working paper

11 30

The Financial Stress Index: Identification of Systemic Risk Conditions

Mikhail V. Oet, Ryan Eiben, Timothy Bianco, Dieter Gramlich, and Stephen J. Ong

FEDERAL RESERVE BANK OF CLEVELAND

Working Paper 11-30 November 2011*

The Financial Stress Index: Identification of Systemic Risk Conditions

Mikhail V. Oet, Ryan Eiben, Timothy Bianco,
Dieter Gramlich, and Stephen J. Ong

This paper develops a financial stress index for the United States, the Cleveland Financial Stress Index (CFSI), which provides a continuous signal of financial stress and broad coverage of the areas that could indicate it. The index is based on daily public-market data collected from four sectors of the financial markets—the credit, foreign exchange, equity, and interbank markets. A dynamic weighting method is employed to capture changes in the relative importance of these four sectors as they occur. In addition, the design of the index allows the origin of the stress to be identified. We compare the CFSI to alternative indexes, using a detailed benchmarking methodology, and show how the CFSI can be applied to systemic stress monitoring and early warning system design. To that end, we investigate alternative stress-signaling thresholds and frequency regimes and then establish optimal frequencies for filtering out market noise and idiosyncratic episodes. Finally, we quantify a powerful CFSI-based rating system that assigns a probability of systemic stress to ranges of CFSI outcomes.

Keywords: Systemic risk; financial stress; financial stress index; early warning system.

JEL classification: G01, C22, C25, C52, C53, E32, E37.

*Original version: August 2009. Second version: October 28, 2011.
This version: November 14, 2011.

Mikhail V. Oet is at the Federal Reserve Bank of Cleveland (mikhail.oet@clev.frb.org); Ryan Eiben is at Indiana University-Bloomington (reiben@indiana.edu); Timothy Bianco is at the Federal Reserve Bank of Cleveland (timothy.bianco@clev.frb.org); Dieter Gramlich is at Baden-Wuerttemberg Cooperative State University (gramlich@dhbw-heidenheim.de); and Stephen J. Ong is at the Federal Reserve Bank of Cleveland (stephen.ong@clev.frb.org). The authors would like to thank Joseph Haubrich, Ben Craig, and Mark Schweitzer for constructive guidance and formative suggestions; James Thomson, Mark Sniderman, Viral Acharya, John Schindler, and Myong-Hun Chang for valuable comments; and participants of the 2010 Deutsche Bundesbank/ Technische Universitat Dresden conference, "Beyond the Financial Crisis," particularly Andreas Jobst and Marcella Lucchetta; the 2010 Committee on Financial Structure and Regulation, particularly Gustavo Suarez and William Keeton; the 2010 Federal Regulatory Interagency Risk Quantification Forum, particularly Steven Burton, William Lang, Evan Sekeris, Christopher Henderson, and Scott Chastain; Federal Reserve Bank of Cleveland Research Department seminars; the Risk-Central Banking New York seminar, "Managing Systemic Risk in Financial Institutions"; the Federal Reserve Bank of Chicago 2009 Capital Markets Conference; and the NBER–Federal Reserve Bank of Cleveland's Research Conference on Quantifying Systemic Risk. They thank Jing Wang, Juan Calzada, and Kent Cherny for data, research assistance, and helpful insights.

Contents

1. Introduction ... 3

2. Identifying systemic risk conditions ... 6

2.1. Historical precedents for identifying systemic risk ... 6
2.2. Do precedents provide sufficient identification for supervisory objectives? ... 8
2.3. Do precedents provide crisis dating sufficient for supervisory objectives? ... 11
2.4. Defining systemic risk through financial stress ... 11
2.5. Supervisory applications: Identifying systemic stress and distress severity ... 14

3. Index construction ... 17

3.1. Description of the data ... 18
3.2. Variable selection – x_{jt} ... 19
3.3. Variable transformation – $\int_{-\infty}^{x_j} f(x_{jt})dx_{jt}$ ... 23
3.4. Variable weighting – w_{jt} ... 24

4. Results and selection of CFSI weighting method ... 27

4.1. Comparison of candidate weighting methods ... 27
4.2. Comparison of candidate weighting methods with expert survey ... 28
4.3. Selection of CFSI from candidate weighting methods ... 31

5. CFSI results ... 37

5.1. Dependability ... 39
5.2. CFSI robustness ... 39
5.3. Decomposition of CFSI ... 40

6. Discussion of the results ... 45

6.1. Duration pattern of systemic stress episodes ... 45
6.2. Comparisons of CFSI with currently available alternates ... 50

7. Applications of CFSI ... 53

7.1. Use as a dependent variable for early warning system ... 53
7.2. Monitoring financial stress ... 55
7.3. Crisis dating ... 60

References ... 61

Appendix A. Description of data ... 65

Appendix B. Stationarity of quarterly CFSI ... 67

Appendix C. CFSI properties ... 68

The Financial Stress Index: Identification of Systemic Risk Conditions

1. Introduction

> In every systematic inquiry (*methodos*) where there are first principles, or causes, or elements, knowledge and science result from acquiring knowledge of these; for we think we know something just in case we acquire knowledge of the primary causes, the primary first principles, all the way to the elements. It is clear, then, that in the science of nature as elsewhere, we should try first to determine questions about the first principles. The naturally proper direction of our road is from things better known and clearer to us, to things that are clearer and better known by nature; for the things known to us are not the same as the things known unconditionally (*haplôs*). Hence it is necessary for us to progress, following this procedure, from the things that are less clear by nature, but clearer to us, towards things that are clearer and better known by nature.
>
> ——*Aristotle, Phys. 184a10–21*

