Abstract

We propose a new nonparametric test to identify mutually exciting jumps in high frequency data. We derive the asymptotic properties of the test statistics and show that the tests have good size and reasonable power in finite sample cases. Using our mutual excitation tests, we empirically characterize the dynamics of financial flights in forms of flight-to-safety and flight-to-quality. The results indicate that mutually exciting jumps and risk-off trades mostly occur in periods of high market stress. Flight-to-safety episodes (from stocks to gold) arrive more frequently than do flight-to-quality spells (from stocks to bonds). We further find evidence that reverse cross-excitations or seeking-return-strategies exhibit significant asymmetry over the business cycle, reflecting the fact that investors appear to be selling gold—rather than bonds—to invest in stocks during good market conditions.

Keywords: Flight-to-safety, Flight-to-quality, Mutual-excitation in jumps, Stock-bond comovement, High-frequency data

JEL: G01, G12, G15, C12, C14, C58
1. Introduction

“U.S. and European stocks dived after a sharp selloff in Chinese shares accelerated, wiping out gains for the year. Oil prices continued to drop, while Treasurys gained as investors sought the relative safety of government bonds.”

(Wall Street Journal, August 24, 2015, 3:55 pm—EST)

“This is a flight to quality and the actual level that the Treasury yield achieves in this environment is not meaningful. This is a time when you dig a deep hole, close your eyes and put your fingers in your ears.”

(David Keeble, Crédit Agricole, Wall Street Journal, August 24, 2015, 4:24 pm—EST)

The global financial system is vulnerable to unforeseen shocks and unexpected events. Sudden shocks may, for instance, amplify uncertainty, which, in turn, exacerbates fear and increases risk perception in the marketplace. The downside effects of financial stress could be even worse, especially when the shocks in one region (or market) spread to others through contagion and market linkages.\(^1\) Changes in market sentiment can quickly lead to the flights to safety (FTS) or flight to quality (FTQ) phenomena highlighted in the quotes above.

The empirical evidence for transmission of shocks across markets and assets is for the large part conducted using frequencies of one day or more; a recent exception being Aït-Sahalia and Xiu (2016). This focus is at odds with the modern concern with the fast-moving behavior of financial markets for many assets. For example, in the United States, the House voted against the bailout of Lehman Brothers on September 29, 2008. As Figure 1 displays, the market reaction to this event was rather swift at the intraday level: the S&P 500 index declined sharply around 18:00 GMT, within minutes after the news shock, and the VIX index—a common proxy for market volatility—elevated substantially.

This paper attempts to fill the gap on estimating transmission in high frequency. In particular, we focus on the cases of FTS from U.S. stock market to gold, and FTQ from U.S. stock market

\(^1\)Researchers define the term financial contagion in different ways; for an overview of these in the empirical literature see e.g., Dungey et al. (2005).
to long-dated Treasury bonds, respectively. While the recent work by Aït-Sahalia et al. (2015) considers contagion through the lens of mutually-exciting Hawkes processes, and Aït-Sahalia and Xiu (2016) examine cojumps amongst different asset classes, there is to our knowledge no high frequency work on the specific identification of FTS or FTQ as distinct categories of transmission between asset classes. These questions are important in understanding changing measures of market risk (embedded in risk spillovers), credit risk (associated with default spillovers) and systemic/system-wide risk (linked to tail comovements). For portfolio managers, understanding the linkages of shocks between asset classes aids in evaluating diversification strategies. The risk exposure of portfolios to sudden shocks depends on the dynamics of these linkages.

[ Insert Figure 1 about here ]

The literature offers two main directions to study the unanticipated transmission of shocks across assets. One strand investigates the sources of shocks that can be linked to information arrivals (Harvey and Huang, 1991, Ederington and Lee, 1993, King and Wadhwani, 1990, Calvo and Mendoza, 2000), liquidity or trading activity (Diamond and Dybvig, 1983, Caballero and Krishnamurthy, 2008, Jiang et al., 2011, Fleming et al., 1998), and real shocks such as the innovations to preferences and consumer behavior in one particular country or asset class (Allen and Gale, 2000). Another body of research focuses on volatility transmission and correlation across markets. Pelletier (2006) extends the work of Allen and Gale (2000) and considers different regime structures in uncovering cross correlations.

Examinations of transmissions across asset classes are rarer, but evidence of this phenomenon is found in Chan et al. (2011). Specifically, related to the FTQ hypothesis, Guidolin and Timmermann (2006, 2007) characterize the nonlinear dynamics of the stock-bond relation, while Chang and Cheng (2016) provide evidence that FTQ occurs primarily to long-dated bonds. Evidence of FTS, particularly for the U.S., is provided in Baur and McDermott (2010). These studies, and those concentrated on single asset classes over international borders, suggest that financial volatility

---

2The literature on this direction is rather vast. For currency market linkages, see, e.g., Diebold and Nerlove (1989), Mahieu and Schotman (1994), Kaminsky and Reinhart (2000); for stock market linkages, see, e.g., Bae et al. (2003), Corradi et al. (2012), Forbes and Rigobon (2002), King and Wadhwani (1990), King et al. (1994), Lin et al. (1994), Karolyi (1995), Bekaert et al. (2005), Hamao et al. (1990), Karolyi and Stulz (1996); and for bond market linkages, see, e.g., Favero and Giavazzi (2002) and Dungey et al. (2000).
exhibits strong connectedness that varies over time with market conditions.

Compared with volatility analysis, however, estimation of the spillover of jumps and cojumps between assets is relatively more challenging. The first possible reason is broadly linked to the underlying process that generates jump-type abnormal shocks. When researchers seek to exploit the patterns in high-frequency data, the observed sharp movements could be associated with jumps in the asset prices (e.g., Andersen et al., 2007a,c; Lee and Mykland, 2008), jumps in volatility (Todorov and Tauchen, 2011) or both. The measurement of cojumps requires the identification of distinct properties (e.g., price cojumps, volatility cojumps or price-intensity cojumps). Moreover, as Aït-Sahalia and Jacod (2012) point out, tail-type jump events tend to occur more frequently at the intradaily—rather than daily or weekly—sampling frequencies. Therefore, a model that fails to characterize jump linkages at a relatively high frequency might underestimate the true dependence in shocks. The second challenge is related to the empirical characteristics of volatility. In financial data, one might notice that the periods of high volatility mingle with the periods of cojumps. This is not surprising because sudden crashes might trigger fear and turbulence, which, in turn, increase market-wise volatility. Such patterns, however, create complications for the estimation of market risk embedded in jump spillovers and volatility transmission. It is thus crucial to separate joint jump shocks from volatility linkages. Perhaps more importantly, those jump-type shocks are likely to propagate over time (e.g., during the crisis), and tail events may spread with delay. In the presence of such regularities, understanding the origin (or direction) of jump spillovers may be difficult and ambiguous. Although correlation/covariance-based measures are certainly of interest in this context, jump-type shock transmission remains elusive.

In this paper, we develop a new approach to measure the transmission of jumps in financial markets. Relying on nonparametric (model-free) testing procedures, this approach allows us to characterize the dynamic properties of shock transmission in greater depth than is possible with correlation/covariance analysis or traditional cojump techniques. In our continuous-time setup,

---

3See Bandi and Renò (2016), who provide evidence that volatility jumps lead to price jumps and such price-volatility cojumps are negatively correlated in the data.

4It is important to note that shock transmission could be attributable to both volatility and jump spillovers. Developing nonparametric tests, Corradi et al. (2012) show that the main channel of shock transmission is through volatility linkages rather than jump spillovers. In parallel to our framework, Corradi et al. (2016) propose a nonparametric test directly for identifying jump spillovers without testing for jumps.
jump-type shocks in one financial asset have the potential to cluster over time and to increase the intensity of jumps in other assets (mutual excitation). We derive the asymptotic behavior of these mutual excitation tests and assess their finite sample performance via simulations. Applying the tests to high-frequency data, we study episodes of financial flights (i.e., risk-on-risk-off trades) in the form of FTS and FTQ in U.S. markets.

We propose a new nonparametric test for identifying common jump arrivals in high-frequency data. Our econometric procedure is particularly related to the multivariate tests of Bollerslev et al. (2008), Jacod and Todorov (2009), Mancini and Gobbi (2012) and Bibinger and Winkelmann (2015). Among these studies, Bollerslev et al. (2008) identify cojumps by utilizing cross-covariance of asset returns. Jacod and Todorov (2009) extend the test of Bollerslev et al. (2008) to a generalized framework based on two possible hypotheses (common jumps or disjoint jumps), while Mancini and Gobbi (2012) is the first to develop the formal truncation method to pin down cojumps. Building on Jacod and Todorov (2009) and Mancini and Gobbi (2012), the cojump analysis of Bibinger and Winkelmann (2015) accounts for non-synchronous trading and market microstructure noise in the data. Perhaps surprisingly, these studies consider jumps as Lévy-type (i.e., either Poisson or pure Lévy), and hence, cojumps occur randomly (in theory) with no memory. Moreover, the statistical problem is typically assumed to be symmetric in the literature, that is, jumps arrive simultaneously and the origins (or at least direction) of shocks may be negligible. We relax these assumptions and minimize the restrictions on jump dynamics.

In a rather tractable specification, we allow the jumps in one asset to increase the jump intensity of the other asset, not necessarily contemporaneously but with lags. This is achieved by extending the univariate (self-excitation) approach of Boswijk et al. (2017). Cojumps are directional events in our identification. In particular we identify three cases of interest between two assets. Case I

\[\text{While our objective is to develop a bivariate cojump test, one can detect the arrival times of common jumps through the univariate jump tests of Lee and Mykland (2008) and Lee and Hannig (2010). Lee and Mykland (2008) test for the presence of big (Poisson-type) jumps, whereas the latter study permits to track both small (i.e. Lévy-type) and big (Poisson-type) jumps in high frequency data. See also Lee and Mykland (2012), who propose a test that separates jumps in prices from market microstructure noise.}\]

Empirically, several studies link cojumps to news announcements. For example, Lahaye et al. (2011) consider various asset classes and find evidence that macro news creates cojumps. Dungey and Hvozdyk (2012) focus on the U.S. Treasury bond market and study cojumps between spot prices and futures. Bibinger et al. (2015) use ECB monetary policy surprises to explain cojump patterns in interest rates. Departing from these studies, Bandi and Renò (2016) identify the periods of price and volatility cojumps: large price jumps are strongly associated with turbulent periods (e.g., Black Monday, the Asian crisis, the Russian crisis) and amplified by jumps in volatility.
represents the case where a negative jump in the risky asset increases the (positive) jump intensity of the safe-haven asset. This case represents the FTS and FTQ episodes. **Case II** represents the potential for the reverse to occur, that is where a jump in the safe-haven asset leads to flight to the riskier asset—we denote this as the *seeking returns strategy* (SRS). Finally, **Case III** characterizes the periods when FTS or FTQ occurs in the face of good news, the safe-haven asset jumps in response to a positive jump in the risky asset. Our approach identifies the jump propagation in space, whereas Boswijk et al. (2017) examine jump propagation in time. We show that the asymptotic behavior of the test statistics in Boswijk et al. (2017) remains fairly similar to that of self excitation tests.

[Insert Figure 2 about here]

In particular, we study the dynamics of financial flights associated with *mutually exciting* jump-type tail events. Researchers often define *flights* as periods of risk-off trades, when market participants flee relatively risky investments (such as stocks) and invest in safe assets such as gold (i.e., flight to safety (FTS)) or assets of higher quality such as Treasury bonds (i.e., flight-to-quality (FTQ)). As widely reported in the media, these episodes typically occur when market conditions abruptly deteriorate due to certain political or news events.\(^7\) Figure 2 illustrates an example of an intraday financial flight that occurred on August 5-6, 2014, when the tension in the Middle East and Ukraine-Russia conflicts increased significantly. In the figure, the Bloomberg trading screen exhibits a sharp sudden decline in the index—around 19:30 GMT—as gold prices surged rapidly.

Studying financial flights is of interest not only for understanding the behavior of investors but also in terms of measuring risk transmission (and premia) within the global financial system. Caballero and Krishnamurthy (2008) demonstrate theoretically that unusual and unexpected events lead to FTQ trades, which is empirically confirmed in Baele et al. (2014) through international FTS. In U.S. markets, Goyenko and Sarkissian (2015) consider U.S. macroeconomic shocks to be drivers of FTQ while Engle et al. (2012) focus on the market microstructure and examine the dynamics of the U.S. Treasury market around FTS periods.\(^8\) Unlike Baele et al. (2014) and Goyenko and Sarkissian

\(^7\)For instance, Bloomberg reported the following headline on July 10, 2014 at 4:03 pm–EST. “Stocks from U.S. to Europe slid as increasing concern over signs of financial stress in Portugal sent investors seeking safety in Treasuries, the yen and gold.”

\(^8\)In this paper, we restrict our attention to the dynamics of FTS and FTQ. It is worth noting that financial flights
(2015), we characterize financial flights from intraday data and apply our empirical analysis to the S&P500 index, long term U.S. bonds and gold as sources of safety and quality, respectively. The identification approach in this paper is non-parametric and flexible enough to capture the different strengths of shocks and time variation in FTS/FTQ regimes. Like Bae et al. (2003), we consider the possibility that severe shocks may propagate differently from small shocks. However, while Bae et al. (2003) utilize extreme value theory based on low-frequency return distribution, we adopt a nonparametric approach relying on continuous time. Our empirical analysis reveals that FTS periods (from stocks to gold) occur more frequently than FTQ cycles (from stocks to bonds). These identified risk-off trades take place mostly in times of turbulence, such as the liquidity case of BNP Paribas, the bankruptcy of Lehman Brothers, the rescue of AIG, the implementation of the Trouble Assets Relief Program, and the intensification of the European debt crisis (2011-2012). In addition, we explore seeking returns strategies (SRS) used by investors when risk appetite increases in the marketplace. We find evidence of such SRS patterns in the data, including periods where FTS, FTQ and SRS occur within the same period. The results suggest that SRS episodes exhibit asymmetry, likely reflecting the fact that as market conditions improve, traders prefer to sell gold, rather than bonds, to invest in stocks.

Our paper extends Aït-Sahalia et al. (2015) and Aït-Sahalia et al. (2014) in several respects. In contrast to the parametric model specifications of these studies, our approach is nonparametric and relies on test statistics that are model-free. This feature allows us to be flexible in characterizing the (bivariate) log-price process: the coefficients of the diffusion terms, volatility components, jump sizes, and jump intensity process can be in any (semimartingale) form. We derive the asymptotic behavior (level and power) of the test statistics and consider different null hypotheses to identify the periods of mutual excitation in the data. A particularly attractive feature is that—unlike the could also come in other forms, such as flight to liquidity (FTL) trades. Vayanos (2004) develops an equilibrium model and shows that risk aversion increases during periods of turmoil, and investors experience a sudden and strong preference for holding bonds. Liquidity premium is time-varying and increases with volatility. Extending this approach, Brunnermeier and Pedersen (2009) theoretically link market liquidity to FTQ. Beber et al. (2009) analyze European bond market and provide evidence that investors demand credit liquidity—rather than quality—in times of crisis: FTL dominates FTQ. From an empirical perspective, Longstaff (2004) studies the risk premium implications of FTL in the U.S. Treasury bond market. Panchenko and Wu (2009), Baele et al. (2010) and Bansal et al. (2014) explore the mechanics and determinants of stock-bond interactions. For other specifications to test for market linkages and stock-bond comovements, see, e.g., Andersen et al. (2007b), Forbes and Rigobon (2002), Hartmann et al. (2004), Chordia et al. (2005), Connolly et al. (2005), Kim et al. (2006) and Karolyi and Stulz (1996).
specification of Aït-Sahalia et al. (2015)—we retain a view of contagion as being entirely unexpected and hence unpriced as consistent with the lower frequency literature; Xiu (2014) demonstrates the role of contagion in option pricing in the Aït-Sahalia et al. (2015) self-exciting Hawkes model. Aït-Sahalia et al. (2015) and Aït-Sahalia et al. (2014) consider a finite jump activity for the underlying Hawkes-type return variation. We relax this assumption in our methodological setup: jumps can exhibit infinite activity with mutual excitation. We can thus generalize the dynamics of jump-type financial shocks in a rather comprehensive way. Shocks can be large (and rare) or small (and frequent). Furthermore, we vary the jump size (i.e., the excitation threshold) and, hence, capture different levels of jump propagation in the form of weak, mild or severe transmissions.

The remainder of this paper is organized as follows. Section 2 introduces the formal setup, the base methodology and testing procedures. In Section 3, we present the Monte Carlo simulations to assess the finite sample performance of the tests. Section 4 describes the data and reports our empirical results. We check the robustness of our results in Section 5. Section 6 concludes.

2. Methodology

2.1. Excitation dynamics of jump-type shocks

Our objective is to identify the transmission of shocks from one asset to another. Before introducing our formal setup, we first illustrate the excitation dynamics as in Aït-Sahalia et al. (2015) and Aït-Sahalia et al. (2014). Consider two assets \((d = 1, 2)\) and let the log-price of each asset follow a semimartingale Itô:

\[
dX_{d,t} = b_{d,t} dt + \sigma_{d,t} dW_{d,t} + \xi_{d,t} dN_{d,t}, \quad d = 1, 2, \tag{1}
\]

where \(b_{d,t}\) is the drift term, \(\sigma_{d,t}\) is the stochastic volatility component, and \(W_{d,t}\) denotes a standard Brownian motion. It is possible to consider that \(W_{1,t}\) and \(W_{2,t}\) are correlated. In our setup, we let the Brownian components of prices to be independent with some potential correlation between spot volatility of assets (i.e., \(\sigma_{1,t}\) and \(\sigma_{2,t}\)).

---

9 Asymptotically, we can identify the transmission of all jumps either large or small at the researcher’s discretion. In finite samples, however, it is more relevant to focus more on those large jumps (triggering high risk exposure) because small jumps are often subject to larger identification errors. Consequently, one may not be able to precisely identify (the location and arrival) of each small jump from the diffusion component of log-price process. Throughout our study, we thus focus on the examination of the transmission of the large jumps.

10 It is possible to consider that \(W_{1,t}\) and \(W_{2,t}\) are correlated. In our setup, we let the Brownian components of prices to be independent with some potential correlation between spot volatility of assets (i.e., \(\sigma_{1,t}\) and \(\sigma_{2,t}\)).

