have somewhat limited country coverage (see above). All returns are daily and denominated in the local currency.

For the US and UK, there is a pronounced pattern that during FTS episodes, shorter-term bonds underperform the benchmark 10-year government bond, while the longer-term 30-year bond outperforms. This pattern largely remains when looking across all countries but becomes less pronounced. Corporate bonds underperform after controlling for their exposures to the stock market and the government bond market; the underperformance is more significant for lower-rated bonds, although the FTS betas of A- and BBB-rated bonds are quantitatively similar. The finding that AAA bonds slightly over-perform on average is driven entirely by Japan; when Japan is excluded, AAA bonds also underperform with a FTS beta of -0.042. It is interesting to note that the betas of corporate bonds with respect to the long-term government bonds are around 0.4 and slightly smaller for lower ratings, whereas the equity betas are minuscule. Hence, corporate bonds almost surely outperform equities during FTS-episodes.

Finally, in Panel C we consider two types of spread portfolios; first, term spread portfolios consisting of a long position in the 10-year government bond and a short position in either the 1-month cash index or the 2-year government bond, and, second, default spread portfolios consisting of a long position in the AAA corporate bond index (benchmark government bond) and a short position in the BBB corporate bond index (the AAA corporate bond index). The first type of portfolios would perform well when the yield curve shifts down and/or flattens, while the second type of portfolio would perform well when default risks or default risk premiums rise. We find that the term spread portfolios generally outperform during FTS events, consistent with the finding in Panel B that longer-term bonds outperform shorter-term instruments. Turning to the default spread portfolios, the government-AAA portfolio outperforms on FTS days for the US, consistent with fears of increased default risks on those days, but underperforms on average across countries. This average underperformance is largely driven by investor preferences for the regional high-quality corporate bonds over local government bonds in some non-core European countries and New Zealand as mentioned above. In contrast, the AAA-BBB spread portfolio consistently delivers positive abnormal returns on FTS days for all countries.

3.3 FTS and Liquidity

3.3.1 Bond Market Liquidity

Benchmark Treasury bonds are attractive in times of market stress not only for their low level of default risk, but also for their (perceived) high levels of liquidity. Longstaff (2004) shows that the liquidity premium in Treasury bonds can amount to more than 15 percent of their value. Beber et al. (2009) find that while investors value both the credit quality and liquidity of bonds, they care most about their liquidity in times of stock market stress. Of course, it is unclear whether the supply of liquidity in the Treasury bond market is present when it is most desired. It is also not likely present for all bonds. Chordia et al. (2005) find that the liquidity in the Treasury market overall deteriorates during crisis periods. Goyenko and Ukhov (2009) show that bid-ask spreads on Treasury bills and bonds increase during recessions, especially for off-the-run long-term bonds.

Our analysis of how bond (il)liquidity is correlated with FTS is severely hampered by data availability. We therefore only show results for the US. Our first illiquidity measure was proposed by Goyenko and Ukhov (2009) and used more recently in Baele et al. (2010) and Goyenko et al. (2011). It is the average of proportional quoted spreads[11] of off-the-run US Treasury bonds with a maturity of at most 1 year (in percent).[12] This measure is available at the monthly frequency from the start of our sample (1980) till December 2010. The monthly average spread is calculated for each security and then equal weighted across securities. Our daily FTS measures are transformed to monthly indicators by taking the proportion of FTS days within a month. Because the proportional spread is clearly non-stationary over our sample, decreasing from over 0.09% in the early 1980s to less than 0.01% more recently, our estimations use the spread relative to a 6-month moving average as the dependent variable (multiplied by 100). As Panel A of Table 10 shows, we observe a positive and significant increase in the proportional spread (relative to a 6-month moving average) on FTS days.

As a second measure, we use the off/on-the-run spread, calculated as the negative of the daily yield difference between an on-the-run Treasury bond and a synthetic off-

[11] The proportional spread is calculated as the difference between ask and bid prices scaled by the midpoint of the posted quote.

[12] We would like to thank Ruslan Goyenko for making this series available to us.

the-run Treasury security with the same coupon rate and maturity date.[13] On-the-run bonds tend to trade at a premium (lower yield) because investors appreciate their higher liquidity relative to off-the-run bonds (see e.g. Jordan and Jordan (1997), Krishnamurthy (2002), and Graveline and McBrady (2011)). Pasquariello and Vega (2009), among others, show that the off-on-the run spread increases in times of higher perceived uncertainty surrounding U.S. monetary policy and macroeconomic fundamentals. The second row of Panel A of Table 10 shows that the off-on-the-run spread increases from about 14 basis points in "normal" times to more than 24 basis points on FTS days (with the change significant at the 1% level).

As a third measure, we use the root mean squared distance between observed yields on Treasury bonds with maturities between 1 and 10 years and those implied by the smoothed zero coupon yield curve proposed by Gurkaynak et al. (2007). This cross-sectional "price deviation" measure was recently used by Hu et al. (2013), who argue that it primarily measures liquidity supply. When arbitrageurs have unrestricted risk-bearing capacity, they can supply ample liquidity and can quickly eliminate deviations between bond yields and their fundamental values as proxied by the fitted yield curve. When their risk-bearing capacity is impaired, liquidity is imperfect and substantial deviations can appear. Fontaine and Garcia (2012) propose a similar measure. Hu et al. (2013) show that their "noise measure" is small in normal times but increases substantially during market crises. The noise measure is on average only 3.6 basis points, but increases to over 10 basis points during crises. Yet, this measure also shows a long-term trend downwards from the early 80s till the end of the 90s. We therefore investigate its value relative to a 150-day moving average. The final row of Panel A shows that the noise measure increases on FTS days relative to its 150-day moving average by about 1.2 basis points (which is significant at the 1% level).

Our overall findings on bond liquidity are consistent with the detailed results in a recent paper by Engle et al. (2012), who use (high-frequency) order book data for on-the-run 2, 5, and 10 year notes from early 2006 till mid-2010. They analyze Treasury bond liquidity in stress times using a FTS threshold measure inspired by this paper to identify stress. They find trading volumes, the number of trades, and net buying volumes to be substantially higher on FTS days, especially for shorter-term (2-year) notes. However, they find market depth, a measure of the willingness to provide liquidity, to be much lower on FTS days and to thin out more quickly for the 5 and 10-year notes than for the 2 year notes. The combination of decreasing depth and

[13] See Section 6 in Gurkaynak et al. (2007) for a discussion on how to calculate the synthetic yields. Our measure is adjusted for auction cycle effects.

high price volatility on FTS days suggests that even though liquidity demand shoots up, high market volatility makes dealers substantially more conservative with their liquidity supply, as they attempt to reduce adverse execution risk. Hence, their paper concludes that insufficient liquidity supply causes bond market illiquidity in stress times.

3.3.2 Equity Market Liquidity

Brunnermeier and Pedersen (2009) develop a theory where a (severe) market shock interacts with (evaporating) funding and market liquidity, with liquidity provision being curtailed particularly in volatile assets such as equities. The extant empirical work seems to confirm this intuition. For example, Chordia et al. (2005) find that equity market liquidity deteriorates together with that in the Treasury market during crisis periods; Naes et al. (2011) find that equity market liquidity systematically decreases during (and even before) economic recessions.

Here, we link our FTS measures to three measures of equity market illiquidity, namely the effective tick measure developed in Goyenko et al. (2009) and Holden (2009), the price impact measure of Amihud (2002), and the reversal measure of Pastor and Stambaugh (2003). Goyenko et al. (2009) and Holden (2009) estimate the effective bid-ask spread from prices using a price clustering model. The "Effective Tick measure" is the probability-weighted average of potential effective spread sizes within a number of price-clustering regimes divided by the average price in the examined time interval. Amihud (2002) examines the average ratio of the daily absolute return to the dollar trading volume on that day, which measures the daily price impact of order flow. Pastor and Stambaugh (2003) use a complex regression procedure involving daily firm returns and signed dollar volume to measure (innovations in) price reversals, both at the firm and market levels. In the tradition of Roll (1984), price reversals are interpreted to reflect the bid-ask spread. Aggregate measures for each of these indicators are equally-weighted averages of monthly firm-level estimates that are in turn estimated using daily firm-level data within a month. Unreported time series graphs reveal that the Amihud and Pastor-Stambaugh series are stationary, so we report level regression results. However, the effective tick measure starts a downward trend at the end of the 80s-early 90s, rendering the series non-stationary. We therefore investigate the series relative to a 6-month moving average.

Results in Panel B of Table 10 suggest that illiquidity in the US equity market increases substantially and significantly during FTS. The FTS coefficients are very large relative to the means in normal periods, as reflected by the constants in the regressions. Do note though that the monthly nature of the data implies that the full estimated effect will never materialize, as this measures the effect of a month in which all days are FTS, which never happens. The maximum FTS value is 0.65, which is obtained for November 2008.

3.4 FTS and Commodities

In Table 11, we report regression coefficients from a regression of the daily S&P GSCI benchmark commodity index returns, which measure returns on commodity futures contracts worldwide, on the joint FTS dummy while controlling for global equity market exposure. We consider broad indices (Commodity Total, Energy, Industrial Metals, Precious Metals, Agriculture, Livestock) and subindices (Crude Oil, Brent Crude Oil and Gold). The table has the exact same structure as the previous tables for bonds and equities, except for the last but one column, which reports the average exposure (beta) to global equity market returns. We note that commodity prices generally decline on FTS days, ranging from on average minus 6 basis points for Livestock to minus 52 basis points for Crude Oil, with Agriculture being an exception registering an average increase of 18 basis points. The decreases are statistically significant for the great majority of country/commodity pairs, including for Agriculture for many countries. There is one, not entirely surprising, exception: precious metals and its main component, gold. Both have positive FTS betas of on average 12 and 13 basis points, respectively. In both cases, the interquartile ranges are strictly positive, and the FTS betas are significant in 11 and 13 of the 23 countries, respectively. Note that Precious Metals and Gold have non-trivial positive global market betas, therefore the positive FTS exposures may be partially offset by negative market returns during a FTS spell. In fact, when we do not control for equity market exposure,[14] the FTS beta for Precious Metals become negative (minus 3 basis points on average) and for gold it drops to on average 6 basis points; both are statistically significant in only 1 country. In comparison, all other commodities have positive market exposures and therefore their systematic risk exacerbates the negative effect of their FTS exposure.