In this paper, we describe the design and features of the Cleveland Financial Stress Index (CFSI) series, originally constructed in early 2009. Reviewing precedents in the literature on identifying systemic distress conditions (section 2), we find that theoretical precedents focusing on crises provide insufficient identification for supervisory objectives. We introduce the concept of financial stress in section 2 and discuss CFSI construction in section 3. We show that CFSI adds a number of useful innovations to the literature on financial stress and conditions indexes.[6]

Clarity of construction is among the main contributions of the CFSI approach to alternative series. This clarity is reflected in CFSI's key design features: *modular construction*; *weighting methodology*; and *component market factors.*

Modular construction. CFSI's construction is modular, with clear, tractable financial markets. This allows modular expansion in response to the emergence and availability of important new financial markets.

Weighting methodology. The selection of weighting methodology maintains the clarity of CFSI's construction. In contrast to alternative recent indexes, excepting Illing and Liu, we do not

[6] In 2009, the Federal Reserve Bank of Cleveland's financial stress series was the only choice for developing the SAFE EWS. By the end of 2010, 12 alternative financial stress indexes existed.

pre-select a weighting methodology. Proceeding from first principles, we are skeptical of making a priori choices of indexing method; instead, we test several alternatives. In this, we extend Illing and Liu's exploration of various construction approaches.[7] Section 4 discusses the selection and results of CFSI's weighting method and describes the testing process.

Component market factors. We correct certain inconsistencies in the Illing and Liu construction method and introduce several observable-component market factors that describe new dimensions of stress in the financial markets (section 3).

The CFSI benchmarking process described in section 4 makes four innovative contributions to the literature. The researcher, having established a measure of financial stress, must be able to support the claim that it is dependable and free from specification problems, such as incorrect functional form, omission of relevant explanatory variables, inclusion of redundant variables, etc. Our first contribution to the literature is an innovative, objective method for verifying and benchmarking financial stress. To benchmark financial stress, the method utilizes the related concept of financial risk as volatility. General volatility indexes, which are readily available in most financial markets, represent aggregate measures of market volatility and usually mix financial and nonfinancial firms. The CFSI benchmarking process establishes a set of independent market volatility benchmarks to verify the performance of the CFSI's measure of financial stress. Our second contribution is the use of volatility benchmarks for dating episodes of systemic stress. Our third is the operational definition of systemic risk, which allows implementation of stress-episode dating. Our fourth is a demonstration of the benchmarking method in an expanded interpretation of past U.S. financial episodes and an objective assessment of current observations. Section 4 shows that the construction method for CSFI is optimal under a variety of monitoring cycles that range from weekly to quarterly.

In section 5, which describes our results, we discuss the dependability, robustness, and decomposition of CFSI. We show that with a quarterly series, CFSI dependably filters out idiosyncratic stress episodes, which makes it useful as a dependent variable in a systemic risk EWS. We consider CFSI's time series properties and establish the stationarity of the quarterly series, which is particularly important in the context of a EWS. We then establish a dynamic

[7] In explaining the financial stress index for Canada, Illing and Liu established the weighting-method selection process as a paradigm of search that minimized selection bias among the weighting methods. We are skeptical about several recent papers on index construction that discuss only one weighting method and offer no support for minimizing selection bias.

benchmark model for forecasts of systemic risk conditions using only the CFSI series itself.[8] This benchmark model further extends CFSI's capacity for monitoring and forecasting systemic risk. In addition, it establishes a minimum performance standard for subsequent EWS development.[9] The decomposition of CFSI into its components allows for intriguing interpretations of economic conditions and permits detailed observations of the effects of regulatory measures to reduce systemic risk through specific financial-stress components.

Section 6 discusses some significant CSFI results. First, we consider the evidence for a structural connection between financial deregulation and the pattern of systemic stress episodes. Our evidence suggests that although the frequency of systemic stress episodes remains consistent before and after U.S. financial deregulation, the duration pattern of systemic stress episodes changes. We observe that after deregulation, the speed of systemic stress propagation is reduced (the benefit of risk diversification for individual institutions); however, the length of recovery from systemic stress also slows substantially (the penalty of universal banking). We also compare CSFI to alternative stress indexes.

In section 7, we address various applications of CFSI: EWS, monitoring, and identifying stress episodes. We begin by discussing additional data and technical considerations to improve the use of CSFI as a dependent variable in a systemic-risk EWS. Effective application of a stress index for an early warning of systemic stress hinges on the index capacity to differentiate idiosyncratic risk. A desirable frequency for EWS would minimize the presence of idiosyncratic stress episodes. We proceed, therefore, by investigating the optimal stress-signaling regime and CFSI frequency[10] for use in EWS of systemic risk and monitoring systemic financial stress. Finally, we ask three surveillance and policy questions: how can a CFSI measure be usefully

---

[8] By contrast, Hatzius et al. (2010) and Brave and Butters (2011) seek to monitor and forecast economic activity and to develop financial conditions indexes. These studies confront the critical question of how to distinguish financial stress from the various cyclical effects of economic activity. Hatzius selects, from among the candidates, the index with optimal performance in forecasting economic growth (e.g., GDP). Unlike these concurrent independent studies, we wish to use FSI for forecasts of systemic banking risk rather than forecasts of economic activity. Economic conditions enter CFSI exogenously. The selected CFSI utilizes a credit-weighting method that dynamically reflects changing economic conditions by changing weights.