---
each asset $d = 1, 2$. The intensity of Hawkes-type jumps follows the dynamics

$$d\lambda_{d,t} = \check{\alpha}_d(\lambda_{d,\infty} - \lambda_{d,t})dt + \sum_{l=1}^{2} \check{\beta}_{d,l}dN_{l,t}, \quad d, l = 1, 2,$$

(2)

with conditions $\check{\alpha}_d > \check{\beta}_{d,l} > 0$ and $\lambda_{d,\infty} > 0$ for $d, l = 1, 2$.\footnote{These conditions ensure that intensities of shocks follow stationary Markov processes.} We can rewrite (2) as

$$\lambda_{1,t} = \lambda_{1,\infty} + \int_{-\infty}^{t} \check{\beta}_{1,1} e^{-\check{\alpha}_1(t-s)}dN_{1,s} + \int_{-\infty}^{t} \check{\beta}_{1,2} e^{-\check{\alpha}_1(t-s)}dN_{2,s},$$

(3)

$$\lambda_{2,t} = \lambda_{2,\infty} + \int_{-\infty}^{t} \check{\beta}_{2,1} e^{-\check{\alpha}_2(t-s)}dN_{1,s} + \int_{-\infty}^{t} \check{\beta}_{2,2} e^{-\check{\alpha}_2(t-s)}dN_{2,s},$$

(4)

where $\lambda_{1,t}$ and $\lambda_{2,t}$ are the shock intensity processes of assets 1 and 2, respectively. The intuition behind the setup in (3)–(4) is the following. Consider the jump dynamics of asset 1 (i.e., Equation (3)). Whenever a sudden shock ($dN_{2,s} = 1$) occurs in asset 2, $\lambda_{1,t}$ jumps with magnitude $\check{\beta}_{1,2}$. The jump intensity $\lambda_{1,t}$ then mean-reverts back towards $\lambda_{1,\infty}$ at speed $\check{\alpha}_1$. This specification delivers two types of feedback mechanism. First, the jump intensity of each asset changes over time and responds to past jumps (via $\check{\beta}_{1,1}$ and $\check{\beta}_{2,2}$). This is the self-excitation of jumps. Second, jump events in one asset propagate across markets and increase the chance of future shocks in other assets (via $\check{\beta}_{1,2}$ and $\check{\beta}_{2,1}$). The latter effect is the mutual excitation that we wish to capture. The next section introduces our modeling framework.

### 2.2. The model setup

Throughout, we fix a probability space $(\Omega, \mathcal{F}_t, \mathbb{P})$ where $\Omega$ is the set of events in financial markets, $\mathcal{F}_t : t \in [0, T]$ is (right-continuous) information filtration for investors, and $\mathbb{P}$ is a data-generating measure. We follow Boswijk et al. (2017) and consider the Grigelionis decomposition of $X_{d,t}$ in (1). That is,

$$X_{d,t} = X_{d,0} + \int_{0}^{t} b_{d,s}ds + \int_{0}^{t} \sigma_{d,s}dW_{d,s} + x_d * (\mu_{d,t} - \nu_{d,t}) + (x_d - h(x_d)) * \mu_{d,t},$$

(5)

Simulation results (unreported for brevity) reveal that the degree of correlation between $\sigma_{1,t}$ and $\sigma_{2,t}$ does not affect the finite sample properties of our mutual excitation tests substantially.
where \( b_d = (b_{d,t}) \) and \( \sigma_d = (\sigma_{d,t}) \) are locally bounded, \( \mu_d \) is the jump measure of \( X_d \) and \( \nu_d \) is its jump compensator that adopts the following decomposition

\[
\nu_{d,t}(dt, dx) = F_{d,t}(x)dt.
\]

We further assume that the predictable random measure \( F_{d,t} \) can be factored into two parts:

\[
F_{d,t}(dx) = f_{d,t}(x)\lambda_{d,t-} \, dx. \tag{6}
\]

Here, the predictable function \( f_{d,t}(x) \) controls the jump size distribution and \( \lambda_{d,-} = (\lambda_{d,t-})^{12} \) is the stochastic jump intensity or stochastic scale with

\[
\lambda_{d,t} = \lambda_{d,0} + \int_0^t b'_{d,s}ds + \int_0^t \sigma'_{d,s}dW_{d,s} + \int_0^t \sigma''_{d,s}dB_{d,s} + \delta_{d,1} * \mu_{1,t} + \delta_{d,2} * \mu_{2,t} + \delta'' * \mu_{d,t}, \tag{7}
\]

where \( B \) is a standard Brownian motion independent of \( W \), \( \mu_{d,t}^+ \) is orthogonal to \( \mu_1 \) and \( \mu_2 \), and \( \delta_{d,1}, \delta_{d,2}, \delta'' \) are predictable.\(^{13} \) Boswijk et al. (2017) identify self-excitation through the common jumps between a log-price process \( X \) and its own jump intensity \( \lambda \).\(^{14} \) We extend their framework to test for mutual excitation in jumps across assets. We consider a function of \( X_1 \) and \( \lambda_2 \) as follows.

\[
U(H)_t = \sum_{0 \leq s \leq t} H(X_{1,s-}, X_{1,s+}, \lambda_{2,s-}, \lambda_{2,s+}). \tag{8}
\]

The idea is then to choose a function \( H \) for \( \mathbb{R} \times \mathbb{R} \times \mathbb{R}_{++}^* \times \mathbb{R}_{++}^* \) such that \( U(H)_T \) behaves distinctly when \( X_1 \) and \( \lambda_2 \) (the jump intensity of \( X_2 \)) do or do not cojump within the interval \([0, T]\). For

\(^{12} \)For each \( t > 0 \), \( \lambda_{d,t-} := \lim_{s \uparrow t} \lambda_{d,t} \). Similarly, \( \lambda_{d,t+} := \lim_{s \downarrow t} \lambda_{d,t} \).

\(^{13} \)The compensator of \( \xi_{d,t}dN_t \) is given by \( f^\xi_{d,t}(x)\lambda_{d,t-}dsdt \), where \( f^\xi_{d,t} \) is the jump size density function of \( \xi_{d,t} \). Given that, it has the same form of \( \nu_{d,t}(dt, dx) = dt \otimes F_{d,t}(x) \) with \( F_{d,t} \) given by (6). In Equation (1), we assume that \( f^\xi_{d,t} \) is the density function of finite activity jumps, while we will later assume that the function \( f_{d,t} \) in (6) is that of infinite activity jumps.

\(^{14} \)See Aït-Sahalia and Hurd (2016) for an alternative model of optimal portfolio selection when jumps are mutually exciting. While Li et al. (2016, 2014) use jump regressions to identify jump size dependence (based on constant intensity), we focus on the interaction between jumps and time-varying jump intensity process. In this direction, Corradi et al. (2014) propose a self-excitement test for constant jump intensity in asset returns.
instance, one may choose a function \( H \) such that

\[
H(x_1, x_2, y_1, y_2) = 0 \iff x_1 = x_2 \text{ or } y_1 = y_2,
\]

(9)

or

\[
H(x_1, x_2, y_1, y_2) \neq 0 \iff x_1 \neq x_2 \text{ and } y_1 < y_2,
\]

(10)

or even more specifically, one can consider

\[
H(x_1, x_2, y_1, y_2) = \begin{cases} 
> 0 & \text{if } x_1 \neq x_2 \text{ and } y_1 < y_2, \\
= 0 & \text{if } x_1 = x_2 \text{ or } y_1 = y_2, \\
< 0 & \text{if } x_1 \neq x_2 \text{ and } y_1 > y_2.
\end{cases}
\]

(11)

**Remark 1.** If at a time point \( t \), we have \( |\Delta X_{1,t}| = 0 \), then no matter which above condition \( H \) satisfies, the value of \( H \) at this time point is zero. If \( |\Delta X_{1,t}| \neq 0 \), then \( H \) will take different values according to whether \( \Delta \lambda_{2,t} \) is equal to, greater than, or smaller than zero.\(^{15}\)

**Remark 2.** A complication may arise when \( |\Delta X_{1,t}| |\Delta \lambda_{2,t}| > 0 \) and \( |\Delta X_{2,t}| \neq 0 \). That is, the positive jump in \( \lambda_{2,t} \) may come from the jump in \( X_{1,t} \), \( X_{2,t} \) or both. In this case, it is not possible to disentangle the effect of \( X_{1,t} \) and that of \( X_{2,t} \). However, many common jumps indicate that other driving forces might trigger these common jumps.\(^{16}\)

One challenge here is that the jump intensity process is not observable. To determine the value of \( H \) at each jump time of \( X_1 \), we need to estimate the spot values of \( \lambda_2 \) before and after this jump time. To achieve this goal and define mutual excitation in jump-type shocks, we make the following assumption for \( F_{d,t} \) (in (6)).

**Assumption 1.** Suppose the drift and volatility processes \( b_{d,t} \) and \( \sigma_{d,t} \) \((d = 1, 2)\) are locally bounded. Assume that there are three (nonrandom) numbers \( \beta_d \in (0, 2) \), \( \beta_d' \in [0, \beta_d) \) and \( \gamma > 0 \),

\(^{15}\)In practice, we believe that it is quite unlikely to have \( \Delta \lambda_{2,t} < 0 \) when \( |\Delta X_{1,t}| \neq 0 \). In other words, the likelihood that a jump in \( X_1 \) will decrease the intensity of \( X_2 \) is very small because price jumps rarely stabilize financial market conditions. Also note that the sign of the function \( H \) (in Equations (9)-(11)) is zero (positive), when \( \Delta \lambda_{2,t} = 0 \) (\( \Delta \lambda_{2,t} > 0 \)). If \( \Delta \lambda_{2,t} < 0 \), the sign of \( H \) in (11) becomes negative, whereas the sign of \( H \) in (9) remains positive.

\(^{16}\)It is worth noting that we are not making causal inference here. We are particularly interested in tracing whether or not the jumps in \( X_1 \) will be accompanied by a positive jump in the intensity of \( X_2 \).
and a locally bounded process \( L_t \geq 1 \), such that, for all \((\omega, t)\),

\[
F_{d,t} = F'_{d,t} + F''_{d,t},
\]

where

(a) \( F'_{d,t}(dx) = f_{d,t}(x)\lambda_{d,t} dx \) with \( \lambda = (\lambda_{d,t}) \) given by (7), \( \lambda_{d,t} \leq L_t \) and

\[
f_{d,t}(x) = \frac{1 + |x|^\gamma h_d(t, x)}{|x|^{1+\beta}},
\]

for some predictable function \( h_d(t, x) \), satisfying

\[
1 + |x|^\gamma h_d(t, x) \geq 0, \quad |h_d(t, x)| \leq L_t.
\]

(b) \( F''_{d,t} \) is a measure that is singular with respect to \( F'_{d,t} \) and satisfies

\[
\int_{\mathbb{R}} (|x|^\beta \wedge 1) F''_{d,t}(dx) \leq L_t.
\]

Given these dynamics, we assume that the log-price process \( X_{d,t} \) in (1) is observed at discrete points in time. For each asset \((d = 1, 2)\), the continuously compounded \( i \)-th intraday return of a trading day \( t^{\text{day}} \) is given by \( r_{i, t^{\text{day}}} \equiv X_{t+i\Delta_n} - X_{t+(i-1)\Delta_n} \), with \( i = 1, \ldots, M \) and trading days \( t^{\text{day}} = 1, \ldots, T^{\text{day}} \). Let \( M \equiv \lfloor 1/\Delta_n \rfloor \) denote the number of intraday observations in one day. \( \Delta_n = 1/M \) is then the time between consecutive observations, the inverse of the observation frequency. To characterize mutual excitation, we introduce a testing procedure that relies on certain functional forms and jump intensity estimators. Extending the methodological framework of Boswijk et al. (2017), we begin by defining these measures.\(^{18}\)

**Definition 1.** Let \( U(H)_t \) denote the functional of \( X_t \) in (1). The estimator of \( U(H)_t \) is then given by

\[
U(H, k_n)_{t}^{12} = \sum_{i=k_n+1}^{[t/\Delta_n]-k_n} H(X_{1,i-1}, X_{1,i}, \tilde{\lambda}(k_n)_{2,i-}, \tilde{\lambda}(k_n)_{2,i})1_{\{\Delta^n_{\infty}|X_1|>\alpha\Delta^n_{\infty}\}},
\]

where \( k_n \) is an integer, \( \Delta^n_{\infty} \) is the time between consecutive observations for a constant \((0 < \omega < 1/2)\) (see e.g., Aït-Sahalia and Jacod, 2012) and \( \alpha\Delta^n_{\infty} \) is the truncation threshold indexed by \( \alpha \) (> 0) number of standard deviations of the continuous part of \( X_t \) in (1).

\(^{17}\)For each trading day \( t^{\text{day}} \), \( r_{i, t^{\text{day}}} \) is hence the \( i \)-th intraday return between consecutive price observations. For simplicity and ease of notation, we drop the superscript \( \text{day} \) in \( r \) and instead use \( r_{i,t} \).

\(^{18}\)In Assumption 1, \( L_t \) gives an upper bound of \( \lambda_{d,t} \) and \( |h_d(t; x)| \). The condition \( L_t \geq 1 \) means that this upper bound is not close to zero. This value can be replaced by any given positive number. At least for our purpose, whether or not \( |h_d(t; x)| \) is larger than one does not matter substantially. We adopt the condition \( L_t \geq 1 \) from Aït-Sahalia and Jacod (2009). We refer our readers to this study for more details on this issue.
In the functional form (16), $\hat{\lambda}(k_n)$ is the estimator of the (spot) jump intensity $\lambda_t$ given by

$$
\hat{\lambda}(k_n)_{2,i} = \frac{\Delta_n^{\omega \beta_2}}{k_n \Delta_n} \sum_{j=i+1}^{i+k_n} g \left( \frac{|\Delta_n^{\alpha} X_2|}{\alpha \Delta_n^{\omega}} \right) \frac{\alpha^{\beta_2}}{C_{\beta_2}(1)},
$$

(17)

where $\beta_2 \in (0, 2]$ is the jump activity index, $\Delta_n^\alpha X_2$ is the intraday return of asset 2, and $k_n$ satisfies $(1/K \leq k_n \Delta_n^\beta \leq K)$ for $(0 < \rho < 1)$ and $(0 < K < \infty)$. In (17), $g(\cdot)$ is an auxiliary function that disentangles jumps from the diffusion component. As in Jing et al. (2012), we assume that this function $g(\cdot)$ satisfies

$$
g(x) = \begin{cases} 
|x|^p & \text{if } |x| \leq 1, \\
1 & \text{if } |x| > 1,
\end{cases}
$$

for an even integer $p > 2$ and $(x := |\Delta_n^\alpha X|/\alpha \Delta_n^{\omega})$. For the intensity estimator (17), the quantity $C_{\beta}(1) = 1$ for $g = 1_{\{x>1\}}$ and its general form can be given by

$$
C_{\beta_2}(k_n) = \int_0^\infty (g(x))^{k_n} / x^{1+\beta_2} dx.
$$

In sum, the estimates for the functional form $U(H)$ and jump intensity $\lambda_t$ allow us to construct a test statistic to identify the mutual excitation across financial assets. This implies that we first need to estimate the spot jump intensities (via Equation (17)) and then use the estimates for calculating the measure in Equation (16). Armed with this setup, we present the main hypotheses and the corresponding testing procedures in the next section.

2.3. Hypotheses and testing procedures

Our identification approach for market linkages relies on mutual excitation dynamics. To characterize financial flights and market shock dependence, the first step is to test for the presence of cross-excitation in jumps. Let $X_1$ and $X_2$ denote asset 1 and asset 2, respectively. We examine whether jumps in asset $X_1$ ($X_2$) excite the jump occurrences in asset $X_2$ ($X_1$). To proceed, let $\omega$
denote a specific outcome, i.e., $\omega \in \Omega$. Then, we can state our testable hypotheses formally as

- The jumps of $X_1$ excite the jump intensity of $X_2$:
  $$\sum_{0 \leq t \leq T} \Delta \lambda_{2,t}(\omega) 1_{\{\Delta \lambda_{2,t}(\omega) \geq 0, \Delta X_{1,t}(\omega) \neq 0\}} > 0.$$ 

- $X_1$ has no common jumps with the jump intensity of $X_2$:
  $$\sum_{0 \leq t \leq T} \Delta \lambda_{2,t}(\omega) 1_{\{\Delta X_{1,t}(\omega) \neq 0\}} > 0.$$ 

- The jumps of $X_1$ decrease the jump intensity of $X_2$:
  $$\sum_{0 \leq t \leq T} \Delta \lambda_{2,t}(\omega) 1_{\{\Delta \lambda_{2,t}(\omega) \leq 0, \Delta X_{1,t}(\omega) \neq 0\}} > 0,$$

or, perhaps more intuitively, in the following forms

(A) $H_0$: No jump excitation from $X_1$ to $X_2$ vs. $H_1$: Jump excitation from $X_1$ to $X_2$

(B) $H_0$: No jump excitation from $X_2$ to $X_1$ vs. $H_1$: Jump excitation from $X_2$ to $X_1$

(C) $H_0$: Jump excitation from $X_1$ to $X_2$ vs. $H_1$: No jump excitation from $X_1$ to $X_2$

(D) $H_0$: Jump excitation from $X_2$ to $X_1$ vs. $H_1$: No jump excitation from $X_2$ to $X_1$

These hypotheses allow us to investigate not only the jump cascades across regions but also the direction of shocks transmission. Methodologically, we need to make an inference about $\omega$, given a discretely observed sample path over $[0, T]$. The outcome could belong to the “no excitation set” ($\omega \in \Omega_T^{\text{no}}$) or to the “excitation set” ($\omega \in \Omega_T^{\text{mut}}$). Specifically, We let $\mathcal{E} = \mathbb{R}$ in general, $\mathcal{E} = (\epsilon, \infty)$ and $\mathcal{E} = (-\infty, -\epsilon)$. We have

$$\Omega_T^{\mathcal{E},+} = \left\{ \omega \in \Omega_T^\mathcal{E} : \sum_{0 \leq t \leq T} \Delta \lambda_t(\omega) 1_{\{\Delta \lambda_t(\omega) \geq 0, \Delta X_t(\omega) \in \mathcal{E}\setminus\{0\}\}} > 0 \right\},$$

$$\Omega_T^{\mathcal{E},-} = \left\{ \omega \in \Omega_T^\mathcal{E} : \sum_{0 \leq t \leq T} \Delta \lambda_t(\omega) 1_{\{\Delta \lambda_t(\omega) \leq 0, \Delta X_t(\omega) \in \mathcal{E}\setminus\{0\}\}} < 0 \right\},$$

$$\Omega_T^{\mathcal{E},0} = \left\{ \omega \in \Omega_T^\mathcal{E} : \sum_{0 \leq t \leq T} |\Delta \lambda_t(\omega)| 1_{\{\Delta X_t(\omega) \in \mathcal{E}\setminus\{0\}\}} = 0 \right\}.$$
Then $\Omega_{T}^{(\text{mut})} = \Omega_{T}^{+}$ and $\Omega_{T}^{(\text{no})} = \Omega_{T}^{-} \cup \Omega_{T}^{0}$. We test (A)–(D) by comparing a test statistic to its probability limit under the alternative hypotheses. We first consider the null hypotheses of “no mutual excitation” (i.e., (A) and (B)).