3.5 FTS Episodes and the Macroeconomy

In Table 12, we investigate the comovement between FTS episodes and the macroeconomy by regressing a number of macroeconomic variables on the fraction of days of FTS instances within the month (expressed in decimals). We investigate 3-month

[14] These results are available in an online appendix.

changes in the CPI, industrial production (IP), the unemployment rate, and the OECD leading indicator over the next three months, as well as quarterly changes in real GDP and the investment-to-GDP ratio over the following quarter. For inflation, IP growth, GDP growth, the unemployment rate and investment growth, we also have monthly survey forecasts from Consensus Economics and we examine both the mean and the standard deviation of 12-month ahead individual forecasts 3 month hence.[15] In the lines with variables marked "future", we regress annual growth rates or annual changes in the economic variables over the following year on the fraction of days of FTS instances within the month (expressed in decimals).

Inflation, real GDP growth, and IP growth are significantly lower during FTS episodes for most countries with data available. The average growth rate and the interquartile range across countries are both strictly negative. Unemployment increases significantly in more than half of the countries. Investment as a percentage of GDP also declines in most countries, but only significantly so in 6 countries. During FTS episodes, survey participants on average predict significantly lower real growth and inflation and significantly higher unemployment rates in most of the countries with forecasts data available. Forecast uncertainty, as measured by the cross-sectional standard deviation of individual forecasts, increases significantly for roughly half of the countries.

Inflation, real GDP growth, and IP growth also decline significantly one year after the FTS for most countries, and unemployment increases substantially over the same period. Note that the economic magnitudes of those changes are very large. For example, US real GDP growth is predicted to be 4.4% lower if all days within a month are categorized as a FTS, although the observed effect will be smaller as the percentage of FTS days within a month never exceeds 65%[16]. Finally, a FTS spell is accompanied by a contemporaneous decline in the OECD leading indicator but an increase one year in the future. As the OECD aims to predict the business cycle with a 6 to 9 months lead, this suggests that the economy is expected to rebound within two years. However, while significant in the US and Germany, this pattern holds for only one quarter of the countries.

[15] 12-month constant-horizon forecasts are calculated by interpolating between current- and next-year forecasts. RPI inflation forecasts are used for the U.K. as CPI inflation forecasts only became available in Jan 2004.

[16] The large magnitudes of the mean estimates for future IP growth and future GDP growth are primarily driven by Japan and Norway, both of which feature very low FTS incidences. Excluding those two countries reduces the mean estimates from -13.758 and -14.870 to -5.350 and -8.723, respectively.

4 An Application: Do Hedge Funds Hedge Against FTS?

Hedge funds should in theory provide at least a partial protection against market downturns. Those funds can go long and short and can invest in a wide array of securities and derivative products, which could potentially provide positive returns in all market environments. In fact, the name "hedge funds" suggest that they may "hedge", presumably, and likely in the mind of many investors, against bad times. But do they? In this section, we examine how well hedge funds do in times of market stress, as measured by our FTS dummy.

We use monthly returns (in US Dollars) on the Dow Jones Credit Suisse Hedge Fund indices over the period January 1994 - December 2010. The data comprise an overall index and 13 different hedge fund categories ranging from Convertible Arbitrage to Global Macro. All returns are in excess of the US 3-month Treasury bill yield. Our FTS variable is the monthly FTS incidence, as defined before.

We run time series regressions of the form:

$$r_{i,t} = \alpha_i + \beta_i^{FTS} FTS_t + \beta_i^{'} F_t + \varepsilon_{i,t} \tag{4.1}$$

where FTS_t is the FTS incidence variable and F_t are risk factors. In a first specification, we only use the US equity market and its lag as risk factors, the latter to control for illiquid positions that can only be slowly unwound. A second specification uses the well-known Fung and Hsieh (2004) factors.[17]

The results are reported in Table 13. The middle columns report the coefficients in the regressions with the current and lagged market factors, and the last column reports the FTS beta in the regressions controlling for the Fung and Hsieh (2004) factors (more detailed results are available upon request). A first striking result is that, except for Dedicated Short Bias and Managed Futures, all categories have positive and significant market betas, even the market neutral categories. Lagged betas are often significant as well. This result is not new and has been pointed out by Asness et al. (2001) for hedge funds in general and Patton (2009) for market

[17] The seven factors included in the Fung and Hsieh (2004) model are the excess return on the S&P 500 index; a small minus big factor; the excess returns on portfolios of lookback straddle options on currencies, commodities, and bonds; the yield spread of the US ten-year Treasury bond over the three-month T-bill, adjusted for the duration of the ten-year bond; and the change in the credit spread of the Moody's BAA bond over the ten-year Treasury bond, adjusted for duration. Fung and Hsieh (2004) and Fung et al. (2008) have shown that these factors have considerable explanatory power for fund of funds and hedge fund returns. Our results are robust to including two additional trend chasing factors (short term interest rates; stock index) and emerging markets as additional risk factors.

neutral funds. Note that this means that most hedge fund categories (with the natural exception of dedicated short bias) are bound to perform poorly in negative market return environments. Our FTS beta measures the differential performance after controlling for market risk during FTS events. We find that FTS betas are negative for all but one category (Managed Futures is the exception). They are statistically significant for the overall index, Event Driven, Event Driven Distressed, Event Driven Multi-Strategy, and Fixed Income Arbitrage. Rather strikingly, even the Dedicated Short Bias category has a negative (but insignificant) FTS beta of nearly 4 percent. Note that the effects are economically large, ranging between 1% and 7%. The Managed Futures category seems to be the one hedge fund category that has no systematic exposure to either market returns or FTS events. In the last column of the table, we report the FTS betas from our second specification with the Fung-Hsieh risk model. The results are largely unchanged.

The negative exposure to FTS events is very robust and not easily explained. For example, it is not simply hedge funds not responding well to high volatility environments. In unreported work, we replace our FTS dummy by a high volatility regime dummy (drawn from our bivariate regime switching model), but find the effects to become much weaker both in economic and statistical terms. Boyson et al. (2010) suggest that hedge funds may experience contagion effects in response to large adverse shocks to asset and hedge fund liquidity, whereas Sadka (2012) has shown that liquidity risk is an important factor in the cross-section of hedge fund returns. Given that liquidity tends to dry up during FTS spells, the negative FTS betas may reflect a liquidity effect. However, when we add the liquidity factors examined in Section 3.3 to the regression, we again find that our evidence regarding FTS betas remains unchanged, even though some hedge fund categories display significant liquidity betas.

We conclude that hedge funds, with the exception of the Managed Futures category, do not hedge against FTS events. We do point out that we find the alphas of all categories, except Dedicated Short Bias and Emerging Markets, to be positive and statistically significant, which remains true under the Fung-Hsieh model. Thus, relative to our risk model, hedge funds on average generate alphas of 30 to 100 basis points (per month). Titman and Tiu (2011) find that the best-performing hedge funds have lower R^2's with respect to various systematic factors, but they do not consider tail risk exposures. Indeed, Bollen (2013) documents that funds with low R^2's relative to standard risk factors may have higher alphas on average but also a higher probability of failure, potentially consistent with the funds having higher FTS betas. Jiang and Kelly (2012) also find that hedge funds that lose value during high tail risk episodes earn higher average returns than funds that are hedged against tail risks. Our finding that most hedge fund categories have significantly positive alpha's and negative and frequently significant FTS betas is potentially consistent with their results if the high alpha funds dominate their aggregate fund categories. However, the managed futures category appears to be an exception, as it has no significant FTS exposure but still generates an overall significantly positive alpha. Similar findings are documented by Cao et al. (2014), who examine the performance of hedge fund categories in "good" and "bad" times and find that the Global Macro, Managed Futures, and Multi-Strategy styles provide investors with especially valuable hedges against bad times. It is also conceivable that the broad categories might have masked individual hedge fund effects, where some high quality funds do indeed hedge against FTS events. It would be of interest to examine the performance of individual hedge funds with respect to their behavior during FTS episodes in more detail.

5 Conclusions

We define a flight to safety event as a day on which bond returns are positive, equity returns are negative, the stock bond return correlation is negative, and there is market stress as reflected in elevated equity return volatility. Using only daily data on equity and bond returns, we identify FTS episodes in 23 countries. On average, FTS episodes comprise less than 3% of the sample, and bond returns exceed equity returns by about 2.5 to 4% on those days. FTS events are mostly country-specific as less than 25% can be characterized as global. Nevertheless, our methodology identifies major market crashes, such as October 1987, the Russia crisis in 1998 and the Lehman bankruptcy as FTS episodes. FTS episodes coincide with increases in the VIX and the TED spread, decreases in consumer sentiment indicators in the US, Germany and the OECD and appreciations of "safe-haven" currencies such as the Yen, the Swiss franc, and the US dollar. In equity markets, the financial, basic materials and industrial industries under-perform in FTS episodes, but the telecom industry outperforms. In bond markets, money market securities and corporate bonds have negative "FTS-betas". Liquidity deteriorates on FTS days in both equity and bond markets. Most commodity prices decrease sharply during FTS episodes, whereas the gold price measured in dollars increases slightly. Both economic growth and inflation decrease immediately following a FTS spell, and this decrease extends to at least one year after the spell.

We hope that our results will provide useful input to theorists positing theories regarding the origin and dynamics of flights to safety, or to asset pricers attempting to uncover major tail events that may drive differences in expected returns across different stocks and/or asset classes. They could also inspire portfolio and risk managers to look for portfolio strategies that may help insure against FTS-events, especially since we show that standard hedge fund strategies do not provide such insurance.

References

Acharya, Viral V., Lasse H. Pedersen, Thomas Philippon, and Mattthew Richards, 2012, Measuring systemic risk,, *CEPR Discussion Paper* No. DP8824.