[9] This is discussed more fully in Oet, Eiben, Gramlich, Miller, and Ong (2011).

[10] Although CFSI is constructed using daily data, monitoring frequency produces different CFSI series, each of which aggregates the data within its monitoring horizon. Increasing the frequency of CFSI monitoring increases the presence of idiosyncratic events within CFSI. Successful application of CFSI as a dependent variable for an early warning system requires the use of an index that can effectively sift out idiosyncratic stresses. We show that a quarterly CFSI possesses this useful quality for an EWS application. It is desirable to maximize the frequency of monitoring, while minimizing idiosyncratic events. We describe the testing and selection of the optimal CFSI frequency from the following monitoring windows: quarterly, FOMC-meeting frequency, monthly, bi-weekly, weekly, and daily.

interpreted: what does it communicate about the probability of systemic stress; and what policy-action thresholds should be considered in its use. To this end, we quantify and test a CSFI-based rating system. In the process, we test the applicability of the Bordo, Dueker, and Wheelock (2000) conjecture on the rating classification of stress thresholds. We conclude by demonstrating the use of CFSI benchmarking for dating systemic stress episodes.

2. Identifying systemic risk conditions

2.1. Historical precedents for identifying systemic risk[11]

In the early 1980s and 1990s, the concepts of systemic risk and systemic crises tended to be synonymous, leading to binary measurement of systemic risk—either crisis or no crisis—and identification relied on professional consensus.[12]

As this period reveals, systemic risk conditions manifest differently in the banking system, in a broader set of financial companies, or in securities and FX markets.[13] Thus, there is obviously some "subjectivity associated with banking crisis identification."[14] Investigating the definitions applied in 13 research studies, Ishihara (2005)[15] finds six different types of financial crises, then defines and measures them individually.[16] Because excessively narrow definitions may lead to inconsistent policies, and crises are increasingly multidimensional, the author suggests a broader concept for conceptualizing and assessing them.

Literature from the 1990s and 2000s focuses on the search for a reassessment and a new definition of systemic risk. Caprio and Klingebiel (1996) and Demirgüç-Kunt and Detragiache (1998) highlight the fact that systemic risk is related to the point at which most of financial firms' capital is exhausted. In their broad survey of systemic risk, De Bandt and Hartmann (2000) define a systemic crisis as an "event that affects a considerable number of financial

---

11 Parts of an earlier version of this section appear in Gramlich, Miller, Oet, and Ong (2010).

12 See, for example, Demirgüc-Kunt and Detragiache (1998) and Kaminsky and Reinhart (1999). Professional consensus is established by precedent and acceptance in the relevant literature.

13 EWSs in finance started in the 1990s with models for predicting currency and national debt crises; specific EWSs for banking system distress have been proposed more recently, for example, by Berg, Borensztein, and Pattillo (2004), pp. 4, 7.

14 Davis and Karim (2008), p. 97.

15 Ishihara (2005), p. 8.

16 The types of crises are: banking liquidity, banking solvency, balance of payments, currency, external debt, growth rate, and financial crisis.

institutions or markets in a strong sense."[17] Another descriptive dimension is introduced by Hendricks, Kambhu, and Mosser (2007),[18] who emphasize disrupted transmission structures as characteristics of systemic crises; in these structures, "systemic risk is the movement from one stable (positive) equilibrium to another stable (negative) equilibrium." The extent and speed of this movement depend mainly on the system's complexity and the shift from classical, bank-based crises to more recent, market-based financial crises. Similarly, Kambhu, Weidman, and Krishnan (2007)[19] refer to systemic risk as a "tendency toward a rapid and large transition from one stable state to another, possibly less favorable, state." They point out that the physical and financial worlds are both characterized by nonlinear, complex adaptive systems.

To avoid the "post-crisis bias" resulting from a false assessment of recovery phases, Bussière and Fratzscher (2002) introduced a three-state classification of crises based on a multinomial logit model. These concepts of systemic risk, however, have several drawbacks. The binary and three-regime approaches ignore market stresses that approached (but never met) crisis standards; they also exclude situations that were successfully managed but might otherwise have become crises.[20]

Consequently, more recent research suggests a richer approach to systemic financial risk as a continuous variable, with crisis as an extreme value. Bordo, Dueker, and Wheelock (2000) develop the concept of "an index of financial conditions" (henceforth FCI), examining whether aggregate price shocks are useful for dating financial instability. Using regime-switching models, the authors measure FCI as an index of destabilized financial conditions based on "unanticipated movements in the aggregate price level or inflation rate."

Further studies extend this scheme. The recent literature includes two alternative approaches with different end-user objectives. In the first, systemic indexes, designed to predict economic conditions, develop as FSIs along the lines of the original Bordo, Dueker, and Wheelock (2000) research. An FCI derives potential financial stress by combining different price vectors on financial markets, principally vectors related to interest rates and equity prices.[21] Here, financial conditions are most often described in terms of deviation from a long-term trend and are

---

[17] De Bandt and Hartmann (2000), p. 11.
[18] P. 65.
[19] P. 6.
[20] The IMF (2009, Responding), p. 145, emphasizes that binary variables do not measure the intensity of the stress.
[21] An overview is given by Swiston (2008), pp. 3–5.

measured in standard deviations from the mean.[22] English, Tsatsaronis, and Zoli (2005), Rosenberg (2008), and Swiston (2008) link FCIs to subsequent bank lending standards and from there to macroeconomic activity and inflation. Financial conditions are thus connected to overall fluctuations in the economy, and FCI is used for predicting economic up- and downturns.