**Theorem 1.** Let $\omega \in (0, 1/2)$. Suppose that Assumptions 1, 2, 3 (see Appendix A) hold with $p > \frac{1 - \omega \beta_{2}}{1/2 - \omega}$, $q \geq 2$ and $q' \geq 1$, and

$$1 - \omega \beta_{2} < \rho < (1 - \omega \beta_{2}) + 2\phi' \wedge 2\phi'' \wedge \frac{1}{2} \omega \beta_{2},$$

(18)

In addition, if either one of the following conditions is satisfied,

(a) $H(x_1, x_2, y_1, y_2) = 0 \text{ for } |x_1 - x_2| \leq \epsilon$, where $\epsilon > 0$;

(b) $q' > 2$, $2\omega \beta_{1} < 1$ and $\rho \leq 1 - \omega \beta_{2} + (q_{c}' - 1 + \omega (2q_{d}' - \beta_{1})) \wedge 2\omega (q_{c}' + q_{d}' - \beta_{1})$.

then, for any fixed $t > 0$, under $H_{0}$ of (A), we have

$$t_{n}^{12} := \sqrt{\frac{k_{n} \Delta_{n}}{\Delta_{n}^{\omega \beta}}} \frac{U(H, k_{n})^{12}_{T}}{U(G, k_{n})^{12}_{T}} \xrightarrow{\mathcal{L}_{st.}} \mathcal{N}(0, 1),$$

(19)

where $U(G, k_{n})_{T}$ is the consistent estimator of the conditional variance with

$$G(x_1, x_2, y_1, y_2) = \frac{\alpha \beta C_{\beta}(2)}{(C_{\beta}(1))^{2}} (y_1 H_{3}'(x_1, x_2, y_1, y_2)^{2} + y_2 H_{4}'(x_1, x_2, y_1, y_2)^{2}),$$

(20)

where $H_{3}'$ and $H_{4}'$ stand for the first partial derivatives of the function $H(\cdot)$ with respect to its 3rd and 4th arguments, respectively. For testing this null hypothesis, we employed $H(p, 1)$ defined in (29). However, under $H_{1}$ of (A), we have $|t_{n}^{12}| \xrightarrow{p} \infty$.

Therefore, the following critical region has an asymptotic level $\alpha$ for testing the null hypothesis of “no mutual excitation” (i.e., $\Omega_{T}^{(\text{no})}$), and asymptotic power 1 for the alternative (i.e., $\Omega_{T}^{(\text{mut})}$):

$$C_{n}^{\text{no}} = \{t_{n}^{12} > z_{\alpha}\},$$

(21)

where $\mathbb{P}(W > z_{\alpha}) = \alpha$ and $W$ is a standard normal random variable.

Under the null hypothesis of “mutual excitation” (i.e., (C) and (D)), the asymptotic behavior of the test statistic is as follows.

**Theorem 2.** Assume $P(\Omega_{T}^{(-)}) = 0$ and that the same assumptions as in Theorem 1 hold, but with Assumption 2 replaced by Assumption 3 (see Appendix A). In addition, assume $H$ satisfies the following degenerate condition:

$$y_1 = y_2 \implies \|H_{3}'(x_1, x_2, y_1, y_2)\| + \|H_{4}'(x_1, x_2, y_1, y_2)\| = 0.$$  

(22)

If either condition (a) in Theorem 1 or the following (b')

$$q' > 2, 2\omega \beta_{1} < 1, \rho \leq 1 - \omega \beta_{2} + ((q_{c}' - 1 + \omega (2q_{d}' - \beta_{1}))/2) \wedge 2\omega (q_{c}' + q_{d}' - \beta_{1})$$
is satisfied, then the following critical region has an asymptotic level $\alpha$ for testing the null hypothesis $\Omega_{T}^{\text{mut}}$:

$$C_n^{\text{mut}} = \{|R_n| > z_{\alpha} \sqrt{V_n}\},$$

(23)

where

$$R_n = \frac{U(H, wk_n) - U(H, k_n)}{U(H, k_n)}$$

and

$$V_n = \frac{\Delta_n^{\rho}}{k_n} \left( \frac{m - 1}{w} U(G, k_n) \right)^2.$$

(24)

Moreover, choose a sequence of positive numbers $v_n$ such that

$$v_n \to 0, \quad \text{and} \quad \frac{k_n v_n \Delta_n}{\Delta_n^{\rho}} \to \infty,$$

and set $V'_n = V_n \wedge v_n$. Then, the critical region

$$C_n^{\text{mut}} = \{|R_n| > z_{\alpha} \sqrt{V'_n}\}$$

(25)

has an asymptotic level $\alpha$ for testing the null hypothesis $\Omega_{T}^{\text{mut}}$ and asymptotic power 1 for the alternative $\Omega_T^{(0)}$.

Therefore, we have two statistics to test for mutual excitation between jumps. Under the null hypothesis of no mutual excitation (i.e., (A) and (B)), we have the statistic in (19) with $H(p, 1)$ given in (29), and under the null hypothesis of excitation (i.e., (C) and (D)), we use the statistic (24) with $H(p, 2)$ given in (29). $H(p, 2)$ satisfies the degenerate condition (22).

3. Monte Carlo study

Having presented our testing procedures and hypotheses, we now assess the finite sample properties of the test statistics. Throughout this section, we consider an observation length of one week and a sampling frequency of 5 seconds. We thus set $T = 5/252$ and $M = 23400$, giving $\Delta_n = 1/23400$. Following Aït-Sahalia and Jacod (2009) and Jing et al. (2012), we further choose the values $\varpi = 1/3$, $\rho = 0.6$ and $\beta_2 = 1.25$. We conduct each simulation with 1000 replications.\textsuperscript{21}

3.1. Simulation setup

Our simulation setup is based on a bivariate Hawkes process with mutually exciting jumps. As in Aït-Sahalia et al. (2015) and Boswijk et al. (2017), we consider the following data-generating

\textsuperscript{21}It is worth mentioning that our test statistics are constructed over a finite time span. Hence, we are particularly detecting realized mutual excitation in jumps. In other words, if we do not find mutual excitation on $[0, T]$, this does not necessarily imply that the probability of mutual excitation is zero.
process:

\[
\begin{aligned}
    dX_{1,t} &= \sigma_{1,t}dW_{1,t} + \lambda_{2,\infty}dY_{1,t} \\
    dX_{2,t} &= \sigma_{2,t}dW_{2,t} + \lambda_{2,t-}dY_{2,t} \\
    d\sigma_{1,t}^2 &= d\sigma_{2,t}^2 = \kappa(\theta_1 - \sigma_{1,t}^2) + \eta_1\sigma_{1,t}dB_{1,t} \\
    d\lambda_{2,t} &= \kappa_1(\lambda_{2,\infty} - \lambda_{2,t})dt + \eta_\lambda dB_{2,t} + \xi_1I(\Delta X_{1,t}>\epsilon),
\end{aligned}
\]

where the Brownian motions \((W_{1,t}, W_{2,t}, B_{1,t}, B_{2,t}')\), and the \(\beta\)-stable jump processes \((Y_{1,t}, Y_{2,t})\) are assumed to be independent. We consider \(\mathbb{E}[dW_{1,t}dB_t] = \phi dt\), which allows us to capture a potential leverage effect between prices and volatility dynamics. For the volatility process (third line), we follow Jing et al. (2012) and set \(\kappa = 5\), \(\theta_1 = 1/16\), \(\eta_1 = 0.5\) and \(\phi = -0.5\). For the jump intensity process (fourth line), we set \(\kappa_\lambda = 1400\), \(\eta_\lambda = 200\) and \(\epsilon = 100\sqrt{\theta}\Delta_n^\infty\) as in Boswijk et al. (2017). \(\lambda_{2,\infty}\) further denotes the constant jump intensity, and we calibrate this value to generate pre-specified values of the tail probability for 0.25%. That is,

\[
P(|\lambda_{2,\infty}\Delta_n^\infty Y_1| \geq \alpha\Delta_n^\infty) \approx \frac{2c_{\beta_2}\lambda_{2,\infty}\Delta_n^{\beta_2}}{\beta_2(\alpha\Delta_n^\infty)^{\beta_2}},
\]

where

\[
c_{\beta_2} = \frac{\Gamma(\beta_2 + 1)}{2\pi} \sin\left(\frac{\pi\beta_2}{2}\right).
\]

We set \(\alpha = 5\sqrt{\theta}\), and \(\beta_2\) denotes the jump activity index. With the choice of \(\beta_2 = 1.25\), the calibrated value for \(\lambda_{\infty}\) is approximately 20.\(^{22}\)

The intuition behind our simulation setup is the following. When we set \(\xi = 0\), there is no mutual excitation in jumps from one asset \((X_1)\) to another \((X_2)\). In this case, we can thus consider (i) a null hypothesis of no mutual excitation (i.e., (A)) or (ii) a null hypothesis of mutual excitation (i.e., (C)).\(^{23}\) For each null, we can check the size and power of the testing procedures. Similarly, when we set \(\xi > 0\), jumps can excite as long as the price changes are large enough (i.e., when

\(^{22}\) In Section 5, we check the sensitivity of the test statistics to different \(\beta_2\) values. The results for \(\beta_2 = 1.5\) and \(\beta_2 = 1.75\) are rather similar.

\(^{23}\) The best small sample properties for these combinations of tests are achieved by examining hypotheses (A) and (C). As our simulation will reveal, an examination of (B) and (D) gives relatively weaker evidence in terms of size and power.
$|\Delta_n X_{1,t}| > \epsilon$). To choose the excitation parameter $\xi$, we utilize the following form

$$
P(\{(\lambda_{2,\infty} + \xi)\Delta_n^n Y_1 | \geq \alpha \Delta_n^n \}) \approx \left(\frac{\lambda_\infty + \xi}{\lambda_\infty}\right)^{\beta_2} = (1 + \xi/\lambda_\infty)^{\beta_2}. \tag{28}
$$

For $\xi = 12$, the ratio in (28) is around 1.80. This implies that the mutual excitation effect will increase the tail probability by 80% (in relative terms), which is statistically and economically significant. Given this setup, the test statistic is given by

$$
t_{12}^n := \sqrt{\frac{k_n \Delta_n \cdot U(H, k_n)_{T}}{\Delta_n^{\alpha \beta_2} \sqrt{U(G, k_n)_{T}}}} \begin{cases}
\xrightarrow{p} -\infty & \omega \in \Omega_T^{(-)}, \\
\xrightarrow{-} \mathcal{N}(0,1) & \omega \in \Omega_T^{(0)}, \\
\xrightarrow{p} +\infty & \omega \in \Omega_T^{(+)},
\end{cases}
$$

where the function $H$ takes the following form

$$
H(p, q) = \begin{cases}
g_0(x_2 - x_1) \cdot (y_2 - y_1)^q \cdot 1_{\{|x_2 - x_1| \geq \epsilon\}}, & p = 0; \\
g_{2,p}(x_2 - x_1) \cdot (y_2 - y_1)^q \cdot 1_{\{|x_2 - x_1| \geq \epsilon\}}, & p > 2;
\end{cases} \tag{29}
$$

where $\epsilon > 0$ and—following Boswijk et al. (2017)—we consider two alternative functions for $H(p, q)$: $H(6, 1)$ and $H(0, 1)$.\(^{24}\) The next section reports the finite sample properties of our testing procedures.

### 3.2. Simulation results

To check the power and size properties, we consider four null hypotheses—presented in Section 2.3 ((A) to (D)). We proceed as follows. In the next section, we set the excitation parameter $\xi = 0$ in order to test for no mutual excitation in all directions, that is, either from $X_1$ to $X_2$ or from $X_2$ to $X_1$. In Section 3.2.2, we conduct simulations when $\xi = 12, 50$ and test for the mutual excitation in jumps.\(^{25}\) In the final section, we adjust our simulation setup to test specifically for the financial flights from one asset to another.

\(^{24}\)We follow Jing et al. (2012) and set $p = 6$. Intuitively, the choice of $p$ in (29) reflects the weight for large jumps. The choice of $p = 6$ puts equal weights to large jumps.

\(^{25}\)Before testing for mutual excitation in the data, we first carried out a pretest for the null of no realized jumps on $[0, T]$. Therefore, we apply our testing procedures to only periods having jumps.
3.2.1. Size and power properties: testing for no mutual excitation

We start by considering the null hypothesis (A) such that jumps in asset $X_1$ do not excite the jumps in asset $X_2$. Figure 3 shows the size and power of the tests for no mutual excitation in jumps. The solid line in the upper-left panel of the figure indicates that the size of the test is rather good. That is, when $\xi = 0$ (no excitation), the percentages of rejection, (i.e., $t_{n}^{12} > z_{\alpha}$), are all close to their corresponding nominal level, $\alpha$.

[ Insert Figure 3 about here ]

In a similar way, we further analyze the size of the test under the null hypothesis that $X_2$ jumps do not excite $X_1$ jumps (hypothesis (B)). The upper-right panel of Figure 3 shows that the test also exhibits a well-behaved size in this direction of jump transmission. The Monte Carlo rejections are close to the nominal level of the test at a 5-second sampling frequency.

We can examine the power of the test by considering the same excitation parameter $\xi = 0$. Intuitively, if there is no mutual excitation in jumps (i.e., $\xi = 0$), we should reject the null hypotheses of mutual excitations (corresponding to hypotheses (C) and (D)). The dotted lines in the upper panels of Figure 3 display the power performance. Two results emerge from the panels. First, under the null hypothesis that jumps in $X_1$ excite the jumps in $X_2$ (hypothesis (C)), the power of the test is strong (dotted line). Second, if we consider a null hypothesis of (reverse) mutual excitation from $X_2$ to $X_1$ (hypothesis (D)), then the test still has good power (dotted line in the upper-right panel).\(^{26}\)

3.2.2. Size and power properties: testing for mutual excitation

We now assess the size and power of our tests in the presence of mutual excitation between jumps. We set the excitation parameter as $\xi = 50$, which is a reasonable value implying that mutual excitation (originated in asset $X_1$) will increase the tail event likelihood (in $X_2$) by more than 80%. Figure 4 displays the size and power of the tests for mutual excitation in jumps.

[ Insert Figure 4 about here ]

\(^{26}\)This interpretation requires care, however, since $\xi = 0$ in the DGP does not necessarily imply an excitation from $X_2$ to $X_1$. As an extension, one can set a threshold for the jump intensity process of $X_2$ (by considering another excitation parameter) and examine whether this intensity process excites large jumps in $X_1$.  

19
The upper-left panel of Figure 4 (dotted line) shows that the test has good size under the null hypothesis that jumps in \( X_1 \) excite the jumps in \( X_2 \) (i.e., hypothesis (C)). Put differently, if there are jumps in the data, we do not reject the null hypothesis of jump excitation from \( X_1 \) to \( X_2 \). The percentages of rejection in simulations are fairly close to the corresponding nominal level \( \alpha \). The upper-left panel of Figure 4 (solid-line) further indicates that the power of our test is reasonable. That is, in the presence of mutual excitation, if we consider the null that there is no mutual excitation from \( X_1 \) to \( X_2 \) (hypothesis (A)), the test has fairly high rejection rates (solid line).

With the choice of \( \xi = 50 \) (presence of mutual excitation), we plot in Figure 4 (upper-right panel) the size and power performance of our test for the reverse excitation. We now consider the null hypothesis that jumps in \( X_2 \) do not excite the jumps in \( X_1 \) (hypothesis (B)). The percentages of rejection in the figure barely depict a forty-five-degree (dashed) line, indicating a potential size distortion for high nominal levels.

Lastly, we consider hypothesis ((D)), which states that jumps in \( X_2 \) excite jumps in \( X_1 \). The upper-right panel of Figure 4 shows that the power of the test—only under this null—is relatively strong (dotted line). This result is expected because if the test has good size under the null hypothesis ((C)), then it should be powerful to reject the null hypothesis ((D)). As the lower panels of Figures 3 and 4 indicate, the size and power properties of the testing procedures are fairly robust to different specifications of the functional form \( H(p, q) \)—in Equation (29)—of the test statistic (i.e., \( H(0,1) \) vs. \( H(6,1) \)). Moreover, while sampling at a lower frequency (e.g., 1 minute) deteriorates the power, its impact on size appears to remain limited (see, e.g., the upper panels in Figures 5 and 6).

To properly assess the finite sample performance of the test statistics, we further report the size and power in Tables 1 and 2, respectively, for various sampling frequencies from 5-seconds to 5-minutes. For both size and power, we compute the simulation rejection rates in the absence...
and presence of mutually exciting jumps.\textsuperscript{27} For brevity, we report the test results only for the direction of mutual excitation from $X_1$ to $X_2$. We consider three excitation parameters such that $\xi = 0, 12, 50$, where the tail event likelihood is 80% with the choice of $\xi = 12$.

Under the null of no mutual excitation (i.e., when $\xi = 0$), Panel A of Table 1 shows that the empirical size of the test is fairly good, especially at high frequencies, such as 5-seconds or 10-seconds. As the sampling frequency becomes lower (1-minute or 5-minutes), the rejection rates tend to increase slightly (e.g., 0.07 and 0.12 at the 5% and 10% levels in columns four and five, respectively). Nevertheless, we observe that the size distortion appears to shrink when we choose $H = H(6; 1)$ (rejection rates in second rows) and consider a discontinuous functional form of $g(\cdot)$ given in (17) (rejection rates in the last three columns).

We further investigate the size of our test under the null hypothesis that jumps are mutually exciting. To check this hypothesis, we consider two alternative excitation parameters: $\xi = 12$ and $\xi = 50$. With the choice of $\xi = 12$, Panel B of Table 1 shows that the size distortion at 1% theoretical level is severe with the choice of $H = H(0; 1)$ (column three). This result particularly holds at low sampling frequencies, such as 5-minutes. Similar to the case of no excitation (i.e., Panel A), the use of discontinuous function for $g(\cdot)$ helps reduce once again these high rejection rates (columns six to eight in Panel B). For the sampling frequencies from 30-seconds to 5-minutes, while the test is oversized at 5% with the choice of $H = H(0; 1)$ (0.084, 0.106, 0.128 in column four), the rejection rates in simulations decrease and get closer to 5% theoretical level with the choice of $H = H(6; 1)$ (0.048, 0.045, 0.026 in column seven).

Compared to the case with $\xi = 12$ (Panel B), the size performance of the test seems to be better when we set $\xi = 50$ in simulations (Panel C in Table 1). Contrary to the rejection rates reported in Panel B, we do not observe strong size distortion in Panel C (see columns three to five). The rejection rates of the test range between 0.010 to 0.017, and these rates are fairly close to 1% theoretical level, irrespective of the choice of sampling frequency (column three in Panel C). For

\textsuperscript{27}It is of course important to note that the test statistic changes according to the null hypothesis considered. Specifically, when we assess the size under the null of no mutual excitation, we use the test statistic given by Equation (19) in the main text with the critical region (21). Under the null of mutual excitation, however, we use the test statistic given by Equation (24) with the critical region (25). When we assess the power under the null of no mutual excitation, we use the test statistic given by Equation (24) with the critical region (25). Under the null of mutual excitation, we instead use the test statistic given by Equation (19) in the main text with the critical region (21).
all 1%, 5% and 10% levels considered, the size of the test is particularly good at 5-seconds when we set $H = H(0; 1)$.

We turn to assess the power performance of the mutual excitation tests. We first consider the null hypothesis of no mutual excitation in jumps. Under this null, Panel A of Table 2 reports that the power for rejecting the presence of mutual excitation is relatively high, particularly at 5-seconds sampling frequency. For 1% theoretical level, the rejection rate is around 0.999–1.000, irrespective of the choice of $H$, that is either $H = H(0; 1)$ (columns three to five) or $H = H(6; 1)$ (columns six to eight). For low sampling frequencies, however, the power of test weakens and rejection rates tend to decrease under the null of no mutually exciting jumps.