Adrian, Tobias, and Markus K. Brunnermeier, 2011, CoVar, *NBER Working Paper* No. 17454.

Allen, Linda, Turan G. Bali, and Yi Tang, 2012, Does systemic risk in the financial sector predict future economic downturns?, *Review of Financial Studies* 25, 10.

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

Ang, Andrew, and Geert Bekaert, 2002a, International asset allocation with regime shifts, *Review of Financial Studies* 15, 1137–1187.

Ang, Andrew, and Geert Bekaert, 2002b, Regime switches in interest rates, *Journal of Business and Economic Statistics* 20, 163–182.

Ang, Andrew, and Dennis Kristensen, 2012, Testing conditional factor models, *Journal of Financial Economics* 106, 132–156.

Asness, Clifford, Robert Krail, and John Liew, 2001, Do hedge funds hedge?, *Journal of Portfolio Management* 28.

Baele, Lieven, Geert Bekaert, and Koen Inghelbrecht, 2010, The determinants of stock and bond return comovements, *Review of Financial Studies* 23, 2374–2428.

Baker, Malcolm, and Jeffrey Wurgler, 2006, Investor sentiment and the cross-section of stock returns, *Journal of Finance* 61, 1645–1680.

Bansal, Naresh, Robert A. Connolly, and Christopher T. Stivers, 2010, Regime-switching in stock index and Treasury futures returns and measures of stock market stress, *Journal of Futures Markets* 30, 753–779.

Barsky, Robert B., 1989, Why don't the prices of stocks and bonds move together, *American Economic Review* 79, 1132–1145.

Baur, Dirk G., and Brian M. Lucey, 2009, Flights and contagion–an empirical analysis of stock-bond correlations, *Journal of Financial Stability* 5, 339–352.

Beber, Alessandro, Michael W. Brandt, and Kenneth A. Kavajecz, 2009, Flight-to-quality or flight-to-liquidity? evidence from the euro-area bond market, *Review of Financial Studies* 22, 925–957.

Bekaert, Geert, and Eric Engstrom, 2009, Asset Return Dynamics under Bad Environment Good Environment Fundamentals, *NBER Working Paper* No. 15222.

Bekaert, Geert, Eric Engstrom, and Yuhang Xing, 2009, Risk, uncertainty, and asset prices, *Journal of Financial Economics* 91, 59–82.

Bollen, N. P., 2013, Zero-r2 hedge funds and market neutrality, *Journal of Financial and Quantitative Analysis* 48, 519–547.

Boyson, Nicole M., Christof W. Stahel, and Rene M. Stulz, 2010, Hedge fund contagion and liquidity shocks, *Journal of Finance* 65, 1789–1816.

Brownlees, Christian T., and Robert F. Engle, 2011, Volatility, correlation and tails for systemic risk measurement, Working Paper.

Brunnermeier, Markus K., and Lasse H. Pedersen, 2009, Market liquidity and funding liquidity, *Review of Financial Studies* 22, 2201–2238.

Caballero, Ricardo J., and Arvind Krishnamurthy, 2008, Collective risk management in a flight to quality episode, *Journal of Finance* 63, 2195–2230.

Cao, Charles, David Rapach, and Guofu Zhou, 2014, Which hedge fund styles hedge against bad times?, *Working Paper* .

Chordia, T., A. Sarkar, and A. Subrahmanyam, 2005, An empirical analysis of stock and bond market liquidity, *Review of Financial Studies* 18, 85–129.

Connolly, Robert, Chris Stivers, and Licheng Sun, 2005, Stock market uncertainty and the stock-bond return relation, *Journal of Financial and Quantitative Analysis* 40, 161–194.

Davies, Robert, 1987, Hypothesis testing when a nuisance parameter is present only under the alternative. 33-43, *Biometrika* 74, 33–43.

Engle, Robert F., Michael Fleming, Eric Ghysels, and Giang Nguyen, 2012, Liquidity, volatility, and flights-to-safety in the u.s. Treasury market: Evidence from a new class of dynamic order book models, *Federal Reserve Bank of New York Staff Report 590* .

Fontaine, Jean-Sébastien, and René Garcia, 2012, Bond liquidity premia, *Review of Financial Studies* 25, 1207–1254.

Fung, William, and David A. Hsieh, 2004, Hedge fund benchmarks: A risk based approach, *Financial Analyst Journal* 60, 65–80.

Fung, William, David A. Hsieh, Narayan Y. Naik, and T. Ramadorai, 2008, Hedge funds: Performance, risk, and capital formation, *Journal of Finance* 63, 1777–1803.

Gilbert, Ethel S., 1968, On discrimination using qualitative variables, *Journal of the American Statistical Association* 63, 1399–1412.

Goyenko, Ruslan, and Sergei Sarkissian, 2012, Treasury bond illiquidity and global equity returns, *Journal of Financial and Quantitative Analysis* forthcoming.

Goyenko, Ruslan, Avanidhar Subrahmanyam, and Andrey Ukhov, 2011, The term structure of bond market liquidity and its implications for expected bond returns, *Journal of Financial and Quantitative Analysis* 46, 111–139.

Goyenko, Ruslan, and Andrey D. Ukhov, 2009, Stock and bond market liquidity: A long-run empirical analysis, *Journal of Financial and Quantitative Analysis* 44, 189–212.

Goyenko, Ruslan Y., Craig W. Holden, and Charles A. Trzcinka, 2009, Do liquidity measures measure liquidity?, *Journal of Financial Economics* 92, 153–181.

Graveline, Jeremy J., and Matthew R. McBrady, 2011, Who makes on-the-run Treasuries special?, *Journal of Financial Intermediation* 20, 620–632.

Gurkaynak, Refet S., Brian Sack, and Jonathan H. Wright, 2007, The U.S. Treasury yield curve: 1961 to the present, *Journal of Monetary Economics* 54, 2291–2304.

Hamilton, James D, 1994, *Time Series Analysis* (Princeton University Press, New York).

Holden, Craig W., 2009, New low-frequency spread measures, *Journal of Financial Markets* 12, 778–813.

Hollo, Daniel, Manfred Kremer, and Marco Lo Duca, 2012, Ciss - a composite indicator of systemic stress in the financial system,, *ECB Working Paper* No. 1426.

Hu, Xing, Jun Pan, and Jiang Wang, 2013, Noise as information for illiquidiity, *Journal of Finance* 68, 2341–2382.

Jiang, Hao, and Bryan Kelly, 2012, Tail risk and hedge fund returns, *Chicago Booth Research Paper* No. 12-44.

Jordan, Bradford D., and Susan D. Jordan, 1997, Special repo rates: An empirical analysis, *Journal of Finance* 52, 2051–2072.

Kim, Chang-Jin, 1994, Dynamic linear models with Markov-switching, *Journal of Econometrics* 60, 1–22.

Krishnamurthy, Arvind, 2002, The bond/old-bond spread, *Journal of Financial Economics* 66, 463–506.

Longstaff, Francis, 2004, The flight to liquidity premium in US Treasury bond prices, *Journal of Business* 77, 511–526.

Mayfield, E. Scott, 2004, Estimating the market risk premium, *Journal of Financial Economics* 73, 465–496.

Naes, Randi, Johannes A. Skjeltorp, and Arne Odegaard, 2011, Stock market liquidity and the business cycle, *Journal of Finance* 66, 139–176.

Pasquariello, Paolo, and Clara Vega, 2009, The on-the-run liquidity phenomenon, *Journal of Financial Economics* 92, 1–24.

Pastor, Lubos, and Robert F. Stambaugh, 2003, Liquidity risk and expected stock returns, *Journal of Political Economy* 111, 642–85.

Patton, Andrew J., 2009, Are "market neutral" hedge funds really market neutral?, *Review of Financial Studies* 22, 2495–2530.

Perez-Quiros, Gabriel, and Allan Timmermann, 2001, Business cycle asymmetries in stock returns: Evidence from higher order moments and conditional densities, *Journal of Econometrics* 103, 259–306.

Roll, Richard, 1984, A simple implicit measure of the effective bid-ask spread in an efficient market, *Journal of Finance* 4, 1127–1139.

Sadka, Ronnie, 2012, Hedge-fund performance and liquidity risk, *Journal of Investment Management* 10, 60–72.

Teugels, Jozef L., 1990, Some representations of the multivariate bernoulli and binomial distributions, *Journal of Multivariate Analysis* 32, 256–268.

Titman, Sheridan, and Cristian Tiu, 2011, Do the best hedge funds hedge?, *Review of Financial Studies* 24, 123–168.

Vayanos, Dimitri, 2004, Flight to quality, flight to liquidity, and the pricing of risk, *NBER Working Paper* No. 10327.

White, Halbert, 1980, A heteroskedasticity-consistent covariance-matrix estimator and a direct test for heteroskedasticity, *Econometrica* 48, 817–838.