In the second approach, systemic indexes, which pursue the supervisory objective of averting risk manifestations in the financial system, develop along the lines of Illing and Liu (2003, 2006) as financial stress indexes (FSIs). Illing and Liu examine financial stress "as a continuous variable with a spectrum of values, where extreme values are called a crisis," allowing more information to be contained in the stress measure and avoiding some arbitrary boundaries for the beginnings and ends of crises.[23] Exploring systemic risk in Canada from a supervisory perspective, Illing and Liu (2006) provide an overview of different observable variables used to assess crises originating in the banking, foreign exchange, debt, and equity sectors, as well as multi-sector, composite crises. They show how stress measures vary between and within the crisis categories, sometimes referring to more subjective or objective criteria. The authors compare nine differently constructed indicators, concluding that the most appropriate are based on standard observable variables and weighted by the respective sector's "share of total credit in the economy."[24] Their index relies principally on spreads, betas, and interest rates, with the level of financial stress determined as a weighted aggregation of various sub-indexes. Hanschel and Monnin (2005) use the same type of stress index to investigate systemic risk in Switzerland.

In short, the historical precedents in systemic risk identification show the evolution of a debate among academics, policymakers, and financial practitioners. The literature shows that the definition of systemic risk varies according to the underlying objectives.

2.2. Do precedents provide sufficient identification for supervisory objectives?

Selection of a working definition of systemic risk is a prerequisite to identifying and measuring it. From supervisors' point of view, systemic risk may be referred to as the risk of financial institutions' correlated default, strongly affecting the system's risk capital and liquidity,

---

22 For example, Bloomberg uses a set of three vectors—money market rates, bond market spreads, and equity prices—equally weighted and calculated for the 1994–2008 period. See Rosenberg (2008), p. 8.

23 Illing and Liu show that crises in Canada have been influenced by three broad sets of issues: country-specific, and North American issues, as well as issues elsewhere.

24 Illing and Liu (2006), p. 255.

with subsequent negative feedback effects on real markets. Thus, a useful approach in terms of supervisory objectives is to identify systemic risk through a continuous index measure of financial stress. Operationally, a continuous index definition must allow resolution of the ensuing crisis identification problems, specifically the precise timing of episodes and the differentiation of their relative severity.

To establish further the desirable features of systemic risk definition for supervisory objectives, it is useful to confirm exactly why the precedents that do not establish a continuous measure of systemic risk are insufficient. Demirgüç-Kunt and Detragiache (2005) define systemic crises as events in which at least one of four conditions is present: large-scale nationalization occurs; emergency measures are taken to assist the banking system; the cost of the rescue operations equals at least 2 percent of GDP; and non-performing assets equal at least 10 percent of total assets. Somewhat similarly, Laeven and Valencia (2008) define systemic crises as events characterized by the existence of at least one of four occurrences: deposit runs;[25] introduction of deposit freezes or blanket guarantees; liquidity support; and bank interventions.[26] Reinhart and Rogoff (2008) require at least one of two conditions: bank runs,[27] and emergency measures taken to assist the banking system.[28]

If we consider the first set of precedents in light of supervisory EWS objectives, we encounter some serious problems in defining systemic episodes. The definitions include either en-masse bank insolvencies or government interventions, which are inconsistent with the supervisory need for a definition that allows time to avert negative outcomes. In addition, these definitions cannot describe the continuous states of the banking system or differentiate the severity of systemic episodes.

Boyd, De Nicolò, and Loukoianova (2010) offer an additional critique, pointing out that defining systemic risk from a crisis perspective involves mixing economy-driven shocks with governmental responses. If the effects of governmental actions were not integrated, systemic conditions would develop much earlier, and conventional indicators would recognize them too late. To remedy this, Boyd, De Nicolò, and Loukoianova propose measuring systemic bank

---

25 Defined as a monthly percentage decline in deposits exceeding 5 percent.

26 Defined as the ratio of monetary authorities' claims on banks as a fraction of total deposits of at least 5 percent and at least double the ratio compared to the previous year.

27 Defined as the public sector's closure, merging, or takeover of one or more financial institutions.

28 Defined as closure, merging, takeover, or large-scale government assistance of an important financial institution (or group of institutions), marking the start of a string of similar outcomes for other financial institutions.

shocks (SBS) through extremes of observable drops in real lending and deposits. They construct two types of SBS indicators dictated by data availability and distribution across all countries, indexing extreme drops in real lending and bank deposits (25^{th} and 10^{th} percentiles). SBSs are then measured by means of a theoretical model in which banking problems are produced by exogenous shocks to the industry. The authors emphasize the importance of separating systemic conditions from the ensuing governmental response. Significantly, the proposed method for identifying systemic banking shocks utilizes observable institutional data. From a supervisory EWS perspective, this is a considerable improvement. By focusing on systemic banking shocks, the authors address the problems associated with identifying crises. Using observable institutional data reduces the problems posed by long lags between the emergence of observable bank shocks and en masse bank insolvencies. Empirically, however, the study's data limitations and lack of U.S.-specific dating constrain its effectiveness as a supervisory EWS.