We then look into the power of the test when there are mutually exciting jumps in the data. As in the size analysis (Table 1), we check the power for two pre-selected excitation parameters: $\xi = 12$ and $\xi = 50$. Panels B and C of Table 2 report the rejection rates for the absence of mutual excitation. The results indicate that the power is low when $\xi = 12$ (Panel B). Nevertheless, as the excitation signal gets stronger (such as $\xi = 50$ in Panel C), the rejection rates rise at all sampling frequencies. With the choice of high sampling frequency (such as 5-seconds and 10-seconds), the power of the test is reasonable (only) at 10% theoretical level.

### 3.2.3. Testing for financial flights

It is important to note that mutual excitation from one asset ($X_1$) to another ($X_2$) does not necessarily imply a “financial flight”. In particular, a flight-to-safety episode occurs when a large negative return (or jump) in one asset (e.g., stocks) coincides with a large positive return (or jump) in the other asset (e.g., bonds). We are now interested in checking the size and power of our tests by simulating financial flights. To do so, we slightly modify our simulation (DGP) setup and consider a case where the jump intensity of asset $X_2$ increases only with negative jumps in asset $X_1$. Figure 7 displays the size and power of our tests for financial flights between two assets. We report the main results in three cases.

[ Insert Figure 7 about here ]

**Case I (flight from $X_1$ to $X_2$).** This is the case of a flight to safety (FTS) or flight-to-quality (FTQ) episode. Specifically, we consider a null hypothesis that negative jumps in asset $X_1$ do not increase
(positive) jump intensity in asset $X_2$. Under this null hypothesis and with the choice of $\xi = 0$, the left panel of Figure 7 shows that the size of the test is quite good (solid line). For the null hypothesis of a flight from $X_1$ to $X_2$, the testing power appears to be very strong (dotted line).

*Case II (flight from $X_2$ to $X_1$).* This case corresponds to the episodes of potential reverse cross-excitation or *seeking returns strategy* (SRS) when we consider $X_1$ and $X_2$ as the risky and safe-haven assets, respectively. The middle panel of Figure 7 displays the simulation results. For the null hypothesis that there is no financial flight from $X_2$ to $X_1$, the size of the test is rather good (solid line). Intuitively, this finding is also consistent with the results under *Case I* such that if there is a flight from $X_1$ to $X_2$, we should not simultaneously observe a (reverse) flight from $X_2$ to $X_1$.

*Case III (financial flight when the news is good).* In practice, FTS or FTQ episodes occur due to negative news events or unexpected surprises. That is, when the bad news hits the market, investors tend to sell off risky assets (such as stocks) and shortly invest in safe assets (such as gold or U.S. Treasury bonds). Nevertheless, if the news is good for market participants, then both assets might exhibit positive jumps. In this case, the intensity of jumps in one asset (i.e., $X_2$) can increase with positive jumps in the other asset (i.e., $X_1$). We denote this as bull-market-flights (BMF).

We now focus on this possibility and test for mutual excitation (or flight) only when positive jumps propagate. The right panel of Figure 7 demonstrates that our testing procedure is fairly capable of capturing such patterns in the simulated data. Under the null hypothesis that only positive jumps are mutually exciting, the test delivers a good power performance (dotted line). Similarly, if we consider a null hypothesis of no positive jump excitation (from $X_1$ to $X_2$), we do not reject this hypothesis (solid line).

[ Insert Table 3 about here ]

For all these three cases, Panels A and B of Table 3 report the size and power of our financial flight tests, respectively. Panel A indicates that the rejection rates in simulations are fairly close to their 1%, 5% and 10% nominal levels. These rejection frequencies remain mostly stable for all cases, irrespective of the choice of functional form $H$ (results in second rows) or $g(\cdot)$ (columns five to seven). Similar to the size performance, Panel B confirms that the power of the excitation
test to identify flights (in any form) is reasonably high. The rejection frequencies—under the null hypothesis of no financial flight—are close to 1.00 in almost all cases. For Cases II and III, we observe slight power loss, however, at 1% level with Monte Carlo rejection rates around 0.976, 0.979 (columns two and five).

4. Empirical analysis

In this section, we study the dynamics of financial flights revealed by our excitation tests. After describing the database, we present and discuss our empirical findings.

4.1. Data

We use data on the S&P 500 stock index, 30-year U.S. Treasury bond futures and gold futures for the periods covering January 1, 2007 to December 31, 2013. The SIRCA Thompson Reuters Tick History database provides transaction prices throughout the trading days for each asset class. We use 5-minute sampling frequency to apply the tests over quarterly (business cycle) horizon.28 Although there are a number of studies examining behavior at frequencies of 30-seconds and higher, the difficulty in choosing a sampling frequency sufficient for multiple asset types has led to typically lower frequency applications in the literature (see e.g., Andersen et al., 2007a; Dungey et al., 2009 and Lahaye et al., 2011, who use 5-minute and 15-minute data, respectively). Nevertheless, we examine in Section 5.4 the robustness of our baseline results to the choice of sampling frequency and horizon. As we discuss later, the empirical properties of excitation patterns and financial flights remain qualitatively the same at high sampling frequencies (such as 5-seconds) and short time horizons (e.g., weekly rather than quarterly).

As is typical in the literature, we omit trading days with too many missing values or low trading activity.29 Similarly, we delete weekends, certain fixed/irregular holidays, empty intervals and consecutive prices.30 We further adjust the financial market data according to daylight savings

---

28 The raw bond futures and stock index datasets include all open-close, high-low prices. In our empirical analysis, we use closing prices.

29 See, e.g., Dewachter et al. (2014), Lahaye et al. (2011).

30 These holidays include the New Year (December 31 - January 2), Martin Luther King Day, Washington’s Birthday or Presidents’ Day, Good Friday, Easter Monday, Memorial Day, Independence Day, Labor Day, Thanksgiving Day and Christmas (December 24 - 26).
time, considering the (trading) time zones of the assets. Appendix B presents the details of the data descriptions and our adjustment procedures.

4.2. Flight to safety episodes: mutual excitation from the U.S. stock market to the gold market

We begin by identifying the Case I episodes of flight to safety (FTS) occurring between S&P500 index and gold prices. Specifically, we capture FTS patterns by testing for the mutual excitation from (negative) jumps in the market index to (positive) jumps in gold prices.

[ Insert Table 4 about here ]

We study financial flights depending on the strength of tail shocks: weak, mild and severe as in Baele et al. (2014). To do this, we control the FTS levels by considering different jump sizes. We also account for the impact of stochastic volatility by choosing threshold values when testing for flights. While lower thresholds tend to increase jump activity or frequency (see, e.g., Todorov and Tauchen, 2010), high threshold values allow us to analyze propagation between large/infrequent jumps.

Based on these criteria, Table 4 reports the frequency (in %) of FTS episodes falling into the categories of mutual excitation, no excitation, and the two contradictory possibilities (i) where both nulls are rejected and (ii) where both nulls are accepted.\textsuperscript{31} We find evidence for mutual excitation with $\alpha = 2$ between jumps in the S&P500 and gold prices (fourth column). The test detects that up to 20–25% of all common jump arrivals are in the form of financial flights from stocks to gold (first panel in the fourth column). The frequency of FTS trades does not increase monotonically with jump size. When financial shocks are weak (corresponding to size 0.001), FTS variation in the data is only around 3%. The FTS activity surges substantially to 25% when jump shocks hitting the markets are relatively stronger, and drops somewhat to around 20% with the more severe jumps. The most evident flight to safety events occur with mild sized jumps. This result aligns with the conclusion of Boswijk et al. (2017), in the sense that smaller (jump) events are less likely to trigger mutual excitation and cause investors to sell off risky assets and invest in safe havens. However,\textsuperscript{31}

\textsuperscript{31}We consider four categories because we have two different tests with opposite null hypotheses. Therefore, the testing results naturally fall into four categories. See e.g., Jacod and Todorov (2009), who adopt a similar categorization when testing cojumps between two price processes.
it does point to the large jumps as having a lesser incidence of transmission. This may reflect the emergence of trading halts or liquidity problems in the presence of severe effects. The fourth column of Table 4 further shows that the evidence of mutual excitation remains mostly unchanged as the volatility thresholds increase for mild and severe shocks, e.g., from $\alpha = 2$ to $\alpha = 5$. However, as $\alpha$ increases, there is stronger evidence of FTS through mutual excitation due to weak shocks.

The evidence for no excitation (reported in the fifth column of Table 4) shows that 11–12.5% of mild to severe shocks do not provoke a FTS response in the S&P 500 to gold. This is consistent across the values of $\alpha$. For weak shocks (corresponding to size 0.001), however, the evidence of no FTS shocks is more pronounced, up to 59% for cases of $\alpha = 2$. The evidence supports the supposition that weak shocks do not amplify an FTS reaction in relatively calmer (lower $\alpha$) markets; consistent with the findings of Baur and McDermott (2010).

Columns 6 and 7 report the percentage of periods with indeterminate FTS regions for completeness. While rejecting both null hypotheses is likely around 11–15% of all cases, there are many periods in which our testing procedures accept both null hypotheses for excitation effects (44–54% in the last column). As Baele et al. (2014) argue, one explanation for this finding could be that when systematic and idiosyncratic jumps occur jointly within the same periods, mutual excitation occurs and, hence, financial flights become intractable. We observe such regularities in the mutually exciting jumps of S&P 500 index and gold prices.

The real advantage of our approach is that—having identified the presence of mutual excitation within a sub-sample period, we are able to investigate the timing the arrivals of financial flights between stocks and gold. First, we identify our Case I-implied FTS periods in which we observe flows from equities to the gold market. These FTS episodes represent risk-off trading schemes. Second, we consider the Case II regimes of risk-on trades, the seeking-returns-strategy, SRS, which implies a (reverse) financial flight from safe-havens to risky assets (i.e., from gold to stocks). Intuitively, while FTS episodes capture the market conditions in which investors increase their holdings in safe assets, SRS regimes reflect the periods when risk appetite surges in the marketplace. Table 5 presents the (business cycle) periods of flight-to-safety (FTS) and seeking-returns-strategy (SRS) identified by the mutual excitation tests.
Table 5 indicates that, over the sample years 2007-2014, the second and third quarters of 2007 are distinct FTS states (columns 2 and 4). The identified jump excitation spells surround the beginning of the liquidity crisis when BNP Paribas froze the redemption for three investment funds on August 9, 2007. The data exhibit FTS patterns even in the last quarter of 2007 (columns 4 and 5), reflecting the persistence of market fear and stress. Negative (jump-type) shocks hitting stock markets excite positive jumps in gold. However, at the same time, there is also evidence of SRS behavior particularly in 2007/Q2. This likely reflects that investors were not clear how wide-spread the crisis was going to become at this point, and hence much of the upheaval in markets was evident in the U.K. and European markets, as opposed to the U.S. data investigated here.

While the first half of 2008 contains neither FTS nor SRS periods (rows 2008/Q1 and 2008/Q2), the third quarter (July-September 2008) is linked solely to FTS regimes. One explanation for this result could be related to investors’ judgment about tail risk events. The events tied to the bankruptcy of Bear Stearns (mid-March 2008) did not significantly lead to risk-off flights from U.S. stocks to safe assets. The collapse of Lehman Brothers, however, appears to create severe FTS spells (row 2008/Q3) with high degrees of mutual excitation between equities and gold (i.e., $\alpha = 5$). For the periods following the last quarters of 2009, Table 5 shows the clusters of FTS and SRS regimes at different strength levels ($\alpha = 2, 5$). For example, while the third and fourth quarters of 2009 can be characterized as FTS cycles, the last quarter of 2009 and early 2010 contain SRS regimes. The table reveals similar FTS/SRS regularities until the end of 2011. Regardless of the choice of flight level (i.e., either $\alpha = 5$ or $\alpha = 2$), we do not find strong evidence of FTS in 2012. In 2013/Q1 there is evidence across the spectrum of tests of SRS behaviour. This is consistent with the narrative of the time, as investors sought to return to equity markets in the light of increasing returns, and a new high in the S&P 500 leading to its recovery to pre-crisis levels.

4.3. Flight-to-quality episodes: cross-excitation between stocks and bonds

The previous section identifies flight-to-safety episodes—between stocks and gold—based on the mutual excitation tests. In this section, we analyze the excitation dynamics in U.S. capital markets, and instead characterize Case I as flight-to-quality (FTQ) spells propagating from stocks to bonds. We begin by testing for mutual excitation between negative jumps in the stock prices and positive jumps in bond prices. Table 6 reports the test results for the S&P 500 index and
long-term (30-year) U.S. Treasury bond futures.

The first category of mutual excitation shows that flights from the U.S. stock market to the bond market are relatively infrequent, occurring in under 10% of the quarters (reported in column 4 of the table). The frequency of FTQ regimes is unaffected by either the strength of the volatility via choice of $\alpha = 3$, or the choice of $H = H(0; 1)$ or $H = H(6; 1)$. The strength of the jump does matter. More severe jumps result in stronger mutual excitation behavior between S&P 500 and 30-year bond futures. Overall, these results imply that FTQ occurs mostly when the flight level is strong (i.e., for larger jump size) and is little affected by volatility. The frequency of evidence for Case II episodes of no excitation (column 5) weakens with increasing volatility $\alpha$, but this is most marked with weaker jumps. As with the FTS results analysed above, there is relatively little occurrence of the contradictory cases where the tests both reject their nulls (shown in column 6) and high frequency of both tests accept their nulls (shown in column 7). In the latter case, this is mollified where jumps are weak and volatility is low.

We now consider evidence for the occurrence and frequency of FTQ events. Table 7 presents the arrival times of the detected FTQ spells together with the corresponding SRS episodes between stocks and bonds. The results in this table show that the occurrences of FTQ are quite clearly identified as arriving in 2007/Q2 and Q3, 2011/Q2 and 2013/Q2. The reverse flights are identified very infrequently; the main occurrence being in 2008/Q3. That is, we identify a flight from bonds to stocks during the period of Lehman Brothers’ collapse as the only consistent SRS spell. This result is consistent with the findings of Chang and Cheng (2016), who find that the bond markets are the source of shocks transmitted to the equity market during the crisis period.

Overall, our results show that FTQ occurs less frequently than FTS over the entire sample period 2007–2014 (left panels in Tables 7 and 5, respectively). The evidence for Case I type events of FTQ or FTS is less pronounced with smaller jump sizes, inferring that market participants are less likely to ignore smaller shocks of this nature. When the jump level rises so does the evidence for FTQ, while FTS is highest with moderate-sized jumps. The relatively low incidence of FTQ episodes may account for this difference. The levels of $\alpha$ have less effect, that is, the results are less influenced by an increased severity of jump shocks relative to stochastic volatility. The evidence
for Case II type events, where there is no mutual excitation, shows that for both cases of bonds or gold as the alternative asset, weak jump shocks during lower \( \alpha \) periods are relatively likely to provoke a non-excitation response. As \( \alpha \) increases, however, this declines considerably more than for larger sized jumps. Combining this information, we find that while weak shocks are less likely to trigger a Case I response of either FTS or FTQ, as \( \alpha \) rises, there is declining evidence of a negative equity market shock provoking a no-excitation response via Case II. This suggests the information content of FTS and FTQ come from the shocks themselves, but that information may have more impact in determining a response in a more volatile market.

SRS exhibits asymmetric patterns linked to FTS and FTQ trades. Specifically, comparing the right panels of Table 5 with those in Table 7, we observe that SRS from safe-havens occurs more frequently than SRS from bond investments, especially if the jump threshold is high (\( \alpha = 5 \)). That is, when large adverse shocks hit financial markets, investors tend to move away from risk towards safety of gold and quality of bonds (i.e, FTS and FTQ spells, respectively). Our results suggest that, as the market jump turmoil settles down, traders typically move back to risky assets (i.e. stocks) by mostly selling gold rather than long-term U.S. debt. While negative jumps in gold prices excite positive jumps in the S&P 500 index, evidence of a reverse excitation effect—from bonds to stocks—is weak.

Given the evidence of FTS, FTQ and SRS episodes, we use the mutual excitation tests to identify financial flights when the market (trading) environment is favorable for buying strength. This form of flight—denoted here as bull market flight (BMF)—corresponds to Case III (Section 3.2.3) and entails mutual excitation between the positive price jumps of two assets. We calculate the frequency of trading quarters associated with BMF spells in the data. Table 8 reports the fractions for each directional fund flow across S&P 500, bond futures and gold. Panels A and B of the table reveal weak evidence of mutual excitations contributing to BMF when we test the flight from positive S&P 500 jumps to positive bond and gold jumps (i.e., 3%–7% of all periods). Contrary to this pattern, the data exhibit a strong (reverse) excitation effect—in
the form of BMF—from gold and bond jumps to S&P 500 jumps (Panels C and D, respectively). Approximately 15% (22%) of quarters in the entire sample coincide with episodes in which positive gold (bond) jumps lead to positive S&P 500 jumps. One explanation for this finding could be related to the role of interest rates in determining excitation channels. When financial assets move into bull-market territory (due to positive news shocks), low interest rates are likely to push down cost of equity, raising the present value of future cash flows to firms. This environment may in turn encourage investors to shift their portfolios from gold and bonds to stocks (Panels C–D), and hence reverse excitation transmission remains rather limited (Panels A–B).\textsuperscript{32}

5. Robustness checks and extensions

In this section, we assess the robustness of our results and consider extensions in several respects. First, we account for the impact of intraday periodicity in volatility. To achieve this, we follow Erdemlioglu et al. (2015) by filtering out the periodic component of spot volatility from the data and apply our testing procedures to filtered log-returns. Second, we examine whether market microstructure noise impairs our testing procedures. Third, we investigate the choice of parameter values in affecting the size and power of the excitation tests. Fourth, we extend our baseline empirical analysis (i.e., based on quarterly horizons and 5-minutes data) to the case with weekly horizons and 5-seconds data. Through this analysis, we show that our main conclusions remain qualitatively the same with respect to trading horizon and sampling frequency. We finally complete this section by examining the sensitivity of detected FTQ patterns to the short/long maturity of U.S. Treasury bonds. For brevity, we report only the main findings here.\textsuperscript{33}

5.1. The impact of intraday periodicity

To study the effect of periodic volatility, we focus on the analysis of FTS and FTQ identification. In the presence of periodicity, the distributional characteristics of filtered returns (Figure 8) and unfiltered returns (Figure 9) remain qualitatively similar.\textsuperscript{34} As we apply our testing procedures on

\textsuperscript{32}It is worth noting that the majority of the potential cases of BMF here are found in the indeterminate space where both tests accept their nulls.

\textsuperscript{33}These results are available upon request.

\textsuperscript{34}To account for periodicity, we standardize raw returns \((r_{t,i})\) by intraday periodicity estimates \((\hat{f}_{t,i})\), i.e., \(r_{t,i}^* = r_{t,i}/\hat{f}_{t,i}\), where \(\hat{f}_{t,i}\) estimates the periodic component of the volatility. As in Boudt et al. (2011) and Erdemlioglu
filtered returns, intraday periodicity appears to increase the number of episodes of indeterminate region (where both nulls are rejected), whereas its impact on mutual excitation is limited. For a given shock magnitude (e.g., $\alpha = 2$), results with periodicity-filtered returns are quite similar to those with our former results.

Considering FTQ trades, we find that the impact of periodic volatility on identified frequencies remains limited. Perhaps surprisingly, periodicity-filtered tests indicate even stronger evidence of FTQ regimes in the data. Taken together, while periodic volatility barely affects our mutual excitation results, joint-hypothesis tests of rejecting both nulls reveal mixed evidence. Overall, our main conclusions remain unchanged, and the impact of periodicity on our testing procedures appears to be negligible.