A Calculation of Joint FTS Dummy

Assume $\{X_i, i = 1, 2, ..., n\}$ is a sequence of Bernoulli random variables, where

$$P\{X_i = 0\} = q_i, \quad P\{X_i = 1\} = p_i$$

where $0 < p_i = 1 - q_i < 1$. The multivariate Bernoulli distribution is then represented by

$$p_{k_1,k_2,...,k_n} := P\{X_1 = k_1, X_2 = k_2, ..., X_n = k_n\}$$

where $k_i \in \{0, 1\}$ and $i = 1, 2, ..., n$. Let $\mathbf{p^{(n)}}$ be a vector containing the probabilities of the 2^n possible combinations of the n individual binary indicators. To define $\mathbf{p^{(n)}}$, we write k (with $1 \leq k \leq 2^n$) as a binary expansion:

$$k = 1 + \sum_{i=1}^{n} k_i 2^{i-1}$$

where $k_i \in \{0, 1\}$. This expansion induces a 1-1 correspondence

$$k \leftrightarrow (k_1, k_2, ..., k_n)$$

so that

$$p_k^{(n)} = p_{k_1,k_2,...,k_n}, \qquad 1 \leq k \leq 2^n$$

Teugels (1990) shows that $\mathbf{p^{(n)}}$ can be calculated as:

$$\mathbf{p^{(n)}} = \begin{bmatrix} 1 & 1 \\ -p_n & q_n \end{bmatrix} \otimes \begin{bmatrix} 1 & 1 \\ -p_{n-1} & q_{n-1} \end{bmatrix} \otimes ... \otimes \begin{bmatrix} 1 & 1 \\ -p_1 & q_1 \end{bmatrix} \sigma^{(\mathbf{n})}$$

where $\sigma^{(\mathbf{n})} = \left(\sigma_1^{(n)}, \sigma_2^{(n)}, ..., \sigma_{2^n}^{(n)}\right)^T$ is the vector of central moments than can be calculated as

$$\sigma_k^{(n)} = E\left[\prod_{i=1}^{n} (X_i - p_i)^{k_i}\right]$$

In our application, $n = 4$, with p_i corresponding to the FTS probability on a particular day based on the threshold model ($i = 1$), the Ordinal model ($i = 2$), and smoothed probabilities from the univariate RS ($i = 3$) and bivariate RS ($i = 4$) models, respectively. The Bernoulli variables X_i, $i = 1, .., 4$ are set to 1 when $p_i > 0.5$, and zero otherwise. The vector of central moments $\sigma_k^{(n)}$ is estimated over the full sample. Our joint FTS dummy is set to one when on that particular day the probability that at least 3 FTS measures signal a FTS is larger than 50%, i.e. when $p_{1,1,1,1} + p_{1,1,1,0} + p_{1,1,0,1} + p_{1,0,1,1} + p_{0,1,1,1} > 0.5$.

Table 1: The Ordinal FTS Measure

This table reports summary statistics for the Ordinal FTS measure discussed in Section 2.2.2. Column (1) reports summary statistics for the threshold level, calculated as the minimum of the ordinal numbers on days that satisfy a set of "mild" FTS conditions. Column (2) reports the percentage of observations that have an ordinal number above this threshold. Column (3) reports how much of those observations have an ordinal measure larger than 50% (calculated as 1 minus the percentage of false positives, i.e. the percentage of observations with an ordinal number above the threshold that do not meet our FTS criteria). Column (4) shows the percentage of observations in the full sample that have an ordinal FTS probability larger than 50%.

	(1) Threshold Level	(2) % observations > Threshold	(3) % (obs > threshold) with FTS prob. > 0.5	(4) % obs with FTS prob. > 0.5
US	0.772	6.9%	75.4%	5.2%
Germany	0.781	6.5%	98.7%	6.4%
UK	0.728	9.0%	65.3%	5.9%
Mean	0.723	10.5%	52.9%	5.2%
Median	0.723	10.3%	57.0%	5.1%
Min	0.650	4.8%	18.6%	2.7%
Max	0.804	19.3%	98.7%	7.9%
Interquartile	0.710	9.3%	39.1%	4.6%
Range	0.728	11.4%	64.9%	6.3%

Table 2: Estimation Results Regime-Switching FTS models

Panel A presents the estimation results for the Univariate 3-state Regime-Switching model described in Section 2.2.3. Panel B reports estimation results for the Bivariate Regime-Switching FTS model with jump terms as described in Section 2.2.4. We show detailed estimation results for the US, as well as the average and top/bottom quartile parameter estimates across all 23 countries. ***, **, and * represent statistical significance at the 1, 5, and 10 percent level, respectively. The FTS duration is expressed in days.

Panel A: Univariate 3-state RS FTS Model

	US	**Average**	**6th**	**17th**
Regime-dependent Intercepts (expressed in daily %)				
μ_1	-0.046***	-0.057	-0.079	-0.039
μ_2	-0.014	-0.020	-0.050	-0.007
μ_3	0.218*	0.249	0.198	0.271
Annualized Volatility Estimates				
σ_1	0.097***	0.105	0.087	0.122
σ_2	0.195***	0.201	0.166	0.217
σ_3	0.465***	0.473	0.408	0.498
FTS duration	36.3	26.7	17.2	35.3
# spells	18	26.4	17	31

Panel B: Bivariate RS FTS Model

	US	**Average**	**6th**	**17th**
Equity: Intercept + Jump Terms (expressed in daily %)				
α_0	0.076***	0.069	0.050	0.085
α_1	-1.275**	-2.367	-2.065	-0.246
α_2	1.732***	3.021	1.257	1.991
Bond: Intercept + Jump Terms (expressed in daily %)				
β_0	0.02***	0.030	0.029	0.033
β_1	-0.360	-0.775	-0.923	-0.332
β_2	-0.691***	-0.243	-0.578	0.068
FTS Estimates (expressed in daily %)				
α_3	-7.863***	-5.216	-7.416	-1.628
β_3	0.0001	0.814	0.000	0.976
ν	0.012***	2.079	0.014	0.047
Beta Estimates				
β_4	0.178***	0.030	-0.005	0.035
β_5	-0.344***	-0.166	-0.197	-0.111
Annualized Volatility Estimates				
$h_s\left(S_t^s = 1\right)$	0.104***	0.110	0.093	0.132
$h_s\left(S_t^s = 2\right)$	0.255***	0.286	0.246	0.324
$h_b\left(S_t^b = 1\right)$	0.021***	0.016	0.013	0.018
$h_b\left(S_t^b = 2\right)$	0.048***	0.036	0.031	0.038
FTS duration	89.9	86.8	58.1	101.2
# spells	24	16.2	10.1	18.5

Table 3: Percentage Number of FTS Instances

This table reports the percentage of days that a FTS is observed according to our aggregate FTS measure (Column 1) and 4 individual FTS measures (Columns 2 to 5). We show country-specific results and summary statistics (average, median, min, max, interquartile range) for our full sample of 23 countries.

Country	(1) Joint Prob.	(2) Threshold	(3) Ordinal	(4) Univ RS	(5) Bivar RS
US	3.00	0.74	5.17	7.98	21.74
Germany	4.14	0.91	6.37	11.31	26.77
UK	3.85	0.80	5.86	9.40	23.17
Switzerland	2.24	0.60	5.68	7.05	6.95
Japan	0.65	0.71	3.07	5.49	12.96
Canada	2.72	0.83	4.74	8.56	19.26
Sweden	4.75	0.70	6.66	14.59	28.20
Australia	1.03	0.78	1.80	3.72	17.71
Denmark	2.33	0.67	2.42	12.00	17.74
France	3.21	1.31	6.34	7.85	17.32
Belgium	3.67	0.74	4.34	8.83	16.66
Italy	2.39	0.94	3.28	8.17	10.16
New Zealand	0.31	0.72	1.82	1.99	1.78
Netherlands	4.75	0.93	5.29	12.18	17.26
Ireland	2.64	0.72	3.69	8.89	14.29
Spain	4.31	1.05	5.67	12.09	23.73
Austria	2.63	1.16	3.08	11.91	13.08
Czech Republic	0.71	0.82	2.59	2.96	5.55
Finland	3.02	0.93	4.76	19.20	14.80
Greece	2.08	0.71	2.52	19.75	13.08
Norway	0.26	0.70	0.16	10.83	3.52
Poland	0.53	0.90	2.07	10.88	3.46
Portugal	3.12	1.01	4.65	8.85	13.75
Average	2.54	0.84	4.00	9.76	14.91
Median	2.64	0.80	4.34	8.89	14.80
Min	0.26	0.60	0.16	1.99	1.78
Max	4.75	1.31	6.66	19.75	28.20
Interquartile	1.29	0.71	2.53	7.88	10.86
Range	3.55	0.93	5.58	11.97	18.88

Table 4: Contribution of each method to the joint FTS measure

This table reports the percentage overlap between the joint FTS dummy and one that is based on the probability that all methods other than the one specified in the column heading signal FTS.

	Threshold	Ordinal	Univ RS	bivar RS
US	96.81%	17.93%	21.12%	17.93%
Germany	98.85%	16.71%	17.87%	16.71%
UK	97.83%	18.01%	18.94%	17.08%
Average	82.89%	24.32%	19.19%	22.23%
Median	93.89%	18.33%	17.87%	17.93%
Minimum	0.00%	11.89%	0.00%	9.29%
Maximum	98.85%	53.85%	37.50%	56.00%
Interquartile	76.12%	15.99%	14.36%	15.38%
Range	95.71%	32.35%	22.40%	21.00%

Table 5: The Incidence of Global FTS

This table reports how many of the local FTS days are global in nature. We define a FTS event to be global when at least two-thirds of all countries experience a FTS on that same day. We report country-specific statistics for the US, Germany, and the UK, and summary statistics (average, min, max, interquartile range) for our full sample of 23 countries.

	Joint Prob. Measure		
	# FTS	# global	% global
US	251	35	13.9%
Germany	347	41	11.8%
UK	322	41	12.7%
Average	187	31	23.0%
Min	13	3	10.1%
Max	398	41	80.0%
Interquartile	94	25	14.0%
Range	275	40	22.9%

Table 6: Transmission of FTS

This table reports the average percentage of countries other than the row country i that are in a FTS conditional on country i being in a FTS. To account for asynchronous trading times, we assume that any spillover from the US/Canada affects Europe contemporaneously or on the following day, and affects the Pacific region on the following day. Pacific countries are assumed to have a same day effect on all other markets. European countries are assumed to have a same-day effect on the US/Canada, and a next-day impact on Pacific markets. The first pair of columns considers all FTS days, while the second pair of columns exclude global FTS days. The third pair of columns reports the percentage among the two days preceding a global FTS that are categorized as a FTS for the specified country. We rank countries (from high to low) based on the third column. Estimates that are more than 2 standard errors (calculated as the cross-sectional standard deviation divided by the square root of 23, the number of countries) above the means are shown in bold.