The continuous index approach, which is similar to that of Bordo, Dueker, and Wheelock (2000), is more promising because financial conditions are never absolutely good or absolutely bad (that is, they resist binary classification) but are relatively better or worse. Distinguishing the relative degree of criticality as a continuous measure is certainly daunting, but very desirable. A binary system of distress identification makes it difficult to distinguish how much worse one episode is relative to another. A measure of criticality in terms of underlying stress is very useful; it enables a relative assessment of stress episodes. Significantly, this approach allows financial conditions to be measured in continuity without requiring a priori definition of systemic conditions. Instead, once the measure is obtained, a continuity of stress measures can be interpreted using a rating system approach and calibrated by comparison with a historical series. Bordo, Dueker, and Wheelock transform variable data by standardizing the measured distance between each observation and the sub-period median divided by the variable's standard deviation. The overall index is aggregated simply as an unweighted average of these standardized distances across component variables. Although the authors do not offer a definition of systemic conditions, they do provide a methodology for measuring them continuously and suggest a rating-system approach to identification.[29]

[29] In reference to the sub-period mean, the authors proposed the following five-state empirical calibration of FCI: *severe distress* – Zt > 1.5 std, *moderate distress* – Zt >0.75 std, *normal* – -0.75 < Zt > 0.75 std, *moderate expansion* —-1.5 < Zt < –0.75 std, *euphoria* – Zt < –1.5 std.

2.3. Do precedents provide crisis dating sufficient for supervisory objectives?

In addition to the identification of stress, the second difficulty for supervisory purposes involves the use of precedents in the literature for dating stress episodes. This, of course, is a corollary of misalignment of objectives in identifying systemic stress conditions. Unlike the literature aiming to date crisis episodes, supervisors are interested in a nuanced dated series of potential systemic stress. Typically, authors seeking identification through systemic crises recognize only two U.S. episodes since 1980: the savings and loan crisis, generally dated to 1988, and the subprime crisis of 2007; other significant stress episodes[30] are conspicuously missing. The second major limitation of existing studies is their reliance on survey-based crisis dating. This leads to subjective interpretation of dating and tends to define systemic conditions through crisis response. Thus, as Boyd, De Nicolò, and Loukoianova point out, the crisis series tend to identify crises too late.

Again, the continuous index measures fit supervisory objectives better. The Bordo, Dueker, and Wheelock (2000) annual FCI index can differentiate three distinct episodes between 1980 and 2000: severe distress in 1982–86 and two periods of moderate distress: 1981 and 1987–92. One important advantage of this approach is the availability of a historically deep index series. Unfortunately, the annual FCI is sub-optimal because its frequency does not enable supervisors to observe any conditions until all data for the year has been collected. The resulting lag seriously undermines the annual FCI's ability to serve supervisors. Clearly, having more frequent data is essential to inform supervisors ex ante.

2.4. Defining systemic risk through financial stress

Given the supervisory objectives of monitoring systemic banking risk in the financial sector for ex ante actions, a financial stress index approach is more fitting than a financial conditions approach. To further clarify the criteria for selecting appropriate stress index components, it is critical to consider the first principles of financial stress index design. For Illing and Liu, "financial stress is defined as the force exerted on economic agents by uncertainty and changing expectations of loss in financial markets and institutions. Financial stress is a continuum."[31]

---

30 For example, the LTCM crisis and the Asian crisis.

31 Illing and Liu (2006), p. 243.

Therefore, a measure of financial stress—the financial stress index—"is a continuous variable with a spectrum of values, where extreme values are called financial crises." Similarly, Hanschel and Monnin state that the financial stress index "represents a continuum of states, describing the banking sector's condition ranging from low levels of stress, where the banking sector is tranquil, to high levels of stress, where the banking sector is in a severe crisis."[32] Illing and Liu further maintain that "if financial stress is systemic, economic behaviour can be altered sufficiently to have adverse effects on the real economy… Stress increases with expected financial loss, with risk (a widening in the distribution of probable loss), or with uncertainty (lower confidence about the shape of the distribution of probable loss)."[33]

It is important to remember that components of the financial stress index must be directly observable in the markets. They can be explained in terms of loss expectations, risk, and uncertainty, among others, but are not equivalent to risk in its standard computational finance sense of statistical volatility. In pursuing an understanding of systemic risk from first principles, we must include only first observations of it. Economically, financial stresses are *observable,* continuous manifestations of "forces exerted on economic agents." This is a critical point in guiding selection of components of a financial stress index. To proceed, we must now consider what types of observable factors describe stress on economic agents.

There is rich set of theoretical precedents showing the importance of particular spreads in the context of micro- and macroeconomic equilibria. Reviewing the seminal studies on this subject, Freixas and Rochet (2008) discuss the importance of external finance spread for monetary policy transmission, affirmed by both theoretical and empirical studies.[34] The critical role of the external finance spread emerges differently in alternative models. In the Bernanke and Gertler (1990) theoretical study, the importance of the spread for financial fragility emerges from the perspective of investment and agency costs. Bernanke and Gertler (1995) consider various types of spread empirically, according to their role in the credit channel of monetary policy transmission. In Holmström and Tirole (1997), the spread results from "scarcity of bank capital." In Bolton and Freixas (2000), the cause is "adverse selection in the capital markets." As Freixas and Rochet (2008) point out, the spread's key role in various channels of monetary policy

32 Hanschel and Monnin (2005), p. 431.
33 Illing and Liu (2006), p. 244.
34 Freixas and Rochet (2008), p. 198: "[The] external finance premium, defined as the wedge between the cost of funds raised externally and the opportunity cost of internal funds, [is] an essential key in understanding of the transmission mechanism."

transmission results from its amplification effect on interest rates and generating the financial accelerator effect.[35]

A survey of the current literature on continuous indexes reveals a lack of consistency in applying the above theoretical contributions to the selection of index components. Current studies of continuous indexes generally allow use of both spreads and the conceptually similar volatility indexes as index components. We choose *not* to mix these two related types of market stress information. A benefit of this choice is the availability of volatility series for unbiased selection and benchmarking among alternative financial stress indexes. The decision is supported partially by the abovementioned theoretical insights into the importance of spreads and partially by empirical reasoning.