5.2. The impact of market microstructure noise

We now assess the effect of market microstructure noise on our mutual excitation tests. While we do not develop noise-robust procedures here—leaving this for future research—we show that the impact of noise appears to be fairly limited.

In the presence of noise, we consider that the true value of the log price is $X(d, t)$ (for assets $d = 1, 2$), as in Equation (1), but in the data we observe $Z(d, t)$. That is (in differential form),

$$dZ_{d,t} = dX_{d,t} + de_{d,t},$$

where $e(t)$ is the additive noise term for each asset $d = 1, 2$. To calibrate the noise parameters, we follow Aït-Sahalia et al. (2012) and set $e_{d,t} = C\sigma_{d,t}\Delta^{1/2}\epsilon_{d,t}$ with $\epsilon_{d,t} \sim N(0, 1)$, and $C = 2.6$ is the noise size. While this setup allows for temporal heteroscedasticity in noise, we also consider a case in which noise is independent of stochastic volatility (i.e., $e_{d,t} = C\epsilon_{d,t}$). Given this noise specification, we repeat our simulations in Section 3 to evaluate the performance of the tests in noisy high-frequency data.

et al. (2015), we use the WSD estimator to filter out periodic components of volatility from the data. The procedures are available upon request.
When we consider temporal heteroscedasticity in noise, the simulations indicate that the tests are still correctly sized and powerful with the choice of \( C = 2 \).\(^{35}\) As we increase the noise strength (from \( C = 2 \) to \( C = 6 \)), empirical size remains unaffected, but we observe a small loss of power when testing for mutual excitation from asset \( X_2 \) to \( X_1 \). These results hold both in the absence (i.e., \( \xi = 0 \)) and presence (e.g., \( \xi = 12 \)) of excitation effects in simulations.

We next consider the case of non-shrinking noise such that noise is homoscedastic and independent of spot volatility. Under this noise specification, our results reveal that the impact of noise is non-negligible. Regardless of the noise size, we find that the tests become undersized and lose their power. Specifically, we notice that the Monte Carlo rejection frequencies are substantially lower than the corresponding theoretical levels, which generates almost flat lines in rejection figures instead of forty-five-degree lines (see e.g., Figure 4).

[ Insert Figure 10 about here ]

Given this simulation evidence, we further check how noise affects the power variation measures that we utilized in our empirical analysis (a) to estimate continuous volatility, (b) to estimate truncated realized volatility, (c) to estimate jump activity index, (d) to locate the arrivals of large jumps of assets \( X_1 \) and \( X_2 \), and (e) to estimate the jump intensity process of \( X_2 \). For this assessment, we follow Phillips and Yu (2006) and show in Figure 10 the volatility signature plots (left panels) as well as the corresponding microstructure noise functions (right panels).\(^{36}\)

Visual inspection of the plots suggests two noticeable patterns. First, the impact of noise on realized gold volatility is limited with (almost) flat noise function (lower panels). Second, while potential noise appears to create small upward bias for bond volatility, the bias is downward (and still small) for the S&P 500 index volatility (middle and upper panels, respectively). The plots also reveal that the noise is typically present at very high-frequency (such as 5-seconds) and the effect dies out fairly quickly. In light of these regularities, it is of interest for future research to focus on

\(^{35}\) For brevity, we do not report the results. The figures showing the size and power of the tests are available upon request.

\(^{36}\) For all assets (S&P 500 index, 30-year U.S. Treasury bond futures and gold), we use 5-seconds time-based bars (as the highest sampling frequency), where we skip empty intervals over the period January 1, 2007 – December 31, 2013 (as in the main empirical analysis). The trading sessions include the day sessions and we select the time zone as GMT to ease comparison with our baseline results. The trading session starts at 14:30 and ends at 21:00. For the 30-year U.S. Treasury bond futures, we consider the same data adjustment and include both pit and electronic trades.
removing noise from data and develop noise-robust mutual excitation tests.

### 5.3. Sensitivity of the tests to parameter values

We turn to the sensitivity analysis of the tests for different parameter values. Figure 11 displays the Monte Carlo rejection frequencies, and Table 9 reports the set of parameters chosen in each sensitivity analysis (from I to VI).

We begin by varying the idiosyncratic tail probability (sensitivity analysis I and II). As we increase the probability of observing a tail event (from 0.25% to 1% and 5%), the upper panels of Figure 11 indicate upward size distortion when we test for the null of mutual excitation (dotted lines). Despite this property, we notice that the power of the test (for rejecting the null of no mutual excitation) increases for high values of tail probabilities (solid lines).

In the second part of our sensitivity analysis, we set the tail parameter to 0.25% but increase the excitation likelihood from 80% to 90% and 99% (sensitivity analysis III and IV in Table 9, respectively). For these values, the excitation parameter $\xi$ equals to 14 and 15 instead of 12 or 50 (i.e., the base parameters). The middle panels of Figure 11 show that the test is correctly sized (forty-five-degree dotted line), whereas the power (for rejecting the null of no mutual excitation) weakens (solid line). This finding is consistent with our previous discussion in the sense that choosing an excitation level $\xi = 12$ may entail power loss and hence we suggest using $\xi = 50$ as an alternative baseline parameter.\(^{37}\)

Finally, we check the behavior of the tests when the jump activity is relatively high (lower panels of Figure 11). For this assessment, we consider two $\beta_2$ values and set $\beta_2 = 1.50$ and $\beta_2 = 1.75$ (V and VI in Table 9). In this case, we observe significant power loss (solid lines) and size distortion (dotted lines). This result may imply that it is rather challenging to disentangle mutually-exciting jumps from other types when detected jumps exhibit high level of infinite-activity with small and frequent moves. Considering all these features, we summarize the main findings in the last two columns of Table 9: the behavior of the mutual excitation tests is fairly correct with the base

\(^{37}\) Of course, it is important to note that the power does not necessarily increase as the excitation parameter becomes larger. In additional simulation exercises (unreported for brevity), we find that the testing power appears to weaken once again when we choose excitation parameter as, for instance, $\xi = 75$ or $\xi = 100$.  

33
parameters (i.e., for $\xi = 12$ and $\xi = 50$), and the power of tests (when DGP has mutually-exciting jumps) is highest for the value $\xi = 50$.

5.4. Varying testing horizon and sampling frequency

In our baseline empirical analysis, we use 5-minutes data and apply the testing procedures to quarterly time horizon. Based on this framework, we provide evidence that financial flights (particularly FTS) often coincide with periods of high market stress (see e.g., Tables 5 and 7). We now characterize these patterns using 5-seconds sampling frequency over weekly rather than quarterly horizons. Figures 12 and 13 plot the test statistics (given by Equation (19)) for FTS and FTQ spells, respectively, over the full sample from January 1, 2007 to December 31, 2013.

The panels of Figure 12 indicate that mutual excitation episodes are significantly associated with the periods of NBER-defined recessions and the intensification of the European debt crisis (2011-2012). The fraction of detected FTQ episodes (see Figure 13) is again smaller than that of FTS spells as discussed in the main text. We also continue to observe that realized FTQ patterns track business cycle variation linked to financial turmoil. These results are overall consistent with our baseline conclusions for financial flights, and our previous finding that the results are robust to increased microstructure noise as might be expected at this high sampling frequency.

5.5. Maturity check for FTQ flows

Our final robustness check attempts to extend the baseline FTQ analysis (Section 4.3) to various maturities of the U.S. Treasury bonds. To better highlight the potential excitation differentials (across short and long end of the curve), we apply the testing procedures to 5-second bond futures data over weekly horizons. Figure 14 compares the identified FTQ flows (from S&P 500 stock index to bond futures) for various maturities ranging between 2-years to 30-years.

Amongst all maturities considered, the 10-year U.S Treasury note exhibits the highest activity of the mutual exciting jumps, followed by the 30-year bond (upper-right).\textsuperscript{38} Similar to the baseline

\textsuperscript{38}It is perhaps important to note that we identify excitation episodes based on the critical regions of the test statistics. For brevity, we do not report these flight periods, yet they are available upon request. As in our baseline

34
findings, FTQ-implied jump propagation varies over the business cycle and occurs mostly in periods of high market uncertainty, such as the BNP Paribas’ suspension and subprime mortgage crisis (i.e., surrounding the weeks 28–32 and 44–120, respectively). Although FTQ patterns with 30-year maturity closely resemble the trend in 10-year note (upper-left versus upper-right), short-dated bond futures (i.e., 2-year) reveal little evidence of excitation effect (lower-right). In line with the conjecture of Engle et al. (2012), this could be explained by the fact that 2-year note exhibits high liquidity (and stronger resilience) in crisis periods, compared to its 5-, 10- and 30-year counterparts. Therefore, the majority of the flights are likely to arrive with the high volatility regime of long-term rather than short-term notes, consistent with the findings of Chang and Cheng (2016).

6. Conclusion

This paper examines the spillover effects of tail-type sudden shocks in global financial markets. Measuring linkages is empirically challenging because shocks are typically difficult to estimate, clustered in a specific time period and subsequently likely to spread across different asset classes or borders. In times of crisis, correlation measures tend to be biased, posing further difficulties in understanding how (and when) markets are connected.

We develop an alternative approach to identify financial contagion in high-frequency financial data. This approach tracks the time-variation, direction and strength of different shock spillovers across assets and regions. In our specification, jump-type events in one asset have the potential to increase the intensity of further jumps in other assets, creating mutual-excitation effects. We propose a new nonparametric testing procedure that allows us to characterize the episodes of mutually exciting shock dynamics.

Using our proposed methodology, we study jump linkages in various forms of financial flight, including flight-to-safety, flight-to-quality, flows of seeking-returns-strategies (i.e., reverse flights) and bull-market flights. We find evidence of asymmetric and time-varying cross-excitation effects in high frequency data across stocks, bonds and gold prices. Our econometric analysis further reveals two distinct features of such flight episodes. First, when risk aversion grips markets, investors

---

analysis (Section 4.3), these flights are categorized as Case I of mutual excitation. To ease visual inspection, we plot in Figure 14 the excitation weeks marked by the values of the (no excitation) test statistic.
demand both safety (gold) and quality (bonds), but they do so in different business cycles and at different frequencies. Second, as market fear and panic subside, risk appetite prompts investors to rebalance their portfolios toward stocks, particularly by selling gold rather than bonds. We reach this conclusion because, while large negative drops in gold prices significantly lead to positive jumps in the S&P 500, excitation from bonds to stocks turns out to be weak. These results and conclusions are mainly robust to alternative parameter values (e.g. $\alpha$), trading horizon (quarterly versus weekly) and sampling frequency (e.g., 5-seconds versus 5-minutes).

This research can be extended in several ways. First, our framework relies on testing procedures entailing bivariate (pair-wise) setting, and hence we focus on testing for high-frequency mutual excitation in jumps of two financial assets. A natural extension of this setup would be to develop a multivariate (testing) approach, which may in turn allow to characterize (system-wise) risk spillovers. In the same line yet on the empirical front, our tests can be applied to high-frequency data on international stock markets. This helps researchers examine in depth volatility versus jump connectedness and consider implications for asset pricing and risk premia (i.e., in the absence and presence of mutual excitation). Another extension could be on the identification of the excitation process. Specifically, we investigated in this study whether or not price jumps in one asset excite the jump intensity of other asset. However, the source of jump excitation could be also volatility-driven, that is, volatility jumps (or volatility of volatility) are also likely to trigger excitation effects. Future research may decompose these features and provide insights on methodological and empirical grounds.

Acknowledgments

We thank the editor and two anonymous referees for excellent comments; Massimiliano Caporin, Pasquale Della Corte, Ozgur Demirtas, Dobrislav Dobrev, Louis Eeckhoudt, Robert Engle, Neil R. Ericsson, Roland Füss, Andrew Harvey, David F. Hendry, Sébastien Laurent, Simone Manganelli, Christopher J. Neely, Marc Paolella, Andrew Patton, Olivier Scaillet, Paul Schneider, Victor Todorov, Giovanni Urga, Lars Winkelmann, Kamil Yilmaz and participants at the 9th Annual SoFiE Conference, 14th Paris December Finance Meeting, IAAE 2016 Annual Conference, Federal Reserve Bank of St. Louis Research Seminar, 16th Oxmetrics Financial Econometrics Conference and Koc University Economics Seminar for valuable comments and helpful suggestions. This paper has previously been circulated under the title “Financial flights, stock market linkages and jump excitation.” Dungey and Matei acknowledge funding from Australian Research Council grant (DP130100168).
A. Proofs

A number of the theoretical constructs follow the (one-dimensional) theorems in Boswijk et al. (2017) in a trivial way. This appendix presents the key assumptions in extending the self-excitation (univariate) framework of Boswijk et al. (2017) to the multivariate case. To prove the main theorems, we require the function $H$ to satisfy certain conditions. We start by summarizing these conditions in the following assumptions. Notably, we require the following

A.1. Assumptions and main results

**Assumption 2.** This follows the same form as Assumption 4 in Boswijk et al. (2017) with the difference that $H$ is $C^1$ on $D(R) \times \mathbb{R}^2_+$ and, for some $q', q'' > \max \beta_1, \beta_2$,

and note the degeneracy condition given in equation (3.12) of Boswijk et al. (2017).

**Assumption 3.** This assumption follows Assumption 5 in Boswijk et al. (2017).

We can replace the local boundedness assumptions in Assumption 1 by boundedness. This can be done in a localization procedure, a standard approach in high frequency econometrics (see e.g., Jacod and Protter, 2011). See also Assumption B.1 in Boswijk et al. (2017). We have the following intermediate result:

**Proposition 3.** Suppose $X$ (either $X_1$ or $X_2$ or both) satisfy Assumption 1 and the following condition hold

$$1 - \omega \beta < \rho < (1 - \omega \beta) + 2\phi' \Lambda 2 \phi'' \wedge \frac{1}{2} \omega \beta. \quad (A.1)$$

Given any finite number of time points $\{t_p\}_{p=1}^P$, the following vector of random variables:

$$\sqrt{k_n \Delta_n} \left( \lambda(k_n)_{t_p} - \lambda_{t_p} \right)_{1 \leq p \leq P}, \quad (A.2)$$

converges stably in law to a vector of Gaussian random variables $(W_{t_{p}})_{p=1}^P$ independent of $\mathcal{F}$, with

$$\mathbb{E}_{t_p}(W_{t_p}^2) = \lambda_{t_p} \frac{\alpha^2 C_\beta(2)}{(C_\beta(1))^2} \text{ and } \mathbb{E}_{t_p \wedge q}(W_{t_p} W_{t_q}) = 0.$$

Note that when considering $U(H, k_n)_t^{12}$, we only need to estimate the jump intensity process $\lambda_{2,t}$. Therefore, we replace $\beta$ by $\beta_2$ in condition (A.1), which yields condition (18).

We complete this subsection by introducing some useful notation. We start by decomposing $X_t = X_{d,t}$ $(d = 1, 2$, which we omit for brevity) as $X_t = X_t^c + X_t^d$, where $X_t^c = X_0 + \int_0^t b_s ds + \int_0^t \sigma_s dW_s$, and

$$X_t^d = x 1_{\{|x| \leq 1\}} * (\mu - \nu)_t + x 1_{\{|x| > 1\}} * \mu_t$$

$$= x 1_{\{|x| > \delta\}} * \mu_t + x 1_{\{|x| \leq \delta\}} * (\mu - \nu)_t - x 1_{\{\delta < |x| \leq 1\}} * \nu_t$$

$$= X_t^d(\delta) + X_t^d(\delta) + B(\delta).$$
Throughout, we adopt the standard notation (see e.g., Jacod and Todorov, 2010, Aït-Sahalia et al., 2017 and Boswijk et al., 2017). For each integer \( m \geq 1 \), we define \((S(m, q) : q \geq 1)\) to be the successive jump times of the Poisson random measure given by

\[
\mu^X([0, t] \times \left\{ x : \frac{1}{m} < \gamma(x) \leq \frac{1}{m - 1} \right\}).
\]

This refers to (5) for \( \mu^X \) and to Assumption B.1 in Boswijk et al. (2017) for \( \gamma(\cdot) \). One may relabel the two-parameter sequence \((S(m, q) : m, q \geq 1)\) to become a single sequence \((T_p : p \geq 1)\) that exhausts the jumps of \( X \). Based on this jump classification, we define

\[
\begin{align*}
A_m &= \{ z : \gamma(z) \leq 1/m \}, \quad \gamma_{m,t} = \int_{A_m} (\gamma(x))^{\beta} F_t(dx), \\
T_m^t &= \{ p : \exists \ p \geq 1 \text{ and } m' \in \{1, \ldots, m\} \text{ s.t. } T_p = S(m', q) \leq [t/\Delta_n]\Delta_n \}, \\
i(n, p) &= \text{the unique integer such that } T_p \in (t^n_{i(n,p)-1}, t^n_{i(n,p)}), \\
J(n, m, t) &= \{ i(n, p) : p \in T_m^t \}, \quad J'(n, m, t) = \{1, \ldots, [t/\Delta_n]\} \setminus J(n, m, t),
\end{align*}
\]

\[
\Omega_{n,m,t} = \bigcap_{p \neq q, p,q \in T_m} \{ T_p < [t/\Delta_n]\Delta_n - 3k_n\Delta_n, \text{ or } T_p > 3k_n\Delta_n \text{ and } |T_p - T_q| > 6k_n\Delta_n \}.
\]

Observe that, for any \( m \),

\[\lim_{n \to \infty} \mathbb{P}(\Omega_{n,m,t}) = 1.\]

We further decompose \( X \) as follows

\[X_t = X'(m)_t + L(m)_t + J(m)_t,\]

where

\[
\begin{align*}
b(m)_t &= \begin{cases} b_t - \int_{(A_m)_{x<}} xF_t(dx) & \text{if } \beta > 1, \\
b_t, & \text{if } \beta \leq 1, \end{cases} \\
L(m)_t &= \begin{cases} \int_0^t \int_{A_m} x(\mu - \nu)(ds, dx) & \text{if } \beta > 1, \\
\int_0^t \int_{A_m} x\mu(ds, dx), & \text{if } \beta \leq 1, \end{cases}
\end{align*}
\]

\[X'(m) = X_0 + \int_0^t b'(m)sds + \int_0^t \sigma_s dW_s, \quad J(m)_t = \int_0^t \int_{(A_m)_{x<}} x\mu(ds, dx).\]

In the sequel, let \( K \) be a finite positive number, the value which may vary from line to line.