	All FTS days		excl. Global FTS		1-2 days before Global FTS	
	Rank	Perc	Rank	Perc	Rank	Perc
UK	1	**44.3%**	1	**45.5%**	2	**32.9%**
Germany	2	**42.5%**	2	**41.3%**	4	**31.7%**
Sweden	3	**39.9%**	3	**37.9%**	5	**30.5%**
US	4	**33.7%**	5	**31.3%**	2	**32.9%**
Netherlands	5	**30.9%**	6	**29.9%**	14	17.1%
France	6	**29.8%**	4	**31.4%**	1	**34.1%**
Canada	7	**29.6%**	7	**28.2%**	6	**26.8%**
Austria	8	**28.9%**	8	**28.0%**	7	**25.6%**
Spain	9	27.8%	9	**26.8%**	14	17.1%
Switzerland	10	26.9%	10	24.3%	10	22.0%
Italy	11	25.4%	15	19.6%	10	22.0%
Belgium	12	24.1%	12	21.8%	16	13.4%
Finland	13	23.7%	11	21.8%	18	11.0%
Ireland	14	23.3%	14	20.2%	12	19.5%
Portugal	15	21.1%	13	20.2%	10	22.0%
Denmark	16	18.8%	16	15.1%	8	**23.2%**
Greece	17	15.1%	17	11.5%	14	17.1%
Australia	18	13.1%	18	10.4%	17	12.2%
Czech Republic	19	8.8%	20	5.1%	19	9.8%
Poland	20	8.5%	22	2.4%	20	8.5%
Japan	21	7.9%	19	7.8%	23	1.2%
Norway	22	4.2%	23	2.3%	21	2.4%
New Zealand	23	2.2%	21	2.8%	21	2.4%
Summary Statistics						
Average	-	23.1%	-	21.1%	-	18.9%
Stdev	-	11.8%	-	12.5%	-	10.2%
Regional Averages						
North America	-	31.7%	-	29.8%	-	29.9%
Europe	-	24.7%	-	22.5%	-	20.0%
Europe - Developed	-	26.7%	-	24.8%	-	21.3%
Europe - Emerging	-	8.6%	-	3.8%	-	9.1%
Dev. Europe - euro	-	26.6%	-	24.8%	-	21.0%
Dev. Europe - non-euro	-	26.8%	43	25.0%	-	22.2%
Europe - GIIPS	-	22.5%	-	19.6%	-	19.5%
Pacific	-	7.7%	-	7.0%	-	5.3%

Table 7: FTS Dummies and Alternative Stress Indicators

This table reports estimates from regressions of changes in implied volatility measures, sentiment variables, safe have currency values and the TED spread on the joint aggregate FTS dummy (instances). VIX and country-specific implied volatility measures (i.e. VIX for US, Canada, VFTS for the UK, VDAX for the other European countries, and VJX for Japan, Australia and New Zealand), safe-haven currency values (i.e. the Swiss Franc, the Japanese Yen and the US dollar) and the TED spread (levels and changes) are available on a daily basis and are regressed on the FTS dummy. The sentiment variables are available on a monthly basis and are regressed on the fraction of FTS days within the month (expressed in decimals). Implied volatility and sentiment variables are expressed in absolute changes. The currency values are expressed in percentage changes (country currency per unit of safe currency). The sentiment variables include the Baker-Wurgler sentiment indicator (purged of business cycle fluctuations) and the Michigan consumer sentiment index which measure sentiment in the US, the Ifo Business Climate indicator (sentiment in Germany) and the (country-specific) OECD consumer confidence indicator (seasonally-adjusted). The TED spread data is not available for Ireland, Austria, Czech Republic and Poland. We show slope parameter estimates for the US, Germany and UK, as well as the average, standard deviation and top/bottom quartile parameter estimates across all 23 countries. The last column shows the number of countries for which the parameters estimates are significant at the 10% level. ***, **, and * represent statistical significance at the 1, 5 and 10 percent level, respectively, using heteroskedasticity-consistent standard errors.

	US	Germany	UK	Mean	Std	6th	17th	Sign.
Implied Volatility								
VIX	3.302***	1.707***	1.862***	1.686	0.891	1.017	1.781	22
Country-Specific	3.302***	2.112***	2.361***	2.542	1.405	1.461	3.302	23
Sentiment								
Baker-Wurgler	-1.182**	-0.248	-0.575	-0.966	1.600	-0.979	-0.068	6
Michigan	-3.676*	-4.289***	-4.621**	-5.993	10.262	-4.289	-2.104	10
Ifo Business	-2.939***	-2.817***	-3.041***	-4.774	4.541	-4.152	-2.502	21
OECD	-0.420***	-0.374***	-0.255***	-0.397	0.359	-0.561	-0.212	19
Currencies								
Swiss Franc	0.057***	0.154***	0.239***	0.398	0.555	0.096	0.305	19
Japanese Yen	0.200*	0.291***	0.464***	0.723	0.690	0.282	0.635	22
US Dollar	-	0.019	0.091**	0.347	0.553	0.067	0.342	19
TED Spread								
Levels	0.274***	0.217***	0.021**	0.006	0.666	-0.087	0.229	17
Changes	0.029***	0.005	0.007	0.018	0.036	-0.001	0.016	5

Table 8: FTS and Equity Portfolios

This table reports estimated slope coefficients from regressions of stock portfolio returns on the joint aggregate FTS dummy. The stock portfolios include Datastream industry portfolios (10 industry classification), MSCI style portfolios (large caps, mid caps, small caps, value and growth), a SMB portfolio (i.e. return of small cap portfolio minus return of large cap portfolio), and a HML portfolio (i.e. return of value portfolio minus return of growth portfolio). The portfolio returns are expressed in percentages on a daily basis and are denominated in their original currencies. In the regressions, we control for beta risks by adding a global factor (world market return) and a local factor (local stock market return). We show slope parameter estimates for the US, Germany and UK, as well as the average, standard deviation and top/bottom quartile parameter estimates across all 23 countries. The last column shows the number of countries for which the parameters estimates are significant at the 10 percent level. ***, **, and * represent statistical significance at the 1, 5 and 10 percent level, respectively, using heteroskedasticity-consistent standard errors.

	US	Germany	UK	Mean	Std	6th	17th	Sign.
Industry Portfolios								
Oil & Gas	-0.061	-0.359	0.045	-0.200	0.513	-0.381	0.110	9
Basic Materials	-0.296***	-0.048	-0.297***	-0.235	0.385	-0.355	-0.027	9
Industrials	-0.118**	-0.026	-0.015	-0.194	0.316	-0.353	-0.052	12
Consumer Goods	0.173***	-0.075	0.093	-0.167	0.359	-0.329	0.093	8
Health Care	0.196***	0.009	0.141**	-0.180	0.867	-0.201	0.079	8
Consumer Services	0.142***	-0.012	0.026	-0.083	0.353	-0.203	0.142	8
Telecom	0.231**	0.365***	0.316***	0.304	0.544	0.181	0.373	17
Utilities	-0.105	-0.083	0.061	-0.085	0.364	-0.354	0.072	7
Financials	-0.288***	-0.224**	-0.234***	-0.244	0.279	-0.387	-0.171	16
Technology	0.107	-0.090	-0.459***	-0.119	0.363	-0.252	0.117	7
Style Portfolios								
Large Cap	0.026***	-0.117*	0.048***	0.151	0.327	-0.013	0.172	12
Mid Cap	-0.130***	-0.331***	-0.189***	-0.188	0.192	-0.330	-0.064	11
Small Cap	-0.173***	-0.156	-0.269***	-0.312	0.235	-0.470	-0.156	18
Value	-0.047	-0.246	0.051	-0.040	0.102	-0.099	0.022	3
Growth	0.042	0.007	0.020	0.099	0.162	0.007	0.155	6
SMB	-0.197***	-0.040	-0.318***	-0.470	0.515	-0.692	-0.141	15
HML	-0.089	-0.253	0.030	-0.139	0.201	-0.253	-0.025	6

Table 9: FTS and Bonds

This table reports estimated slope coefficients from regressions of bond yields, spreads and portfolio returns on the joint aggregate FTS dummy. Panel A shows the results for the bond yields and spreads. For the yields, we consider the level, slope and curvature factors of the yield curve. The level factor is calculated as the average of the 3-month bill yield and the 5- and 10-year government bond yields, the slope factor as the 10-year government bond yield minus the 3-month bill yield, and the curvature factor as the sum of the 10-year goverment bond yield and the 3-month bill yield, minus two times the 5-year goverment bond yield. We also show results for the 10-year benchmark government bond yield and the monetary policy target rates. For the spreads, we consider two default spread measures, the yield on the AAA portfolio minus the 10-year government bond yield, and the yield on the BBB portfolio minus the yield on the AAA portfolio. The bond yields and spreads are expressed relative to a 150-day moving average. Regressions reported in Panel A have no other control variables. Panel B shows the results for returns on individual bond portfolios, including JP Morgan cash indices (1, 3, 6 and 12 months), benchmark Datastream government bond indices (2, 5, 20 and 30 years), and BOFA ML coporate bond indices (with AAA, AA, A and BBB ratings). The corporate bond indices are only available for the US, Japan, Canada, Australia, and the Eurozone as a whole. We use the Eurozone corporate bond index for European countries and the corporate bond index of Australia for New Zealand. Panel C considers returns on 4 spread portfolios: the return on the 10-year government bond minus that on the 1-month cash index, the return on the 10-year goverment bond minus that on the 2-year goverment bond, the return on the 10-year goverment bond minus that on the AAA portfolio, and the return on the BBB portfolio minus that on the AAA portfolio. Thus, the first 2 portfolios primarily reacts to changes in the level of yields and term spread, and the latter 2 to changes in default risk. Regressions in Panel B and C control for the 10-year benchmark government bond return. In the regressions for the corporate bonds, we also control for the local stock market return. All yields, spreads and returns are daily and denominated in local currencies. We show estimated parameter loadings on the joint FTS indicator for the US, Germany and the UK, as well as the average, standard deviation and top/bottom quartile parameter estimates across all 23 countries. The last 2 columns show the number of countries for which the parameters estimates are significant at the 10 percent level and the number of countries for which data is available, respectively. ***, **, and * represent statistical significance at the 1, 5 and 10 percent level, respectively, using heteroskedasticity-consistent standard errors.