While both volatility indexes and spreads provide observations of market stress, one critique of their concurrent use in constructing a financial stress index is that they provide qualitatively different insights. Volatility indexes blend the prices of many securities. They hide the causal transmission mechanism by which the factors entering the volatility series influence the stress index, making the mechanism a "black box"[36]—that is, only indirectly observable—and obfuscating the stress index. By contrast, spreads are differences between two related securities. Their definition clarifies the transmission mechanism in spread changes. Empirically, it is also interesting to note that a spread-based financial stress index appears to identify stress episodes more quickly than alternative indexes that use mixed methods (see Fig. 15 in section 6.2). In addition, Table 17 in Appendix C shows that a spread-based financial index frequently leads (that is, has stronger one-way Granger causality than) the volatility indexes in the interbank and credit markets.[37]

Based on careful consideration of fit to study objectives, it is possible to proceed empirically and to modify the components found in earlier financial stress index studies. To do so, we begin by considering that the system consists of financial institutions and financial markets, which "exhibit three forms of interaction: competition, complementarity, and co-evolution."[38] What we can observe continuously are stresses on financial institutions in the financial markets. In each

---

35 Ibid.
36 We use the term "black box" in a sense similar to that of Bernanke and Gertler (1995), "Inside the Black Box: The Credit Channel of Monetary Policy Transmission."
37 In the interbank market, we used the MOVE volatility index. In the credit market, we used the LBOX and BBOX volatility indexes.
38 Song and Thakor (2010), p. 1024.

distinct financial market, we can distinguish the typical products involved in financial institutions' interaction and observe the corresponding applied stresses. These applied stresses would generally be spreads—differences between applied economic forces; for example, arbitrage spreads between risky and "risk-less" products and liquidity (bid–ask) spreads capturing differences between supply and demand.

Following the analysis of Illing and Liu, we can distinguish at least four financial markets: interbank, credit, equity, and foreign exchange. Financial institutions access interbank market to seek direct financing (through bank bonds) and indirect financing (through interbank borrowing), and to manage liquidity and interest rate risk. Reasonable measures of stress in the interbank market therefore would consist of spreads capturing pressures on bank bonds, interbank borrowing, and interbank liquidity. In addition, it may be useful to consider overall stress on the interbank market relative to the overall stock market. In the credit market, financial institutions act as intermediaries for short- and long-term borrowing. Thus, measures of stress in the credit market would include spreads capturing pressures on corporate bonds, commercial papers, and the Treasury yield curve, as well as liquidity (bid–ask) pressures on Treasuries. In addition, the relative stress on U.S. vs. international credit markets may be observed through covered interest-rate parity spreads. In the equity market, it is reasonable to include observable measures that describe the extent to which financial equities in the S&P 500 have dropped over the previous year. Similarly, in the foreign exchange market, we can include observable measures of flight from the U.S. dollar toward a set of foreign currencies.

2.5. Supervisory applications: Identifying systemic stress and distress severity

There are two main supervisory applications for a measure of financial stress: first, identifying and dating episodes of systemic stress, and second, determining whether the financial stress level is critical: in other words, whether observed financial stress constitutes a financial crisis.

2.5.1. Identification of systemic stress

As discussed in Gramlich, Miller, Oet, and Ong (2010), the first of these applications, identification of systemic stress, is based on a continuous notion of what conditions constitute financial stress, namely the *transmission of distress*, the *extent of distress*, and the *type of* *distress,* that is, the financial markets involved. Operationally, systemic stress should be identified by these three conditions. More specifically, *transmission of distress* involves selecting observable market characteristics (for example, spreads and betas) and setting a threshold above which these characteristics are considered to be in distress. *Extent of distress* involves selecting the period of time over which the persistence of distress is observed. Finally, *type of distress* involves choosing markets that in combination may be deemed to raise systemic concerns. The choices made operationally may vary to yield more or less sensitive interpretations of systemic stress, resulting in a flexible identification scheme. Here again, identification must fit the overall objectives.

For purposes of supervisory ex ante monitoring, we wish the identification to be reasonably sensitive. From a supervisor's perspective, the sensitivity of the ex ante identification of systemic stress is important because the supervisory remedy would be based on balancing regulatory actions' costs against their benefits. If the supervisor sets the definition of systemic stress too low, the monitoring of systemic stress will be inefficient. Once a condition is assessed as systemic, supervisory resources will be engaged in the cost–benefit analysis. There is also a risk that some conditions that are inherently not truly systemic will be identified as systemic. In this case as well, the supervisory action may be inefficient—that is, the cost of action would be wasted—because markets would otherwise be able to resolve and stabilize the conditions. Worse, the supervisory action itself could be destabilizing if it caused adverse effects and generated undesirable feedback effects, which might be avoided if normal market mechanisms self-correct. By contrast, if supervisors were to identify systemic stress conservatively, then they would fail to understand and act in timely fashion to arrest its evolution.