### A.2. Proof of Theorem 1

Following the arguments in Boswijk et al. (2017), we decompose \( U(H, k_n) \) (we omit the superscript “12” for simplicity) into two parts, on the set \( \Omega_{n,m,t} \) (with \( X = X_1 \)). That is,

\[U(H, k_n) = \tilde{U}^n(m) + U^n(m),\]
under the conditions of part (b), the remaining part of the proof is that, for all \( t < \infty \) and Compactness yields that, for any \( m \) and \( t \),

\[
\sqrt{\frac{k_n \Delta_n}{\Delta_n^{\alpha/2}}} (\tilde{U}^n(m)_t - \tilde{U}(m)_t, \tilde{U}'^n(m)_t - \tilde{U}(m)_t) \overset{\mathcal{L}_{st.}}{\to} \left( \mathcal{U}_t, \frac{1}{w} \mathcal{U}_t + \sqrt{w-1} \mathcal{U}'_t \right),
\]

where

\[
\mathcal{U}_t = \sum_{p \in T_t} (H'_2(X_{T_p}, X_{T_p^+}, \lambda_{T_p}) W_p^- + H'_4(X_{T_p}, X_{T_p^+}, \lambda_{T_p}) W_p^+),
\]

and

\[
\mathcal{U}'_t = \sum_{p \in T_t} (H'_2(X_{T_p}, X_{T_p^+}, \lambda_{T_p}) W'_p^- + H'_4(X_{T_p}, X_{T_p^+}, \lambda_{T_p}) W'_p^+),
\]

and we refer to Boswijk et al. (2017) for the details. According to the discussion therein, the limiting process \( \mathcal{U}_t \) of \( U(H, k_n)_t \) is a Gaussian martingale conditional on the filtration generated by \( \{X_{1,s}\}_{s \geq t} \) and \( \{X_{2,s}\}_{s \geq t} \) with conditional variance given by

\[
\mathbb{E}(\mathcal{U}_t^2 | \mathcal{F}) = \frac{\alpha^2 C_\beta(2)}{C_\beta(1)} \sum_{s \leq t} (\lambda_{s-h_2}(X_{s-h}, X_s, \lambda_{s-h}, \lambda_h)^2 + \lambda_{s-h}(X_{s-h}, X_s, \lambda_{s-h}, \lambda_h)^2).
\]

We can approximate this limiting variance process by \( U(G, k_n)_t \), where the function \( G \) is given by (20). Also, \( U(G, k_n)_t \) is a consistent estimator of the limiting conditional variance. Then, the statistic \( t^{12} \) defined in (19) converges stably in law to a standard normal random variable under the null hypothesis \((A)\). While under \( H_1 \) of \((A)\), the calculations in Boswijk et al. (2017) show that it will diverge to infinity with probability one. This completes the proof of part (a) of the theorem.

Under the conditions of part (b), the remaining part of the proof is that, for all \( t < \infty \) and 39
$
abla > 0$, we have

$$
\lim_{m \to \infty} \limsup_{n \to \infty} \mathbb{P} \left( \sqrt{\frac{k_n \Delta_n}{2 \beta_2}} \left| \tilde{U}^n(m)_t - \bar{U}(m)_t \right| > \eta \right) = 0, \tag{A.4}
$$

and the same for $\tilde{U}^n(m)_t - \bar{U}(m)_t$. Nevertheless, it is enough to prove (A.4) only.

Define $Y(m) = X'(m) + L(m)$, i.e., the sum of continuous component and “small” jumps of $X$.
Hence, for any $i \in J'(n, m, t)$, we have $\Delta_i^n Y(m) = \Delta_i^n X$. Consequently, we can rewrite $\tilde{U}^n(m)_t$ as

$$
\tilde{U}^n(m)_t = \sum_{i \in J'(n, m, t)} H(X_{i-1}, X_{i-1} + \Delta_i^n Y(m), \lambda(k_n)_{j-k_n-1}, \lambda(k_n)_{j}) 1_{\{|\Delta_i^n Y(m)| > \alpha \Delta_m^n\}}.
$$

As in Boswijk et al. (2017), we introduce the following random variables

$$
\bar{\chi}(m, 2)^n_i = H(X_{i-1}, X_{i-1} + \Delta_i^n Y(m), \lambda, \lambda) 1_{\{|\Delta_i^n Y(m)| > \alpha \Delta_m^n\}}
$$

$$
- \sum_{s \in I(n,i)} H(X_{s-}, X_{s-} + \Delta Y(m)_s, \lambda, \lambda) 1_{\{|\Delta Y(m)_s| \neq 0\}}.
$$

$$
\bar{\chi}(m, 4)^n_i = H(X_{i-1}, X_{i-1} + \Delta_i^n Y(m), \lambda, \lambda) 1_{\{|\Delta_i^n Y(m)| > \alpha \Delta_m^n\}}
$$

$$
- \sum_{s \in I(n,i)} H(X_{s-}, X_{s-} + \Delta Y(m)_s, \lambda, \lambda) 1_{\{|\Delta Y(m)_s| > \alpha \Delta_m^n\}}.
$$

$$
\bar{\chi}(m, 5)^n_i = H(X_{i-1}, X_{i-1} + \Delta_i^n Y(m), \lambda, \lambda) 1_{\{|\Delta_i^n Y(m)| > \alpha \Delta_m^n, |\Delta Y(m)_s| > q_n\}}
$$

$$
- \sum_{s \in I(n,i)} H(X_{s-}, X_{s-} + \Delta Y(m)_s, \lambda, \lambda) 1_{\{|\Delta Y(m)_s| > \alpha \Delta_m^n\}},
$$

where $q_n = [(\alpha \Delta_m^n)^{-1}]$ for some $l \in (1, 1/(2\nu \beta_1))$.

Again, following the procedures of Boswijk et al. (2017) (Steps 3 to 6 in the proof of Theorem 3.2), we can prove that for any $m$

$$
\limsup_{n \to \infty} \sqrt{\frac{k_n \Delta_n}{2 \beta_2}} \sum_{i \in J'(n, m, t)} \mathbb{E} \left( |\tilde{U}^n(m) - \sum_{i \in J'(n, m, t)} \bar{\chi}(m, 4)^n_i| \right) = \limsup_{n \to \infty} K \gamma_m \Delta_n^{(1-\rho)+(1-\nu \beta_2)} = 0.
$$

The last equality holds because we choose $\rho < 1, \nu < 1/2$ and by default $\beta_2 \leq 2$ in our analysis.

Note that the evaluations of $\bar{\chi}(m, 2)^n_i$, $\bar{\chi}(m, 4)^n_i$ and $\bar{\chi}(m, 5)^n_i$ are slightly different from those in Boswijk et al. (2017). For brevity, we omit most calculations here and only highlight the differences. We can replace the results given in (B.6) and (B.7) in Boswijk et al. (2017) by the following

$$
\mathbb{E}(|\bar{\chi}(m, 2)^n_i - \bar{\chi}(m, 4)^n_i|) \leq K \sum_{s \in I(n,i)} \mathbb{E}(\{|\Delta Y(m)_s|^q 1_{\{|\Delta Y(m)_s| \leq \alpha \Delta_m^n\}}\})
$$

$$
\leq K \int_{I(n,i)} \int_{\gamma(x) \leq \alpha \Delta_m^n} \mathbb{E}(\{(\gamma(x))^q F_t(dx)ds\}
$$

$$
\leq K \Delta_n^{\nu(q-\beta_1)} \int_{I(n,i)} \int_{A_m} \mathbb{E}(\{(\gamma(x))^{\beta_1} F_t(dx)ds\}
$$

$$
\leq K \gamma_m \Delta_n^{1+\nu(q-\beta_1)}, \tag{A.5}
$$
and
\[
\mathbb{E} \left( \tilde{\chi}(m, 4)_i^n - \tilde{\chi}(m, 5)_i^n \right)
\leq K \mathbb{E} \left( \left( |X_{s-} - X_{i-1}|^{q'_c} + |X_s - X_{i-1}|^{q'_c} \right) \cdot |\Delta_i^n Y(m)|^{q''} \cdot 1_{\{|\Delta Y(m)| > q_n \}} \right)
\leq K \left( \mathbb{E} \left( |X_{s-} - X_{i-1}|^{q'_c} + |X_s - X_{i-1}|^{q'_c} \right)^2 \cdot \mathbb{E} \left( |\Delta_i^n Y(m)|^{2q''} \right) \right)^{1/2}
\leq K \Delta_n \left( \Delta_n^{q'_c - 1 + \varpi(2q'' - \varpi)} / 2 + \Delta_n^{\varpi(q''_c + q''_d - \beta_1)} \right) \gamma_m,
\tag{A.6}
\]
respectively. Then, it is clear that if \( \rho < 1 - \varpi \beta_2 + 2 \varpi (q' - \beta_1) \), we have
\[
\lim_{m \to \infty} \limsup_{n \to \infty} \frac{k_n \Delta_n}{\Delta_n^{\varpi_2 / 2}} \sum_{i \in J'(n, m, t)} \mathbb{E} \left( |\tilde{\chi}(m, 4)_i^n - \tilde{\chi}(m, 5)_i^n| \right)
= \lim_{m \to \infty} \limsup_{n \to \infty} K \gamma_m \Delta_n^{(1 - \rho - \varpi \beta_2) + \varpi(q - \beta_1)} = 0.
\]
Moreover, as long as \( 1 - \rho - \varpi \beta_2 + (q'_c - 1 + \varpi(2q''_d - \beta_1)) \wedge 2 \varpi (q'_c + q''_d - \beta_1) \geq 0 \),
\[
\lim_{m \to \infty} \limsup_{n \to \infty} \frac{k_n \Delta_n}{\Delta_n^{\varpi_2 / 2}} \sum_{i \in J'(n, m, t)} \mathbb{E} \left( |\tilde{\chi}(m, 4)_i^n - \tilde{\chi}(m, 5)_i^n| \right)
= \lim_{m \to \infty} \limsup_{n \to \infty} K \gamma_m \Delta_n^{(1 - \rho - \varpi \beta_2) / 2} \Delta_n^{q'_c - 1 + \varpi(2q''_d - \varpi)} + \Delta_n^{\varpi(q'_c + q''_d - \beta_1)} = 0.
\]
Finally, (B.8) in Boswijk et al. (2017) becomes
\[
\sum_{i \in J'(n, m, t)} \mathbb{E} \left( |\tilde{\chi}(m, 5)_i^n| \right) \leq K \Delta_n^{\varpi(q'_c + q''_d - \beta_1)} \gamma_m.
\tag{A.7}
\]
Therefore, if \( 1 - \rho - \varpi \beta_2 + 2 \varpi (q' \wedge (q'_c + q''_d) - \beta_1) \geq 0 \), then we have
\[
\lim_{m \to \infty} \limsup_{n \to \infty} \sqrt{\frac{k_n \Delta_n}{\Delta_n^{\varpi_2 / 2}}} \sum_{i \in J'(n, m, t)} \mathbb{E} \left( |\tilde{\chi}(m, 5)_i^n| \right) = 0,
\]
which concludes the proof.

A.3. Proof of Theorem 2

**Step 1.** As in Boswijk et al. (2017), the leading term of \( \tilde{U}^n(m)_t \) is given by
\[
\tilde{\xi}_p^n = \frac{k_n \Delta_n}{2 \Delta_n^{\varpi_2 / 2}} \left( H''_{34} (X_{T_p -}, X_{T_p}, \lambda_{T_p}, \lambda_{T_p}) \tilde{\lambda}(k_n)_{T_p -} - \lambda_{T_p -} \right)^2
+ H''_{44} (X_{T_p -}, X_{T_p}, \lambda_{T_p}, \lambda_{T_p}) \tilde{\lambda}(k_n)_{T_p -} - \lambda_{T_p -} \right)^2
+ 2 H''_{34} (X_{T_p -}, X_{T_p}, \lambda_{T_p}, \lambda_{T_p}) \tilde{\lambda}(k_n)_{T_p -} - \lambda_{T_p -} \tilde{\lambda}(k_n)_{T_p -} - \lambda_{T_p -} - \lambda_{T_p -} \right),
\tag{A.8}
\]
and we obtain the following results:

\[
\frac{k_n \Delta_n}{2 \Delta_n \bar{\beta}_2} \tilde{U}^n(m)_t - \sum_{p \in T^n_m} \tilde{\xi}^n_p \xrightarrow{p} 0.
\]

\[
\sum_{p \in T^n_m} \tilde{\xi}^n_p \xrightarrow{L^1} \bar{U}(m)_t.
\]

\[
\tilde{E}(|\bar{U}(m)_t - \bar{U}_t| \mid \mathcal{F}) \leq K \sum_{s \leq t} |\Delta X_s|^q 1_{\{\Delta X_s \leq m\}} \leq K \gamma m t.
\]

Then, because we consider the same argument as in the previous proof, we prove part (a) of Theorem 2.

**Step 2.** Now we turn to part (b). Following the calculations in Boswijk et al. (2017), we can verify that

\[
\frac{k_n \Delta_n}{\Delta_n \bar{\beta}_2} \sum_i \mathbb{E}(\tilde{\chi}(m, 2)^n_i - \tilde{\chi}(m, 4)^n_i) \leq K t \gamma m \Delta_n^{1 - \rho - \omega \beta_2 + \omega (q' - \beta_1)},
\]

\[
\frac{k_n \Delta_n}{\Delta_n \bar{\beta}_2} \sum_i \mathbb{E}(\tilde{\chi}(m, 4)^n_i - \tilde{\chi}(m, 5)^n_i) \leq K t \gamma m \Delta_n^{1 - \rho - \omega \beta_2 + (q'_c - 1 + \omega (2q'_d - \beta_1))/2 + \omega (q'_c + q'_d - \beta_1)},
\]

\[
\frac{k_n \Delta_n}{\Delta_n \bar{\beta}_2} \sum_i \mathbb{E}(\tilde{\chi}(m, 5)^n_i) \leq K t \gamma m \Delta_n^{1 - \rho - \omega \beta_2 + \omega (q'_c + q'_d - \beta_1)},
\]

where the sum \(\sum_i\) denotes \(\sum_{i \in J(n,m,t)}\). Under condition \((b')\) of Theorem 2, the three terms above are all asymptotically negligible. Then, the conclusion follows the arguments of Boswijk et al. (2017).

**B. Data description and adjustment**

30-year US T-bond futures are traded on the Chicago Board of Trade (CBOT). The original time zone for the US bond futures is based on the Eastern time (EST), and we start sampling from 9:35 EST until 16:00 EST, the last observation of the trading day \(t^{\text{day}}\). These trading hours leave us \(M = 78\) five-minute intervals for the all sample trading days of 30-year US T-bond futures and gold futures.

To correctly match intra-day time intervals, we convert the original time zones to the Greenwich time (GMT) for all asset classes. For the bond futures, we take into account the day sessions as well as the trading at both pit and electronic platforms. We further exclude overnight sessions and apply the automatic rolling method to generate raw database. This method automatically determines the best times to roll of a continuous-time futures price series. The method first computes the daily volume of the both current-month and next-month contracts. Next, the current month contracts are rolled to the next-month contracts, when the volume of the new contracts exceeds the volume of the current ones. We select the contracts as front futures contracts. Those contracts are the contracts that are closest to maturity. The use of the front contracts allows us to process all nearest contracts into a single continuous contract.

The raw database is based on tick data for all asset classes. We create a database for 5-minute intervals by converting tick-by-tick data into intraday 5-minute frequencies. For the S&P 500 index data, we consider the business hours (i.e., the day session) as its period of trading activity, and hence start sampling from 9:35 EST until the market closes at 16:00 EST of the trading day \(t^{\text{day}}\). This
corresponds to 14:35–21:00 in GMT with $M = 78$ 5-minute intervals, as in the 30-year T-bond futures data. In our main empirical analysis, we apply our mutual excitation tests to quarter-by-quarter data (aggregated from 5-minute returns) in order to characterize financial flights over business cycles.
References


Table 1: The size of the mutual excitation tests

Panel A: Under the null of no mutual excitation ($\xi = 0$)

<table>
<thead>
<tr>
<th>Frequency $\Delta t$</th>
<th># of obs.</th>
<th>1%</th>
<th>5%</th>
<th>10%</th>
<th>1%</th>
<th>5%</th>
<th>10%</th>
</tr>
</thead>
<tbody>
<tr>
<td>5-seconds</td>
<td>4680</td>
<td>0.015</td>
<td>0.058</td>
<td>0.102</td>
<td>0.014</td>
<td>0.055</td>
<td>0.109</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.010</td>
<td>0.055</td>
<td>0.114</td>
<td>0.013</td>
<td>0.064</td>
<td>0.116</td>
</tr>
<tr>
<td>10-seconds</td>
<td>2340</td>
<td>0.020</td>
<td>0.071</td>
<td>0.128</td>
<td>0.021</td>
<td>0.073</td>
<td>0.124</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.009</td>
<td>0.051</td>
<td>0.104</td>
<td>0.010</td>
<td>0.055</td>
<td>0.108</td>
</tr>
<tr>
<td>30-seconds</td>
<td>780</td>
<td>0.012</td>
<td>0.053</td>
<td>0.096</td>
<td>0.013</td>
<td>0.050</td>
<td>0.093</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.010</td>
<td>0.065</td>
<td>0.108</td>
<td>0.011</td>
<td>0.050</td>
<td>0.104</td>
</tr>
<tr>
<td>1-minute</td>
<td>390</td>
<td>0.013</td>
<td>0.074</td>
<td>0.123</td>
<td>0.016</td>
<td>0.065</td>
<td>0.118</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.018</td>
<td>0.065</td>
<td>0.115</td>
<td>0.019</td>
<td>0.063</td>
<td>0.108</td>
</tr>
<tr>
<td>5-minutes</td>
<td>78</td>
<td>0.026</td>
<td>0.063</td>
<td>0.112</td>
<td>0.002</td>
<td>0.028</td>
<td>0.081</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.020</td>
<td>0.052</td>
<td>0.097</td>
<td>0.000</td>
<td>0.019</td>
<td>0.066</td>
</tr>
</tbody>
</table>

Panel B: Under the null of mutual excitation ($\xi = 12$)

<table>
<thead>
<tr>
<th>Frequency $\Delta t$</th>
<th># of obs.</th>
<th>1%</th>
<th>5%</th>
<th>10%</th>
<th>1%</th>
<th>5%</th>
<th>10%</th>
</tr>
</thead>
<tbody>
<tr>
<td>5-seconds</td>
<td>4680</td>
<td>0.020</td>
<td>0.030</td>
<td>0.056</td>
<td>0.022</td>
<td>0.035</td>
<td>0.063</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.020</td>
<td>0.030</td>
<td>0.056</td>
<td>0.008</td>
<td>0.015</td>
<td>0.015</td>
</tr>
<tr>
<td>10-seconds</td>
<td>2340</td>
<td>0.040</td>
<td>0.054</td>
<td>0.073</td>
<td>0.023</td>
<td>0.038</td>
<td>0.058</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.040</td>
<td>0.054</td>
<td>0.073</td>
<td>0.007</td>
<td>0.014</td>
<td>0.022</td>
</tr>
<tr>
<td>30-seconds</td>
<td>780</td>
<td>0.076</td>
<td>0.084</td>
<td>0.100</td>
<td>0.030</td>
<td>0.048</td>
<td>0.066</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.076</td>
<td>0.084</td>
<td>0.100</td>
<td>0.003</td>
<td>0.012</td>
<td>0.014</td>
</tr>
<tr>
<td>1-minute</td>
<td>390</td>
<td>0.100</td>
<td>0.106</td>
<td>0.119</td>
<td>0.028</td>
<td>0.045</td>
<td>0.055</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.100</td>
<td>0.106</td>
<td>0.119</td>
<td>0.000</td>
<td>0.009</td>
<td>0.010</td>
</tr>
<tr>
<td>5-minutes</td>
<td>78</td>
<td>0.116</td>
<td>0.128</td>
<td>0.136</td>
<td>0.020</td>
<td>0.026</td>
<td>0.030</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.116</td>
<td>0.128</td>
<td>0.136</td>
<td>0.000</td>
<td>0.001</td>
<td>0.001</td>
</tr>
</tbody>
</table>