	US	Germany	UK	Mean	Std	6th	17th	Sign.	Obs
Panel A: Yields and spreads									
Gov Level	-0.447***	-0.276***	-0.302***	-0.128	0.520	-0.405	-0.109	21	23
Gov Slope	-0.209***	-0.152***	-0.059	-0.123	0.357	-0.258	-0.041	17	23
Gov Curvature	0.357***	0.355***	0.357***	0.191	0.699	0.184	0.428	21	23
Gov 10 Year	-0.442***	-0.293***	-0.264***	-0.134	0.509	-0.342	-0.179	21	23
MP Target Rates	-0.144***	-0.140***	-0.157***	-0.100	0.218	-0.140	0.001	18	23
AAA - Gov 10 Year	0.426***	0.131***	0.136***	0.027	0.449	-0.089	0.262	21	23
BBB - AAA	0.415***	0.666***	0.649***	0.557	0.292	0.379	0.780	20	23
Panel B: Bond portfolio returns									
Cash 1 Month	-0.008***	-0.004***	-0.009***	-0.007	0.003	-0.010	-0.006	15	17
Cash 3 Month	-0.007***	-0.004***	-0.009***	-0.007	0.003	-0.009	-0.005	15	17
Cash 6 Month	-0.008***	-0.003***	-0.010***	-0.006	0.005	-0.010	-0.004	14	17
Cash 12 Month	-0.010**	-0.001	-0.009**	-0.005	0.006	-0.011	-0.001	7	16
Gov 2 Year	-0.019***	0.008*	-0.005	-0.001	0.031	-0.021	0.011	12	21
Gov 5 Year	-0.005	0.034***	0.025**	0.020	0.033	-0.003	0.036	9	23
Gov 20 Year		-0.007	0.110***	0.013	0.045	-0.010	0.014	2	9
Gov 30 Year	0.157***	0.006	0.182***	0.039	0.078	0.006	0.052	3	12

	US	Germany	UK	Mean	Std	6th	17th	Sign.	Obs
Panel B: Bond portfolio returns (continued)									
AAA	-0.031	-0.007	-0.022	0.012	0.262	-0.047	-0.016	3	23
AA	-0.057***	-0.041**	-0.050***	-0.010	0.262	-0.061	-0.043	21	23
A	-0.082***	-0.078***	-0.092***	-0.068	0.279	-0.132	-0.084	21	23
BBB	-0.078***	-0.070**	-0.097***	-0.069	0.280	-0.138	-0.090	20	23
Panel C: Spread portfolio returns									
Gov 10 Year - Cash 1 Month	0.008***	0.004***	0.009***	0.007	0.003	0.004	0.009	15	17
Gov 10 Year - 2 Year	0.019**	-0.008*	0.005	-0.001	0.032	-0.012	0.013	12	23
Gov 10 Year - AAA	0.033*	0.006	-0.010	-0.105	0.275	-0.112	-0.006	11	23
AAA - BBB	0.042*	0.062**	0.065***	0.070	0.048	0.039	0.108	15	23

Table 10: Liquidity and FTS

This table reports estimated slope coefficients from regressions of US bond (Panel A) and equity market (Panel B) illiquidity measures on the joint aggregate FTS dummy. Our bond market illiquidity measures are (1) the monthly effective spread, a cross-sectional monthly average of proportional quoted spreads of Treasury bonds with a maturity of at most one year (in %), (2) the daily Treasury on/off-the-run spread, calculated as the negative of the daily difference in yields between an on-the-run Treasury bond and a synthetic off-the-run Treasury security with the same coupon rate and maturity data (in basis points), and (3) the 'noise' measure of Hu et al. (2013). Our equity market illiquidity measures are monthly cross-sectional averages of (1) the effective tick measure from Holden (2009), (2) Amihud (2002)'s price impact measure, and (3) the negative of the Pastor and Stambaugh (2003) price impact measure. When the measures are non-stationary over the sample, we use values relative to either a 150-day or 6-month moving average. The regressions include only a constant and the FTS measure as independent variable. When the illiquidity measure is only available at the monthly frequency, we regress it to the percentage of FTS days within that month (expressed in decimals). ***, **, and * represent statistical significance at the 1, 5 and 10 percent level, respectively, using heteroskedasticity-consistent standard errors.

	Level	
	α	β_{FTS}
Panel A: Bond Illiquidity Measures		
Proportional Spread	-0.11***	0.44***
Treasury On/Off-the-run Premiums	14.35***	9.95***
Noise Measure Hu, Pan, Wang (2012)	-0.12***	1.22***
Panel B: Equity Illiquidity Measures		
Effective Tick	-0.04**	0.64***
Amihud	2.46***	7.72***
(negative of) Pastor-Stambaugh	0.02***	0.22***

Table 11: FTS and Commodity Prices

This table reports estimated slope coefficients in regressions of the S&P GSCI benchmark commodity index returns (in US dollar) on the joint aggregate FTS dummy, using the global equity market return (in US dollar) as a control. We consider broad indices (Commodity Total, Energy, Industrial Metals, Precious Metals, Agriculture, Livestock) and sub-indices (Crude Oil, Brent Crude Oil and Gold). The returns are expressed in percentage terms on a daily basis and are denominated in US dollar. We show FTS slope parameter estimates for the US, Germany and the UK, as well as the average, standard deviation and top/bottom quartile parameter estimates across all 23 countries. The last two columns report the average market beta as well as the number of countries for which the FTS slope parameter estimates are significant at the 10 percent level, respectively. ***, **, and * represent statistical significance at the 1, 5 and 10 percent level, respectively, using heteroskedasticity-consistent standard errors.

	US	Germany	UK	Mean	Std	6th	17th	Beta	Sign.
Commodity Total	-0.396***	-0.306***	-0.393***	-0.294	-0.469	0.318	-0.590	0.306	20
Energy	-0.501***	-0.373***	-0.454***	-0.444	-0.525	0.340	-0.678	0.332	18
Industrial Metals	-0.052	-0.239**	-0.360***	-0.317	-0.406	0.238	-0.496	0.519	19
Precious Metals	0.467***	0.254***	0.196**	0.123	0.208	0.309	0.116	0.167	11
Agriculture	-0.183*	-0.121	-0.205**	0.182	-0.289	0.399	-0.282	0.253	11
Livestock	-0.084	-0.114**	-0.170***	-0.058	-0.152	0.146	-0.165	0.106	9
Crude Oil	-0.676***	-0.457***	-0.555***	-0.516	-0.611	0.376	-0.784	0.352	18
Brent Crude Oil	-0.521***	-0.201	-0.266*	-0.116	-0.512	0.513	-0.673	0.551	15
Gold	0.470***	0.269***	0.219**	0.134	0.242	0.297	0.134	0.128	13

Table 12: FTS and the Macroeconomy

This table reports estimated slope coefficients from regressions of various macroeconomic variables on the fraction of FTS days within the month or quarter (expressed in decimals). The macroeconomic variables include 3-month changes in the CPI, industrial production (IP), the unemployment rate, and the OECD leading indicator over the next three months, as well as quarterly changes in real GDP and the investment-to-GDP ratio over the following quarter. For inflation, IP growth, GDP growth, the unemployment rate and investment growth, we also have monthly survey forecasts from Consensus Economics and we examine both the mean and the standard deviation of 12-month ahead individual forecasts 3 month hence. 12-month constant-horizon forecasts are calculated by interpolating between current- and next-year forecasts. In the lines with variables marked "future", we regress annual growth rates or annual changes in the economic variables over the following year on the fraction of days of FTS instances within the month (expressed in decimals). We show slope parameter estimates for the US, Germany and the UK, as well as the average, standard deviation and top/bottom quartile parameter estimates across all 23 countries. The second to last column shows the number of countries for which the parameters estimates are significant at the 10 percent level. The last column shows the number of countries for which the real economy data is available. ***, **, and * represent statistical significance at the 1, 5 and 10 percent level, respectively, using Newey-West standard errors (using either 6 quarterly or 24 monthly lags).

	US	Germany	UK	Mean	Std	6th	17th	Sign.	Obs
Inflation	-1.374**	-0.925***	-0.980**	-1.104	0.945	-1.374	-0.761	18	23
Inflation Forecast Mean	-2.718**	-1.842***	-1.792*	-3.227	5.344	-2.623	-0.888	17	23
Inflation Forecast St. Dev.	0.229	0.106***	0.157	0.274	0.555	0.055	0.229	6	16
Future Inflation	-3.624***	-3.094***	-3.524**	-3.307	2.701	-3.958	-2.404	15	23
IP Growth	-3.237	-4.215**	-2.448*	-6.603	8.410	-5.709	-2.448	10	14
IP Growth Forecast Mean	-5.356	-5.099**	-4.261**	-6.339	5.575	-6.848	-3.182	16	22
IP Growth Forecast St. Dev.	0.466	1.242**	0.620**	1.139	1.147	0.471	0.941	8	15
Future IP Growth	-9.314***	-4.057	-5.761	-13.758	19.400	-9.917	-4.057	8	14
GDP Growth	-2.273***	-2.904***	-1.738	-5.072	5.794	-5.665	-1.855	17	23
GDP Growth Forecast Mean	-2.512	-2.593**	-2.469*	-3.785	2.728	-5.312	-2.058	20	23
GDP Growth Forecast St. Dev.	0.149	0.139	0.194***	0.442	0.689	0.116	0.216	8	16
Future GDP Growth	-4.370	-6.379**	-5.268	-14.870	21.934	-13.112	-3.912	13	23
Unemployment	0.849***	0.272**	0.134	0.597	0.598	0.134	1.026	13	23
Unemployment Forecast Mean	0.725	0.473**	0.705**	1.069	0.963	0.473	1.088	7	9
Unemployment Forecast St. Dev.	0.080***	0.007	0.065	0.105	0.108	0.007	0.156	5	9
Future Unemployment	2.815**	0.462	0.619	3.006	3.857	0.732	2.815	12	23
Investment/GDP	-0.960**	-0.318	-0.318	-1.499	2.759	-2.115	-0.083	6	20
Investment Growth Forecast Mean	-2.900	-2.958**	-2.265***	-4.114	3.792	-7.951	-1.936	9	13
Investment Growth Forecast St. Dev.	0.640***	-0.088	0.556***	0.503	0.742	0.061	0.640	5	13
Future Investment/GDP	-2.267	-0.344	-2.268***	-4.794	8.645	-4.143	-0.395	7	20
OECD Leading Indicator	-0.992*	-0.645**	-0.577**	-1.308	1.652	-1.617	-0.577	15	23
Future OECD Leading Indicator	2.211*	2.889*	0.898	0.535	2.115	-0.777	1.671	5	23

Table 13: Hedge Fund Index returns and FTS

This table reports estimated slope coefficients from regressions of returns on different hedge fund indices on a constant (α), the contemporaneous return on the S&P500 ($\beta_m(t)$) and its first lag ($\beta_m(t-1)$), and our joint aggregate FTS dummy (β_{FTS}). The last column reports FTS betas (β^{FH}_{FTS}) from a regression which controls for the Fung and Hsieh (2008) risk factors. ***, **, and * represent statistical significance at the 1, 5, and 10 percent level, respectively, and are based on heteroskedasticity-consistent standard errors.