An example of excessive sensitivity would be the identification of systemic stress on the basis of daily stress fluctuations. Beyond this, it becomes difficult to gauge the right level of sensitivity a priori. The best approach may well be iterative:

1. Select specific choices for *transmission of distress, extent of distress*, and *type of distress*.
2. Evaluate the sensitivity of the resulting systemic stress episode identification. If there are too many false positives, decrease the sensitivity of the identification scheme and re-evaluate.

Exclusion of idiosyncratic events serves as a useful additional constraint on the identification of systemic stress. Setting the extent of distress to a daily or weekly interval would result in a very volatile stress index with too many idiosyncratic stress episodes: market rumors, unsubstantiated fears, political events, etc. As we show in this paper, quarterly extent of distress excludes random and idiosyncratic events and is useful for a supervisory EWS. However, it is possible that this extent of distress is too conservative for the purposes of supervisory monitoring. This line of rationalization helps to constrain the extent of distress somewhere between one week and one quarter.

This reasoning leads us to the initial selection of operational choices for *transmission of distress*, *extent of distress*, and *type of distress* for the purposes of supervisory monitoring using CFSI. For *transmission of distress,* we begin by selecting a distress threshold of one standard deviation above previous quarterly levels. For *extent of distress*, we select two consecutive weeks of distress. For *type of distress*, we focus on distress conditions that affect at least two distinct markets. The choices result in the following operational definition or identification of systemic stress episodes with CFSI: *Systemic stress is characterized by two consecutive weeks of distress above previous quarterly thresholds, or concurrent weekly distress in at least two distinct markets.* The intent of this approach is to set a lower threshold for monitoring systemic stress in order to increase ability to identify, ex ante, episodes that have the potential to become critical.[39]

Applying this scheme results in the identification of 50 systemic stress episodes from the fourth quarter of 1991 to the fourth quarter of 2010. Twenty-five of them were previously identified in the literature; the remaining 25 are market shocks that are known, but typically are not identified in the literature as U.S. financial stress episodes or crises. These episodes include a mix of both systematic and idiosyncratic events,[40] suggesting that the above initial operational definition may be improved further. We discuss the process and results of identification of stress episodes in more detail in section 5.3.

2.5.2. Identifying distress severity

The current literature on continuous index measures, of both financial conditions and financial stress, generally report the indexes as Z_t standardized distances, without providing explicit thresholds for identifying distress severity. Bordo, Dueker, and Wheelock (2000) show

[39] Oet and Eiben (2009).

[40] For example, the Gulf War; $17 billion BCCI collapse; global bond markets reversal of 1994, Japanese bank runs of 1995, etc.

that a continuous index can be used to identify distress severity independent of the dating of systemic conditions, using Z_t standardized distances from the median. They suggest a five-category differentiation of distress: “severe distress,” “moderate distress,” “normal,” “moderate expansion,” and “euphoria.”[41] Certainly, the use of a common scale facilitates comparison among alternative indexes. Establishing thresholds to identify distress severity is an even more useful application of the standardized-distance method of index measurement. In section 5.6 of this paper, we extend Bordo, Dueker, and Wheelock’s idea for measuring distress severity using a probit model of CFSI to calibrate the optimal system for rating the severity of distress.

3. Index construction

Constructing an indicator of coincident stress is, to put it succinctly, a formidable task. A well-constructed financial stress index should be meaningful, not only as a monitoring tool but also within a larger EWS. By contrast, a noisy FSI will prove utterly useless to a supervisor concerned with monitoring systemic risk.

When trying to classify a financial crisis, one necessarily contemplates the extent to which a host of market variables deviate from some long-term trend. The notion that an abnormal event in market A could cause substantial deviations in values within market B (the contagion effect) necessarily complicates matters by introducing a feedback effect. This makes the identification of leading stress indicators difficult in that the researcher’s choice of indicators may move in response to propagation events rather than as the first rumblings of distress. Moreover, in the normal course of business-cycle fluctuations, the regulator expects some movement in market indicators as part of rational resource allocation. The researcher must be wary of selecting indicators that move regularly with the business cycle.

Having selected appropriate indicators, the researcher hopes to aggregate them into a single financial stress index (FSI). A useful framework for accomplishing this task considers two factors for each indicator: what is the precedent set by the indicator’s value and how much does that precedent matter. One may generate an FSI using the equation

---

[41] Bordo, Dueker, and Wheelock (2000, p. 27) “assign Zt larger than 1.5 standard deviations above the sub-period mean to the “severe distress” category; Zt larger than 0.75 standard deviation above the sub-period mean to the “moderate distress” category; Zt falling between +/− 0.75 standard deviation of the mean to the “normal” category; Zt between –0.75 and –1.5 standard deviations of the mean to the “moderate expansion” category; and Zt below –1.5 standard deviations of the mean to the “euphoria” category.

$$FSI_t = \sum_j \left[w_{jt} \times \int_{-\infty}^{x_j} f(x_{jt})\, dx_{jt} \right] \times 100 \tag{1}$$

where the x_{jt} term is the value of indicator j at time t, the integration term is the CDF of indicator j, and the w_{jt} term is the weight given to indicator j in the FSI at time t. Inspired by the continuous financial-stress- index methodology of Illing and Liu (2003, 2006), our FSI is currently constructed using daily data from 11 components that reflect four distinct financial markets: the interbank, foreign exchange, credit, and equity markets. One of the key technical challenges to be overcome by choosing the appropriate weighting methodology is the potential for false alarms; this potential should be balanced against the possibility of missing important events by setting warning standards too high.[42]