Panel C: Under the null of mutual excitation ($\xi = 50$)

<table>
<thead>
<tr>
<th>Frequency $\Delta t$</th>
<th># of obs.</th>
<th>1%</th>
<th>5%</th>
<th>10%</th>
<th>1%</th>
<th>5%</th>
<th>10%</th>
</tr>
</thead>
<tbody>
<tr>
<td>5-seconds</td>
<td>4680</td>
<td>0.017</td>
<td>0.044</td>
<td>0.101</td>
<td>0.005</td>
<td>0.023</td>
<td>0.071</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.017</td>
<td>0.044</td>
<td>0.101</td>
<td>0.012</td>
<td>0.037</td>
<td>0.091</td>
</tr>
<tr>
<td>10-seconds</td>
<td>2340</td>
<td>0.010</td>
<td>0.034</td>
<td>0.086</td>
<td>0.007</td>
<td>0.025</td>
<td>0.066</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.010</td>
<td>0.034</td>
<td>0.086</td>
<td>0.009</td>
<td>0.032</td>
<td>0.078</td>
</tr>
<tr>
<td>30-seconds</td>
<td>780</td>
<td>0.011</td>
<td>0.026</td>
<td>0.044</td>
<td>0.011</td>
<td>0.023</td>
<td>0.045</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.011</td>
<td>0.026</td>
<td>0.044</td>
<td>0.010</td>
<td>0.022</td>
<td>0.044</td>
</tr>
<tr>
<td>1-minute</td>
<td>390</td>
<td>0.011</td>
<td>0.018</td>
<td>0.042</td>
<td>0.011</td>
<td>0.023</td>
<td>0.041</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.011</td>
<td>0.018</td>
<td>0.042</td>
<td>0.003</td>
<td>0.011</td>
<td>0.028</td>
</tr>
<tr>
<td>5-minutes</td>
<td>78</td>
<td>0.015</td>
<td>0.023</td>
<td>0.042</td>
<td>0.005</td>
<td>0.009</td>
<td>0.022</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.015</td>
<td>0.023</td>
<td>0.042</td>
<td>0.000</td>
<td>0.002</td>
<td>0.013</td>
</tr>
</tbody>
</table>

Notes: The table presents the size of the mutual excitation tests. We test for the mutual excitation from $X_1$ to $X_2$. Panel A reports the frequency of rejections in simulations in the absence of mutually exciting jumps (i.e., when $\xi = 0$). Panels B and C in the table report the rejection rates in the presence of mutually exciting jumps (i.e., when $\xi = 12$ and $\xi = 50$, respectively). Under the null of “no mutual excitation”, we use the test statistic given by Equation (19) in the main text with the critical region (21). Under the null of “mutual excitation”, we use the test statistic given by Equation (24) with the critical region (25). The results reported in the first and second rows are based on the choice of $H = H(0; 1)$ and $H = H(6; 1)$, respectively. Columns three to five (six to eight) report the rejection rates for continuous (discontinuous) form of the function $g(\cdot)$ given in (17).
### Table 2: The power of the mutual excitation tests

#### Panel A: Under the null of no mutual excitation ($\xi = 0$)

<table>
<thead>
<tr>
<th>Frequency $\Delta t$</th>
<th># of obs.</th>
<th>1%</th>
<th>5%</th>
<th>10%</th>
<th>1%</th>
<th>5%</th>
<th>10%</th>
</tr>
</thead>
<tbody>
<tr>
<td>5-seconds</td>
<td>4680</td>
<td>0.999</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.999</td>
<td>1.000</td>
<td>1.000</td>
<td>0.998</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>10-seconds</td>
<td>2340</td>
<td>0.985</td>
<td>0.993</td>
<td>0.997</td>
<td>0.981</td>
<td>0.990</td>
<td>0.993</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.946</td>
<td>0.982</td>
<td>0.990</td>
<td>0.954</td>
<td>0.988</td>
<td>0.995</td>
</tr>
<tr>
<td>30-seconds</td>
<td>780</td>
<td>0.713</td>
<td>0.839</td>
<td>0.889</td>
<td>0.669</td>
<td>0.800</td>
<td>0.856</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.454</td>
<td>0.713</td>
<td>0.793</td>
<td>0.482</td>
<td>0.706</td>
<td>0.813</td>
</tr>
<tr>
<td>1-minute</td>
<td>390</td>
<td>0.280</td>
<td>0.527</td>
<td>0.664</td>
<td>0.296</td>
<td>0.469</td>
<td>0.584</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.101</td>
<td>0.316</td>
<td>0.471</td>
<td>0.088</td>
<td>0.311</td>
<td>0.463</td>
</tr>
<tr>
<td>5-minutes</td>
<td>78</td>
<td>0.098</td>
<td>0.116</td>
<td>0.141</td>
<td>0.044</td>
<td>0.072</td>
<td>0.092</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.116</td>
<td>0.131</td>
<td>0.148</td>
<td>0.020</td>
<td>0.023</td>
<td>0.030</td>
</tr>
</tbody>
</table>

#### Panel B: Under the null of mutual excitation ($\xi = 12$)

<table>
<thead>
<tr>
<th>Frequency $\Delta t$</th>
<th># of obs.</th>
<th>1%</th>
<th>5%</th>
<th>10%</th>
<th>1%</th>
<th>5%</th>
<th>10%</th>
</tr>
</thead>
<tbody>
<tr>
<td>5-seconds</td>
<td>4680</td>
<td>0.064</td>
<td>0.228</td>
<td>0.341</td>
<td>0.062</td>
<td>0.212</td>
<td>0.354</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.064</td>
<td>0.228</td>
<td>0.341</td>
<td>0.062</td>
<td>0.212</td>
<td>0.354</td>
</tr>
<tr>
<td>10-seconds</td>
<td>2340</td>
<td>0.053</td>
<td>0.190</td>
<td>0.317</td>
<td>0.054</td>
<td>0.187</td>
<td>0.317</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.053</td>
<td>0.190</td>
<td>0.317</td>
<td>0.054</td>
<td>0.187</td>
<td>0.317</td>
</tr>
<tr>
<td>30-seconds</td>
<td>780</td>
<td>0.013</td>
<td>0.137</td>
<td>0.285</td>
<td>0.017</td>
<td>0.145</td>
<td>0.319</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.013</td>
<td>0.137</td>
<td>0.285</td>
<td>0.017</td>
<td>0.145</td>
<td>0.319</td>
</tr>
<tr>
<td>1-minute</td>
<td>390</td>
<td>0.006</td>
<td>0.096</td>
<td>0.239</td>
<td>0.004</td>
<td>0.105</td>
<td>0.266</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.006</td>
<td>0.096</td>
<td>0.239</td>
<td>0.004</td>
<td>0.105</td>
<td>0.266</td>
</tr>
<tr>
<td>5-minutes</td>
<td>78</td>
<td>0.024</td>
<td>0.087</td>
<td>0.221</td>
<td>0.001</td>
<td>0.051</td>
<td>0.175</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.024</td>
<td>0.087</td>
<td>0.221</td>
<td>0.001</td>
<td>0.051</td>
<td>0.175</td>
</tr>
</tbody>
</table>

#### Panel C: Under the null of mutual excitation ($\xi = 50$)

<table>
<thead>
<tr>
<th>Frequency $\Delta t$</th>
<th># of obs.</th>
<th>1%</th>
<th>5%</th>
<th>10%</th>
<th>1%</th>
<th>5%</th>
<th>10%</th>
</tr>
</thead>
<tbody>
<tr>
<td>5-seconds</td>
<td>4680</td>
<td>0.661</td>
<td>0.855</td>
<td>0.921</td>
<td>0.641</td>
<td>0.834</td>
<td>0.906</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.661</td>
<td>0.855</td>
<td>0.921</td>
<td>0.641</td>
<td>0.834</td>
<td>0.906</td>
</tr>
<tr>
<td>10-seconds</td>
<td>2340</td>
<td>0.561</td>
<td>0.814</td>
<td>0.896</td>
<td>0.524</td>
<td>0.804</td>
<td>0.895</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.561</td>
<td>0.814</td>
<td>0.896</td>
<td>0.524</td>
<td>0.804</td>
<td>0.895</td>
</tr>
<tr>
<td>30-seconds</td>
<td>780</td>
<td>0.424</td>
<td>0.777</td>
<td>0.881</td>
<td>0.399</td>
<td>0.746</td>
<td>0.869</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.424</td>
<td>0.777</td>
<td>0.881</td>
<td>0.399</td>
<td>0.746</td>
<td>0.869</td>
</tr>
<tr>
<td>1-minute</td>
<td>390</td>
<td>0.374</td>
<td>0.727</td>
<td>0.846</td>
<td>0.351</td>
<td>0.689</td>
<td>0.851</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.374</td>
<td>0.727</td>
<td>0.846</td>
<td>0.351</td>
<td>0.689</td>
<td>0.851</td>
</tr>
<tr>
<td>5-minutes</td>
<td>78</td>
<td>0.270</td>
<td>0.692</td>
<td>0.830</td>
<td>0.208</td>
<td>0.619</td>
<td>0.795</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.270</td>
<td>0.692</td>
<td>0.830</td>
<td>0.208</td>
<td>0.619</td>
<td>0.795</td>
</tr>
</tbody>
</table>

Notes: The table presents the power of the mutual excitation tests. We test for the mutual excitation from $X_1$ to $X_2$. Panel A reports the frequency of rejections in simulations in the absence of mutually exciting jumps (i.e., when $\xi = 0$). Panels B and C in the table report the rejection rates in the presence of mutually exciting jumps (i.e., when $\xi = 12$ and $\xi = 50$, respectively). Under the null of "no mutual excitation", we use the test statistic given by Equation (24) with the critical region (25). Under the null of "mutual excitation", we use the test statistic given by Equation (19) in the main text with the critical region (21). As in Table 1, the results reported in the first and second rows are based on the choice of $H = H(0; 1)$ and $H = H(6; 1)$, respectively. Columns three to five (six to eight) report the rejection rates for continuous (discontinuous) form of the function $g(\cdot)$ given in (17).
Table 3: The size and power of the tests for financial flights

<table>
<thead>
<tr>
<th>Case I</th>
<th>1%</th>
<th>5%</th>
<th>10%</th>
<th>1%</th>
<th>5%</th>
<th>10%</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.012</td>
<td>0.062</td>
<td>0.110</td>
<td>0.014</td>
<td>0.058</td>
<td>0.113</td>
</tr>
<tr>
<td></td>
<td>0.017</td>
<td>0.058</td>
<td>0.108</td>
<td>0.010</td>
<td>0.050</td>
<td>0.100</td>
</tr>
<tr>
<td>Case II</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.014</td>
<td>0.041</td>
<td>0.085</td>
<td>0.015</td>
<td>0.048</td>
<td>0.096</td>
</tr>
<tr>
<td></td>
<td>0.008</td>
<td>0.045</td>
<td>0.092</td>
<td>0.002</td>
<td>0.015</td>
<td>0.030</td>
</tr>
<tr>
<td>Case III</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.010</td>
<td>0.054</td>
<td>0.116</td>
<td>0.010</td>
<td>0.060</td>
<td>0.113</td>
</tr>
<tr>
<td></td>
<td>0.010</td>
<td>0.048</td>
<td>0.101</td>
<td>0.010</td>
<td>0.044</td>
<td>0.103</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Case I</th>
<th>1%</th>
<th>5%</th>
<th>10%</th>
<th>1%</th>
<th>5%</th>
<th>10%</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.993</td>
<td>0.998</td>
<td>1.000</td>
<td>0.995</td>
<td>0.999</td>
<td>1.000</td>
</tr>
<tr>
<td></td>
<td>0.980</td>
<td>0.993</td>
<td>0.997</td>
<td>0.981</td>
<td>0.990</td>
<td>0.999</td>
</tr>
<tr>
<td>Case II</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.995</td>
<td>0.999</td>
<td>0.999</td>
<td>0.993</td>
<td>0.998</td>
<td>0.998</td>
</tr>
<tr>
<td></td>
<td>0.976</td>
<td>0.993</td>
<td>0.999</td>
<td>0.976</td>
<td>0.992</td>
<td>0.998</td>
</tr>
<tr>
<td>Case III</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.993</td>
<td>0.998</td>
<td>0.998</td>
<td>0.994</td>
<td>0.998</td>
<td>0.998</td>
</tr>
<tr>
<td></td>
<td>0.977</td>
<td>0.992</td>
<td>0.998</td>
<td>0.979</td>
<td>0.992</td>
<td>0.996</td>
</tr>
</tbody>
</table>

Notes: The table presents the size and power of the excitation tests for financial flights. Panels A and B report the frequency of rejections in simulations. **Case I (FTS/FTQ identification):** testing whether negative jumps in asset $X_1$ excite positive jumps in $X_2$. **Case II (SRS identification):** testing whether negative jumps in asset $X_2$ excite positive jumps in $X_1$. **Case III (bull market flight identification):** testing whether positive jumps in asset $X_1$ excite positive jumps in $X_2$. For all cases, we consider the null hypothesis of no financial flight (i.e., $\xi = 0$). The sampling frequency is 5-seconds. For the size assessment of the test (Panel A), we use the test statistic (19) in the main text with the critical region (21). For the power assessment (Panel B), we use the test statistic given by Equation (24) with the critical region (25). The results reported in the first and second rows are based on the choice of $H = H(0; 1)$ and $H = H(6; 1)$, respectively. Columns two to four (five to seven) report the rejection rates for continuous (discontinuous) form of the function $g(\cdot)$ given in (17).
Table 4: Testing for the mutual excitation from S&P 500 to Gold

<table>
<thead>
<tr>
<th>α</th>
<th>FTS level</th>
<th>Jump size</th>
<th>Flight Cat. (1) (excitation)</th>
<th>Flight Cat. (2) (no excitation)</th>
<th>Flight Cat. (3) (rejecting nulls)</th>
<th>Flight Cat. (4) (accepting nulls)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>weak</td>
<td>0.001</td>
<td>0.00%</td>
<td>59.26%</td>
<td>33.33%</td>
<td>7.41%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3.70%</td>
<td>44.44%</td>
<td>40.74%</td>
<td>11.11%</td>
</tr>
<tr>
<td></td>
<td>mild</td>
<td>0.005</td>
<td>25.93%</td>
<td>14.81%</td>
<td>11.11%</td>
<td>48.15%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>25.93%</td>
<td>11.11%</td>
<td>14.81%</td>
<td>48.15%</td>
</tr>
<tr>
<td></td>
<td>severe</td>
<td>0.009</td>
<td>20.83%</td>
<td>12.50%</td>
<td>12.50%</td>
<td>54.17%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>20.83%</td>
<td>12.50%</td>
<td>12.50%</td>
<td>54.17%</td>
</tr>
<tr>
<td>3</td>
<td>weak</td>
<td>0.001</td>
<td>3.70%</td>
<td>44.44%</td>
<td>37.04%</td>
<td>14.81%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>14.81%</td>
<td>29.63%</td>
<td>29.63%</td>
<td>25.93%</td>
</tr>
<tr>
<td></td>
<td>mild</td>
<td>0.005</td>
<td>25.93%</td>
<td>11.11%</td>
<td>14.81%</td>
<td>48.15%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>25.93%</td>
<td>11.11%</td>
<td>14.81%</td>
<td>48.15%</td>
</tr>
<tr>
<td></td>
<td>severe</td>
<td>0.009</td>
<td>20.83%</td>
<td>12.50%</td>
<td>12.50%</td>
<td>54.17%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>20.83%</td>
<td>12.50%</td>
<td>12.50%</td>
<td>54.17%</td>
</tr>
<tr>
<td>4</td>
<td>weak</td>
<td>0.001</td>
<td>14.81%</td>
<td>25.93%</td>
<td>29.63%</td>
<td>29.63%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>29.63%</td>
<td>18.52%</td>
<td>14.81%</td>
<td>37.04%</td>
</tr>
<tr>
<td></td>
<td>mild</td>
<td>0.005</td>
<td>25.93%</td>
<td>11.11%</td>
<td>14.81%</td>
<td>48.15%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>33.33%</td>
<td>11.11%</td>
<td>11.11%</td>
<td>44.44%</td>
</tr>
<tr>
<td></td>
<td>severe</td>
<td>0.009</td>
<td>20.83%</td>
<td>12.50%</td>
<td>12.50%</td>
<td>54.17%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>25.00%</td>
<td>12.50%</td>
<td>8.33%</td>
<td>54.17%</td>
</tr>
<tr>
<td>5</td>
<td>weak</td>
<td>0.001</td>
<td>25.93%</td>
<td>22.22%</td>
<td>14.81%</td>
<td>37.04%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>25.93%</td>
<td>14.81%</td>
<td>14.81%</td>
<td>44.44%</td>
</tr>
<tr>
<td></td>
<td>mild</td>
<td>0.005</td>
<td>29.63%</td>
<td>11.11%</td>
<td>11.11%</td>
<td>48.15%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>29.63%</td>
<td>7.41%</td>
<td>11.11%</td>
<td>51.85%</td>
</tr>
<tr>
<td></td>
<td>severe</td>
<td>0.009</td>
<td>25.00%</td>
<td>12.50%</td>
<td>8.33%</td>
<td>54.17%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>20.83%</td>
<td>12.50%</td>
<td>8.33%</td>
<td>58.33%</td>
</tr>
</tbody>
</table>

Notes: The table presents the test results of the mutual excitation in jumps from asset X (S&P 500) to asset Y (Gold). The first and second columns indicate various diffusion threshold levels (i.e., α) and degree of flights based on different jump sizes (weak/mild/severe), respectively. We consider four categories of testing procedures. Cat. (1): rejecting the null of no mutual excitation and accepting the null of mutual excitation, Cat. (2): accepting the null of no mutual excitation and rejecting the null of mutual excitation, Cat. (3): rejecting the null of no mutual excitation and rejecting the null of mutual excitation, and Cat. (4): accepting the null of no mutual excitation and accepting the null of mutual excitation. The table reports the frequencies (in %) of trading quarters (i.e. 12 weeks) falling into each category. We set $H = H(0; 1)$ and $H = H(6; 1)$ in upper and lower row frequencies, respectively. The sample covers January 1, 2007 to December 31, 2013 and the sampling frequency is 5-minutes. The significance level of the mutual excitation tests is 0.05.
Table 5: FTS and SRS episodes detected by the mutual excitation tests

<table>
<thead>
<tr>
<th>Years</th>
<th>FTS (1)</th>
<th>FTS (2)</th>
<th>FTS (1)</th>
<th>FTS (2)</th>
<th>SRS (1)</th>
<th>SRS (2)</th>
<th>SRS (1)</th>
<th>SRS (2)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\alpha = 5$</td>
<td>$\alpha = 5$</td>
<td>$\alpha = 2$</td>
<td>$\alpha = 2$</td>
<td>$\alpha = 5$</td>
<td>$\alpha = 5$</td>
<td>$\alpha = 2$</td>
<td>$\alpha = 2$</td>
</tr>
<tr>
<td>2007/Q1</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>2007/Q2</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>2007/Q3</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>2007/Q4</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>2008/Q1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2008/Q2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2008/Q3</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>2008/Q4</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>2009/Q1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2009/Q2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2009/Q3</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>2009/Q4</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>2010/Q1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>2010/Q2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>2010/Q3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>2010/Q4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>2011/Q1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>2011/Q2</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>2011/Q3</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>2011/Q4</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>2012/Q1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>2012/Q2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>2012/Q3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>2012/Q4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>2013/Q1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>2013/Q2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>2013/Q3</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