	α	$\beta_m(t)$	$\beta_m(t-1)$	β_{FTS}	R^2	β^{FH}_{FTS}
Hedge Fund Index (HFI)	0.007***	0.246***	0.042	-0.034**	36.2%	-0.026**
Convertible Arbitrage HFI	0.005***	0.139***	0.119***	-0.015	21.9%	-0.013
Dedicated Short Bias HFI	0.003	-0.848***	-0.072	-0.039	58.5%	-0.038
Emerging Markets HFI	0.004	0.485***	0.098	-0.028	30.8%	-0.021
Equity Market Neutral HFI	0.006***	0.123***	0.113	-0.074	21.5%	-0.074
Event Driven HFI	0.007***	0.211***	0.092***	-0.030***	51.4%	-0.031***
Event Driven Distressed HFI	0.007***	0.223***	0.102***	-0.030***	48.9%	-0.031***
Event Driven Multi-Strategy HFI	0.007***	0.204***	0.089***	-0.032**	43.3%	-0.033***
Event Driven Risk Arbitrage HFI	0.004***	0.12***	0.026	-0.011	26.6%	-0.010
Fixed Income Arbitrage HFI	0.004***	0.093*	0.064*	-0.040***	24.3%	-0.031***
Global Macro HFI	0.01***	0.141**	-0.033	-0.030	7.4%	-0.014
Long/Short Equity HFI	0.006***	0.395***	0.034	-0.022	43.6%	-0.016
Managed Futures HFI	0.005**	-0.065	-0.081	0.002	2.1%	0.010
Multi-Strategy HFI	0.005***	0.104***	0.072***	-0.016	20.1%	-0.013

Figure 1: FTS Incidence and Return Impact for the Threshold Model

This figure plots the average cross-country incidence of FTS for κ ranging from 0 to 3, both in the actual data (solid line) and in data simulated from bivariate normal distributions (dashed line). For a given threshold level κ, we identify a FTS day as a day on which the bond return is κ standard deviations above zero and the equity return is κ standard deviations below zero. We allow for low-frequency variations in country-specific bond and equity return standard deviations as well as in their pairwise correlations, and model it using a symmetric kernel method with a bandwidth of 250 days. In the simulation exercise, we draw for each country stock and bond returns from a bivariate normal distribution with country-specific full-sample means and time-varying standard deviations and correlations, and report for each κ the average FTS incidence across countries. The dotted line plots for each κ the average cross-country return impact on FTS days, where the return impact is calculated as the difference between the country-specific bond and equity returns.

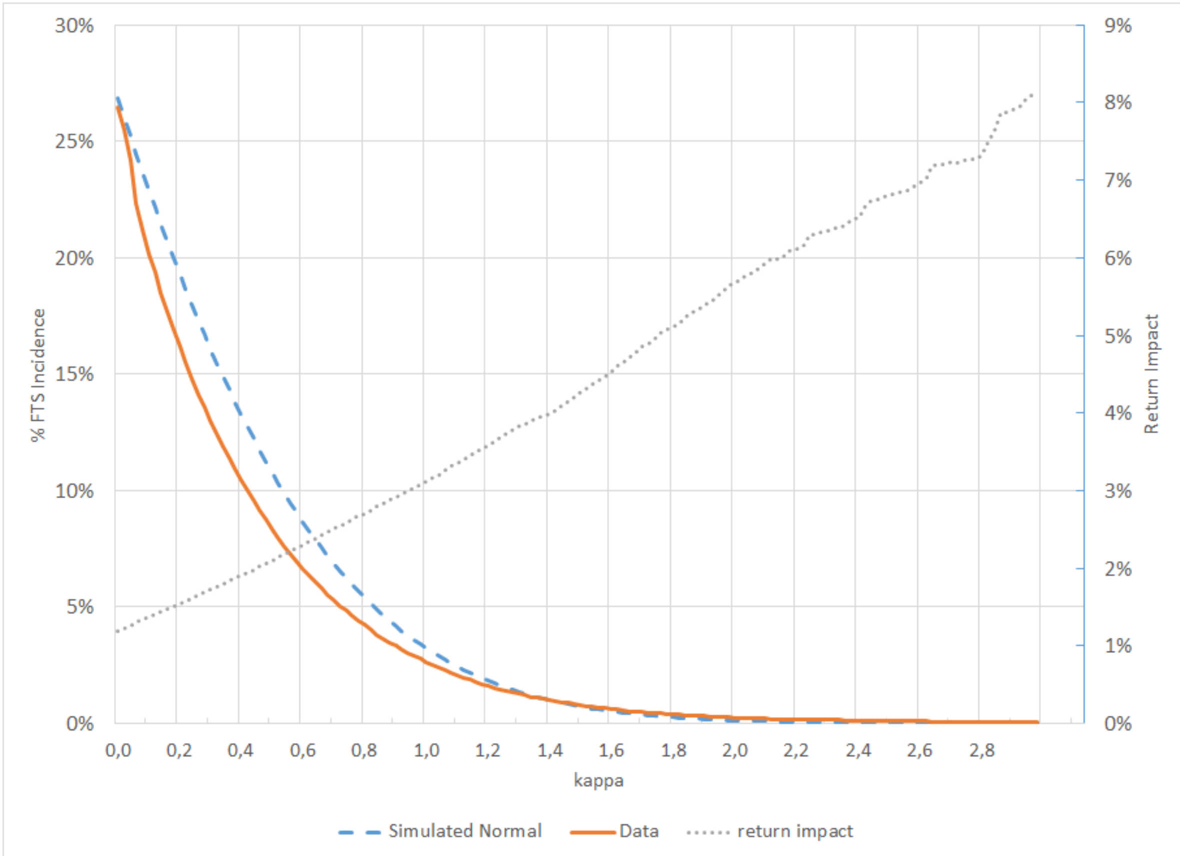


Figure 2: Ordinal Indicator: US, Germany, and UK

The left panels plot the ordinal FTS indices for the US, Germany, and the UK and the corresponding minimum threshold levels, calculated as the minimum of all ordinal values for which the minimal FTS conditions hold. The right panels plot the derived ordinal FTS measures. Values above 0.5 are depicted in black, and values below 0.5 in light grey.

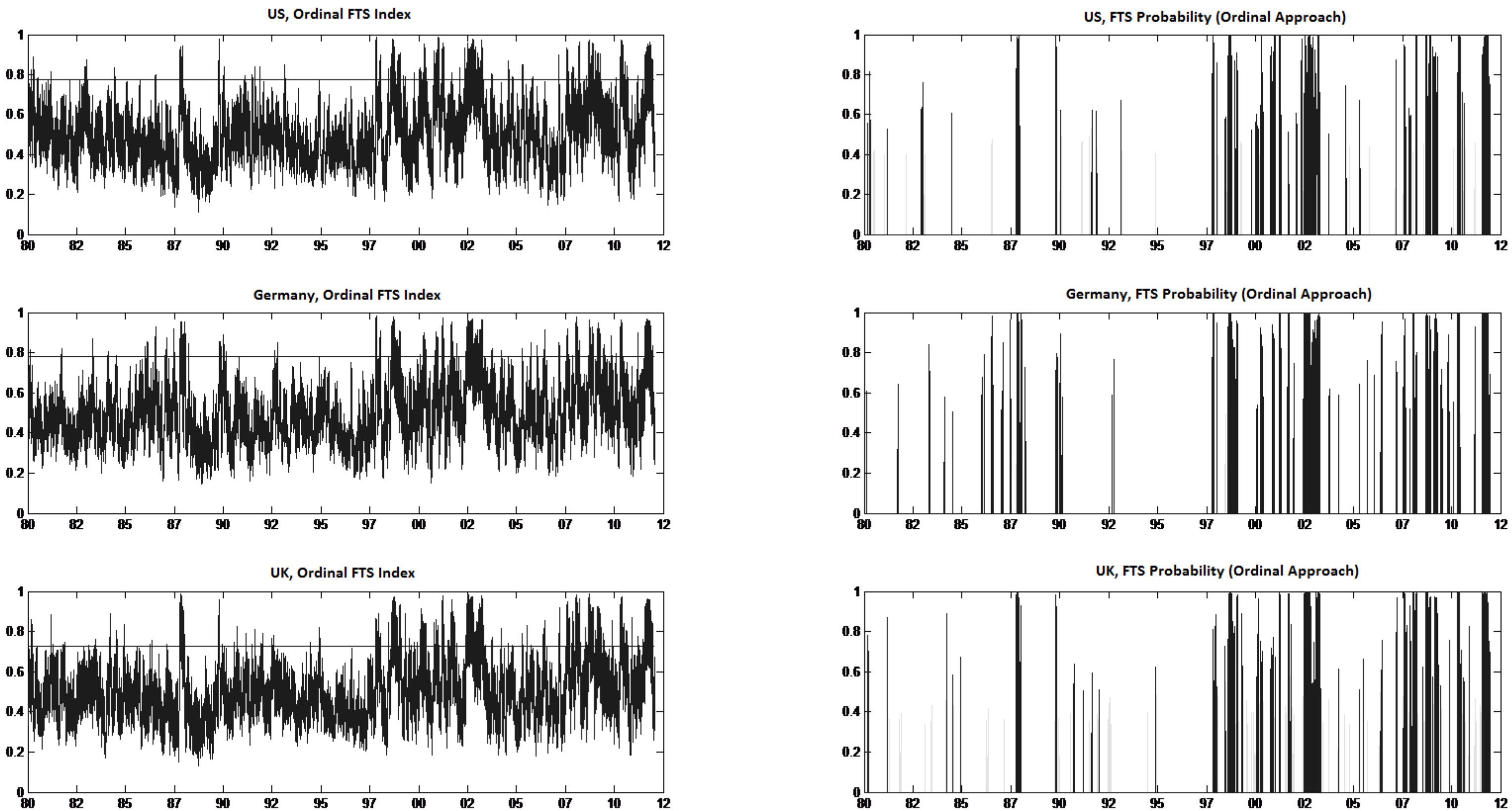


Figure 3: Aggregate FTS Measure and Dummy, US

The top panel of this figure plots the joint FTS probability for the US. The bottom panel plots the corresponding joint FTS dummy, which equals one when the FTS probability is larger than 50%, and zero otherwise.

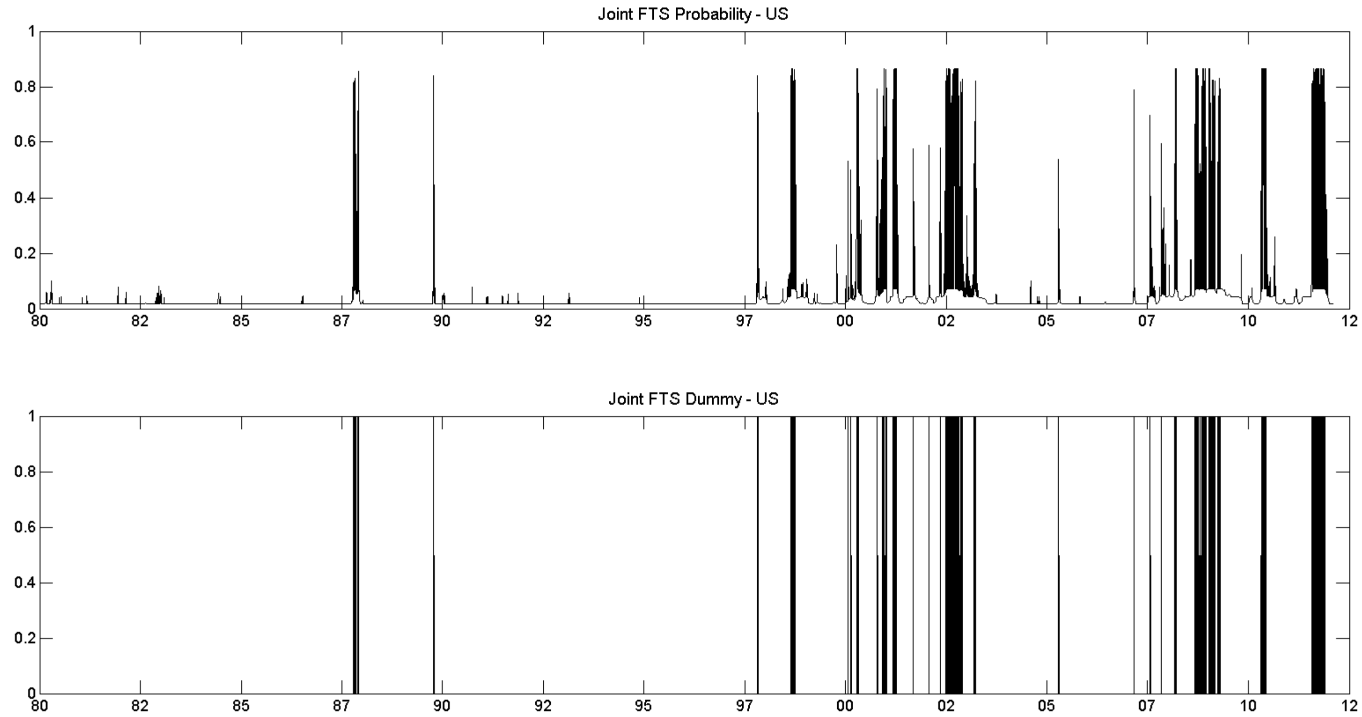
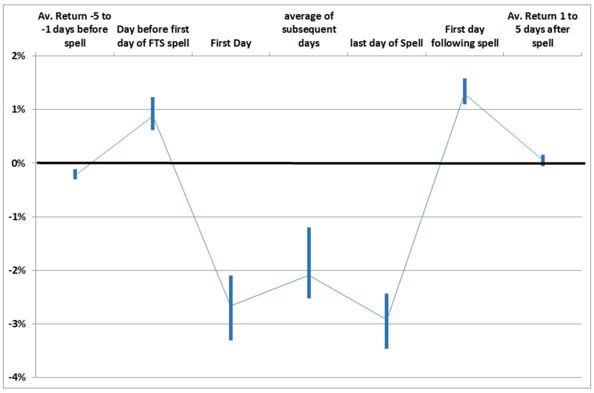


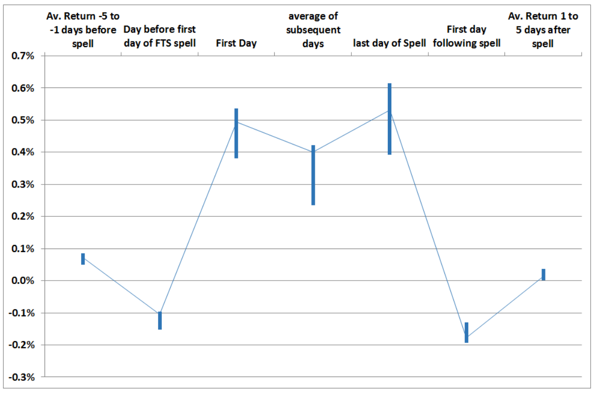
Figure 4: Return Impact before, during, and after FTS

This figure plots the average return across all countries and their cross-country interquartile range before, during, and after a FTS spell for equities (Panel A), bonds (Panel B), and return impact (Panel C) in event time. For each FTS spell, we calculate the average returns (1) in the 5 days before the start of the spell, (2) on the day right before the start of the spell, (3) on the first day of the spell, (4) on any subsequent day within the spell (except that last one), (5) on the last day of the spell, (6) on the first day following the spell, and (7) during the 5 days following a spell.

Panel A: Equities



Panel B: Bonds



Panel C: Return Impact

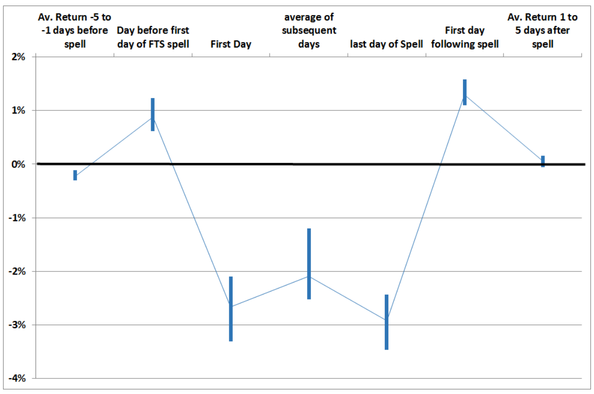


Figure 5: The Persistence of FTS Events

This figure plots the average number of FTS spells that last 1, 2, 3, 4 to 9, 10 to 49, 50 to 99, or 100 days or more based on each of the 4 individual methods as well as for the aggregated measure. For reference, the average number of FTS spells for each methodology is 54 for the threshold model, 139 for the ordinal model, 26 for the univariate regime-switching model, 15 for the bivariate regime-switching model, and 83 for the joint measure.

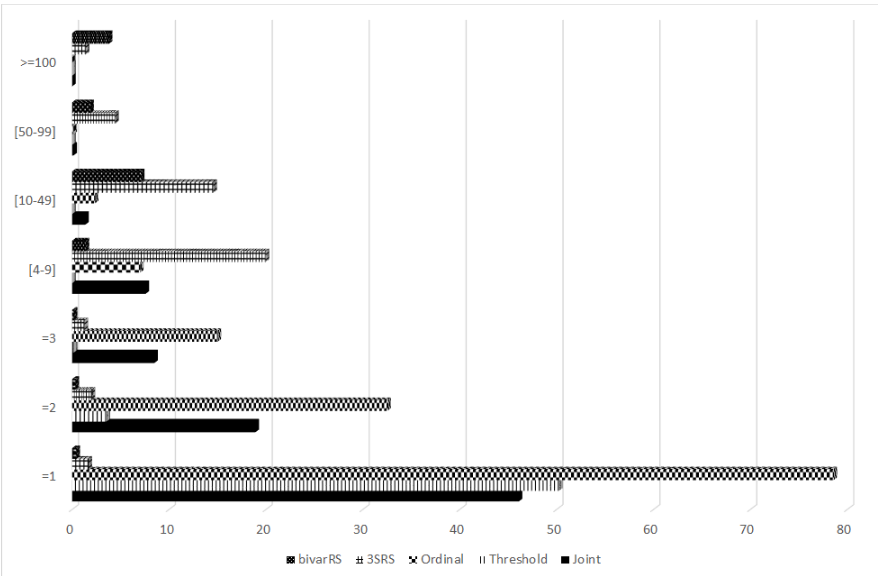


Figure 6: Percentage of Countries in FTS

This figure plots the percentage of countries experiencing a FTS at each point in time.