Notes: The table presents the flight-to-safety (FTS) and seeking-returns-strategy (SRS) episodes identified by the mutual excitation tests. We report the results for two different types of jump thresholds, $\alpha = 5$ and $\alpha = 2$. We set the jump size as 0.001. "✓" denotes the episode of the financial flight detected when the test statistics exceed the critical levels (1.96). On the table, (1) and (2) correspond to measures $H = H(0; 1)$ and $H = H(6; 1)$ used in the estimations, respectively. For all tests, we consider the null hypothesis of no mutual excitation. The sample covers January 1, 2007 to December 31, 2013 and the sampling frequency is 5-minutes. The significance level of the mutual excitation tests is 0.05.
Table 6: Testing for the mutual excitation from S&P 500 to 30-year US Treasury bonds

<table>
<thead>
<tr>
<th>α</th>
<th>FTQ level</th>
<th>Jump size</th>
<th>Flight Cat. (1) (excitation)</th>
<th>Flight Cat. (2) (no excitation)</th>
<th>Flight Cat. (3) (rejecting nulls)</th>
<th>Flight Cat. (4) (accepting nulls)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>weak 0.001</td>
<td></td>
<td>0.00%</td>
<td>92.59%</td>
<td>7.41%</td>
<td>0.00%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.00%</td>
<td>85.19%</td>
<td>7.41%</td>
<td>7.41%</td>
</tr>
<tr>
<td></td>
<td>mild 0.005</td>
<td></td>
<td>3.70%</td>
<td>22.22%</td>
<td>7.41%</td>
<td>66.67%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3.70%</td>
<td>18.52%</td>
<td>7.41%</td>
<td>70.37%</td>
</tr>
<tr>
<td></td>
<td>severe 0.009</td>
<td></td>
<td>8.00%</td>
<td>8.00%</td>
<td>0.00%</td>
<td>84.00%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>8.00%</td>
<td>15.19%</td>
<td>7.41%</td>
<td>67.41%</td>
</tr>
<tr>
<td>3</td>
<td>weak 0.001</td>
<td></td>
<td>0.00%</td>
<td>81.48%</td>
<td>7.41%</td>
<td>11.11%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3.70%</td>
<td>51.15%</td>
<td>11.11%</td>
<td>33.33%</td>
</tr>
<tr>
<td></td>
<td>mild 0.005</td>
<td></td>
<td>3.70%</td>
<td>18.52%</td>
<td>7.41%</td>
<td>70.37%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3.70%</td>
<td>18.52%</td>
<td>7.41%</td>
<td>70.37%</td>
</tr>
<tr>
<td></td>
<td>severe 0.009</td>
<td></td>
<td>8.00%</td>
<td>8.00%</td>
<td>0.00%</td>
<td>84.00%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>8.00%</td>
<td>15.19%</td>
<td>7.41%</td>
<td>67.41%</td>
</tr>
<tr>
<td>4</td>
<td>weak 0.001</td>
<td></td>
<td>0.00%</td>
<td>51.85%</td>
<td>3.70%</td>
<td>44.44%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3.70%</td>
<td>25.93%</td>
<td>7.41%</td>
<td>62.96%</td>
</tr>
<tr>
<td></td>
<td>mild 0.005</td>
<td></td>
<td>3.70%</td>
<td>18.52%</td>
<td>7.41%</td>
<td>70.37%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3.70%</td>
<td>18.52%</td>
<td>7.41%</td>
<td>70.37%</td>
</tr>
<tr>
<td></td>
<td>severe 0.009</td>
<td></td>
<td>8.00%</td>
<td>8.00%</td>
<td>0.00%</td>
<td>84.00%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>8.00%</td>
<td>11.11%</td>
<td>7.41%</td>
<td>77.78%</td>
</tr>
<tr>
<td>5</td>
<td>weak 0.001</td>
<td></td>
<td>7.41%</td>
<td>29.63%</td>
<td>7.41%</td>
<td>55.35%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3.70%</td>
<td>11.11%</td>
<td>7.41%</td>
<td>77.78%</td>
</tr>
<tr>
<td></td>
<td>mild 0.005</td>
<td></td>
<td>3.70%</td>
<td>14.01%</td>
<td>7.41%</td>
<td>74.07%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3.70%</td>
<td>11.11%</td>
<td>7.41%</td>
<td>77.78%</td>
</tr>
<tr>
<td></td>
<td>severe 0.009</td>
<td></td>
<td>8.00%</td>
<td>8.00%</td>
<td>0.00%</td>
<td>84.00%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>8.00%</td>
<td>4.00%</td>
<td>0.00%</td>
<td>88.00%</td>
</tr>
</tbody>
</table>

Notes: The table presents the test results of the mutual excitation in jumps from asset X (S&P 500) to asset Y (30-year US Treasury bond futures). The first and second columns indicate various diffusion threshold levels (i.e. $\alpha$) and degree of flights based on different jump sizes (weak/mild/severe), respectively. We consider four categories of testing procedures. Cat. (1): rejecting the null of no mutual excitation and accepting the null of mutual excitation, Cat. (2): accepting the null of no mutual excitation and rejecting the null of mutual excitation, Cat. (3): rejecting the null of no mutual excitation and rejecting the null of mutual excitation, and Cat. (4): accepting the null of no mutual excitation and accepting the null of mutual excitation. The table reports the frequencies (in %) of trading quarters (i.e. 12 weeks) falling into each category. We set $H = H(0; 1)$ and $H = H(6; 1)$ in upper and lower row frequencies, respectively. The sample covers January 1, 2007 to December 31, 2013 and the sampling frequency is 5-minutes. The significance level of the mutual excitation tests is 0.05.
Table 7: FTQ and SRS episodes detected by the mutual excitation tests

<table>
<thead>
<tr>
<th>Years</th>
<th>FTQ (1)</th>
<th>FTQ (2)</th>
<th>FTQ (1)</th>
<th>FTQ (2)</th>
<th>SRS (1)</th>
<th>SRS (2)</th>
<th>SRS (1)</th>
<th>SRS (2)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>α = 5</td>
<td>α = 5</td>
<td>α = 2</td>
<td>α = 2</td>
<td>α = 5</td>
<td>α = 5</td>
<td>α = 2</td>
<td>α = 2</td>
</tr>
<tr>
<td>2007/Q1</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2007/Q2</td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2007/Q3</td>
<td></td>
<td></td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2007/Q4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2008/Q1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2008/Q2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>2008/Q3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>2008/Q4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>2009/Q1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2009/Q2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2009/Q3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2009/Q4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2010/Q1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2010/Q2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2010/Q3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2010/Q4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2011/Q1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2011/Q2</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2011/Q3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2011/Q4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2012/Q1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2012/Q2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2012/Q3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2012/Q4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2013/Q1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2013/Q2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>2013/Q3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes: The table presents the flight-to-quality (FTQ) and seeking-returns-strategy (SRS) episodes identified by the mutual excitation tests. We report the results for two different types of jump thresholds, α = 5 and α = 2. We set the jump size as 0.001. ✓ denotes the episode of the financial flight detected when the test statistics exceed the critical levels (1.96). On the table, (1) and (2) correspond to measures $H = H(0; 1)$ and $H = H(6; 1)$ used in the estimations, respectively. For all tests, we consider the null hypothesis of no mutual excitation. The sample covers January 1, 2007 to December 31, 2013 and the sampling frequency is 5-minutes. The significance level of the mutual excitation tests is 0.05.
Table 8: Testing for bull market flights using mutual excitation tests

<table>
<thead>
<tr>
<th>Cat. (1) (excitation)</th>
<th>Cat. (2) (no excitation)</th>
<th>Cat. (3) (reject nulls)</th>
<th>Cat. (4) (accept nulls)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Panel A. positive S&amp;P 500 jumps ⇒ positive gold jumps:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.70%</td>
<td>55.56%</td>
<td>3.70%</td>
<td>37.04%</td>
</tr>
<tr>
<td>0.00%</td>
<td>29.63%</td>
<td>7.41%</td>
<td>62.96%</td>
</tr>
<tr>
<td>Panel B. positive S&amp;P 500 jumps ⇒ positive bond jumps:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.70%</td>
<td>37.04%</td>
<td>7.41%</td>
<td>51.85%</td>
</tr>
<tr>
<td>7.41%</td>
<td>7.41%</td>
<td>0.00%</td>
<td>85.19%</td>
</tr>
<tr>
<td>Panel C. positive S&amp;P 500 jumps ⇐ positive gold jumps:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14.81%</td>
<td>14.81%</td>
<td>14.81%</td>
<td>55.56%</td>
</tr>
<tr>
<td>14.81%</td>
<td>3.70%</td>
<td>18.52%</td>
<td>62.96%</td>
</tr>
<tr>
<td>Panel D. positive S&amp;P 500 jumps ⇐ positive bond jumps:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>22.22%</td>
<td>7.41%</td>
<td>11.11%</td>
<td>59.26%</td>
</tr>
<tr>
<td>22.22%</td>
<td>3.70%</td>
<td>14.81%</td>
<td>59.26%</td>
</tr>
</tbody>
</table>

Notes: The table reports the test results of the financial flights in good (bull) market conditions. We set the diffusion threshold as $\alpha = 5$ and the degree of excitation (jump size) is 0.001. The table reports the frequencies (in %) of trading quarters (i.e. 12 weeks) falling into each category. We set $H = H(0; 1)$ and $H = H(6; 1)$ in upper and lower row frequencies, respectively. The sample covers January 1, 2007 to December 31, 2013 and the sampling frequency is 5-minutes. The significance level of the mutual excitation tests is 0.05.

Table 9: Parameter sensitivity of the mutual excitation tests

<table>
<thead>
<tr>
<th>Tail prob.</th>
<th>Excitation prob.</th>
<th>$\beta$</th>
<th>$\lambda$</th>
<th>$\xi$</th>
<th>Size</th>
<th>Power</th>
</tr>
</thead>
<tbody>
<tr>
<td>Base parameters</td>
<td>0.0025</td>
<td>0.80</td>
<td>1.25</td>
<td>20</td>
<td>12</td>
<td>good</td>
</tr>
<tr>
<td></td>
<td>0.0025</td>
<td>0.80</td>
<td>1.25</td>
<td>20</td>
<td>50</td>
<td>good</td>
</tr>
<tr>
<td>Varying tail probability</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sensitivity I</td>
<td><strong>0.010</strong></td>
<td>0.80</td>
<td>1.25</td>
<td>62</td>
<td>37</td>
<td>slightly oversized</td>
</tr>
<tr>
<td>Sensitivity II</td>
<td><strong>0.050</strong></td>
<td>0.80</td>
<td>1.25</td>
<td>223</td>
<td>134</td>
<td>oversized</td>
</tr>
<tr>
<td>Varying tail excitation probability</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sensitivity III</td>
<td>0.0025</td>
<td><strong>0.90</strong></td>
<td>1.25</td>
<td>20</td>
<td>14</td>
<td>good</td>
</tr>
<tr>
<td>Sensitivity IV</td>
<td>0.0025</td>
<td><strong>0.99</strong></td>
<td>1.25</td>
<td>20</td>
<td>15</td>
<td>good</td>
</tr>
<tr>
<td>Varying beta</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sensitivity V</td>
<td>0.0025</td>
<td>0.80</td>
<td><strong>1.50</strong></td>
<td>7</td>
<td>3</td>
<td>incorrect</td>
</tr>
<tr>
<td>Sensitivity VI</td>
<td>0.0025</td>
<td>0.80</td>
<td><strong>1.75</strong></td>
<td>4</td>
<td>2</td>
<td>incorrect</td>
</tr>
</tbody>
</table>

Notes: The table presents the size and power of the mutual excitation tests for different parameter values. Tail prob. is the probability of observing a tail shock in an individual asset (i.e., $X_1$). Excitation prob. is the probability that mutual excitation leads a tail event (i.e., from $X_1$ to $X_2$). $\beta$, $\lambda$ and $\xi$ denote the jump activity index, jump intensity and excitation parameter, respectively. The last two columns report the empirical size and power of the tests.
Notes: Illustration of a jump-type event in the VIX index when House voted against the bailout of Lehman Brothers on September 29, 2008: 5-minute intra-day prices (left panels) and log-returns (right panels). The circles on the right panels indicate a detected jump occurring simultaneously around 18:40 GMT.

Figure 2: S&P 500 index and Comex gold spot prices on August 5-6, 2014

Notes: Illustration of S&P 500 index (white) and Comex gold prices (green) on August 5-6, 2014 when the tension in Middle-east and Ukraine-Russia conflicts increased significantly. Data source and graphics: Bloomberg Analytics.
Figure 3: Size and power of the tests for mutual-excitation with $\xi = 0$

Notes: We set $H = H(0; 1)$ in the upper panels and $H = H(6; 1)$ in the lower panels. The sampling frequency is 5-seconds. The x-axis shows the nominal level of the test, and the y-axis shows the percentage of rejection in the Monte Carlo sample. Red solid and dotted curves (left panels) correspond to excitation from $X_1$ to $X_2$, whereas blue solid and dotted curves (right panels) correspond to excitation from $X_2$ to $X_1$.

Figure 4: Size and power of the tests for mutual-excitation with $\xi = 50$

Notes: We set $H = H(0; 1)$ in the upper panels and $H = H(6; 1)$ in the lower panels. The sampling frequency is 5-seconds. The x-axis shows the nominal level of the test, and the y-axis shows the percentage of rejection in the Monte Carlo sample. Red solid and dotted curves (left panels) correspond to excitation from $X_1$ to $X_2$, whereas blue solid and dotted curves (right panels) correspond to excitation from $X_2$ to $X_1$. 
Figure 5: Size and power of the tests for mutual-excitation with $\xi = 0$ and 1-minute frequency

Notes: We set $H = H(0; 1)$ in the upper panels and $H = H(6; 1)$ in the lower panels. The sampling frequency is 1-minute. The x-axis shows the nominal level of the test, and the y-axis shows the percentage of rejection in the Monte Carlo sample. Red solid and dotted curves (left panels) correspond to excitation from $X_1$ to $X_2$, whereas blue solid and dotted curves (right panels) correspond to excitation from $X_2$ to $X_1$.

Figure 6: Size and power of the tests for mutual-excitation with $\xi = 50$ and 1-minute frequency

Notes: We set $H = H(0; 1)$ in the upper panels and $H = H(6; 1)$ in the lower panels. The sampling frequency is 1-minute. The x-axis shows the nominal level of the test, and the y-axis shows the percentage of rejection in the Monte Carlo sample. Red solid and dotted curves (left panels) correspond to excitation from $X_1$ to $X_2$, whereas blue solid and dotted curves (right panels) correspond to excitation from $X_2$ to $X_1$. 
Notes: Size and power of the mutual excitation tests for financial flights. Left panel (Case I which corresponds to FTS/FTQ identification): testing whether negative jumps in asset \( X_1 \) excite positive jumps in \( X_2 \). Middle panel (Case II which corresponds to SRS identification): testing whether negative jumps in asset \( X_2 \) excite positive jumps in \( X_1 \). Right panel (Case III which corresponds to bull market flight identification): testing whether positive jumps in asset \( X_1 \) excite positive jumps in \( X_2 \). For all cases, we consider the null hypothesis of no financial flight (i.e., \( \xi = 0 \)). The sampling frequency is 5-seconds. The x-axis shows the nominal level of the test, and the y-axis shows the percentage of rejection in the Monte Carlo sample. The left and right panels correspond to excitation from \( X_1 \) to \( X_2 \), whereas the middle panel correspond to excitation from \( X_2 \) to \( X_1 \). In all panels, we set \( H = H(0; 1) \). Solid and dotted curves show the size and power of the tests, respectively.

Notes: Empirical densities and QQ plots of S&P 500 (periodicity-filtered) log-returns in different periods of intraday trading: morning (left panels), afternoon (middle panels) and evening (right panels). The sample covers January 1, 2007 to December 31, 2013 and the sampling frequency is 5-minutes.
Notes: Empirical densities and QQ plots of S&P 500 log-returns in different periods of intraday trading: morning (left panels), afternoon (middle panels) and evening (right panels). The sample covers January 1, 2007 to December 31, 2013 and the sampling frequency is 5-minutes.
Figure 10: Volatility signature plots and microstructure noise function

Figure 11: Parameter sensitivity of size and power

Notes: Size and power of the tests for different parameter values reported in Table 9. We set $H = H(0; 1)$ in all panels. The sampling frequency is 5-seconds. The x-axis shows the nominal level of the test, and the y-axis shows the percentage of rejection in the Monte Carlo sample. Solid and dotted curves correspond to excitation from $X_1$ to $X_2$ (a) under the null of no mutual excitation and (b) under the null of excitation, respectively. See Table 9 for the parameters chosen in each sensitivity analysis, from I to VI.
Figure 12: FTS episodes from S&P500 to Gold

Notes: Detected flight-to-safety (mutual excitation) episodes from S&P500 stock index ($X_1$) to gold ($X_2$). The figure plots the test statistics (given by Equation (19) in the main text) over the full sample (January 1, 2007 to December 31, 2013). X-axis denotes the dates in the sample. Extending our baseline study (relying on 5-minutes frequency and quarterly horizons), we consider weekly horizons with 5-seconds sampling frequency in this analysis. We set $\alpha = 5$ and jump size to 0.001 (0.009) in the upper (lower) panel. The first and second shaded areas represent the BNP crisis (August 2007) and the NBER-defined recession period (December 2007 – June 2009), respectively. For computational reasons, we identify the periods when positive $X_2$ jump excite negative $X_1$ jumps.
Figure 13: FTQ episodes from S&P500 to 30-Year U.S. Treasury Bond

Notes: Detected flight-to-quality (mutual excitation) episodes from S&P500 stock index ($X_1$) to 30-Year U.S. Treasury Bonds ($X_2$). The figure plots the test statistics (given by Equation (19) in the main text) over the full sample (January 1, 2007 to December 31, 2013). X-axis denotes the dates in the sample. Extending our baseline study (relying on 5-minutes frequency and quarterly horizons), we consider weekly horizons with 5-seconds sampling frequency in this analysis. We set $\alpha = 5$ and jump size to 0.001 (0.003) in the upper (lower) panel. The first and second shaded areas represent the BNP crisis (August 2007) and the NBER-defined recession period (December 2007 – June 2009), respectively.
Notes: Detected flight-to-quality (mutual excitation) episodes from long dated U.S. Treasury Bonds ($X_1$) to short-dated U.S. Treasury Bonds ($X_2$) with different maturities (from 2-year to 30-year). The figure plots the test statistics (given by Equation (19) in the main text) over the full sample (January 1, 2007 to December 31, 2013). X-axis denotes the dates in the sample. The sampling frequency is 5-seconds and estimation horizon is weekly. When computing the test statistics, we set $\alpha = 5$ (i.e., FTQ level) and jump size to 0.001 in all panels.