3.1. Description of the data

Our FSI uses daily observable financial-market data to capture the continuity of stress in financial markets. The data is of high quality; its sources are the Federal Reserve FRED database, Bloomberg, and the Bank of England (see Table 13 in Appendix A). Note that data beginning with the first quarter of 1985 was desired and collected in most cases; however, some series were unavailable before certain dates, and no obvious way to improve the data presented itself. The most severe constraint is Bloomberg's 10-year A-rated Bank Bond Index, which is not available before September 26, 1991. Two other binding constraints are the S&P 500 Financials Total Return Index, which is not available before September 13, 1990; and the S&P Financials Price Index, which is not available before September 11, 1989. In principle, it would be possible to reconstruct the missing Bloomberg historical data using alternative data sources and information on index composition and construction. However, at the time of the project's initial development, this was deemed a suboptimal choice. Consequently, the original development data set spans the fourth quarter of 1991 through the first quarter of 2009. Data quality considerations partly offset the depth limitation of the constrained choices described above. Most constituent time series contain some data gaps prior to 1987, but fewer gaps to 1994. Thus, starting the dataset in 1991 avoided the excessive costs of reconstructing time series and minimized data-gap problems.

[42] Gramlich, Miller, Oet, and Ong (2009), p. 17.

3.2. Variable selection – x_{jt}

To be certain that a versatile index of stress has been identified, the researcher aims to represent a spectrum of markets in which stress may originate. As previous research in this field attests, representatives of conditions in credit, foreign exchange, equity, and interbank markets provide substantial coverage of potential stress origination. The indicators for—and construction of—each of the four sectors discussed earlier are outlined below.

I—Interbank markets

(1) Financial beta – This indicator examines the volatility of share prices in the banking sector relative to overall stock market volatility. It is appealing to include banking-sector beta because it describes the strain on bank profitability, and potentially solvency, in light of changes in the profitability of publicly traded companies throughout the economy. The calculation of financial beta is

$$Financial\ Beta_t = {cov(r_t|_{t-1}^{t}, m_t|_{t-1}^{t})}\Big/{var(m_t|_{t-1}^{t})} \tag{2}$$

where r represents banking sector share prices (S&P 500 Financials Total Return Index); m is overall stock market share prices (S&P 500 Total Return Index); and $r_t|_{t-1}^{t}$ and $m_t|_{t-1}^{t}$ are each a set of observations from time t to one year earlier.

(2) Bank bond spread – This spread measures the perceptions of medium- to long-term risk in banks issuing A-rated bonds.[43] It is logical to include this measurement because it indicates the medium- to long-range risk to high-quality bank profits during periods of rational expectations. The calculation is

$$Bank\ Bond\ Spread_t = 10A_t - 10TB_t \tag{3}$$

where $10A$ refers to 10-year A-rated bank bond yields, and $10TB$ refers to 10-year Treasury yields.

(3) Interbank liquidity spread[44] – Examination of the TED spread provides some evidence regarding the perception of counterparty risk in interbank lending. The TED spread—the

43 A rating constitutes a composite computed by Bloomberg for its C07010Y Index (10-year A-rated Bank Bond Index) and comprising equally weighted S&P, Moody's, Fitch, and DBRS components.

44 The interbank liquidity spread as a component of the banking sector is provided by Illing and Liu in extension of the precedent FSI model.

difference between the three-month LIBOR rate and three-month Treasury yields—serves as a general indicator of overall liquidity risk because it reflects the risk premium associated with lending to commercial banks. When market liquidity is scarce, Treasuries trade at a premium and the spread between the two rates increases. The spread also increases when the risk of default on interbank loans rises. As an indicator of reduction in market liquidity, a rising TED spread can also serve as an indicator of downturns in equity markets.

By comparing LIBOR rates with three-month Treasury yields, one can observe periods of rising expectations of default risk on borrowed funds as the TED spread grows, worsening financial stress. The calculation is

$$Interbank\ Liquidity\ Spread_t = 3mo\ L_t - 3mo\ TB_t \tag{4}$$

where $3mo\ L$ represents three-month LIBOR rates and $3mo\ TB$ represents 90-day Treasury bill secondary-market rates.

(4) Interbank cost of borrowing[45] – This indicator measures the degree of apprehension with which banks loan to one another. By taking the 90-day LIBOR–Fed funds spread as a metric, the researcher identifies the risk premium that banks pay to borrow short-term funds from one another. The greater the spread, the more stressful are the conditions in interbank lending markets. The LIBOR–Fed funds spread is conceptually similar to the LIBOR–OIS spread—the latter being the difference between the London Interbank Offered Rate and the Overnight Swap Index rate—and can also be used to measure counterparty risk. In an efficiently functioning market, the difference between the unsecured LIBOR rate and the Fed funds target rate would approach zero through arbitrage. The spread between the two rates indicates either increases in default risk on the part of counterparties in a transaction or a lack of capital requirements for banks to engage in arbitrage. The indicator is calculated as follows:

$$Interbank\ Cost\ of\ Borrowing_t = 3mo\ L_t - FFR_t \tag{5}$$

where $3mo\ L$ is the three-month LIBOR rate, and FFR is the Federal funds target rate.

II—Foreign exchange markets

(5) Weighted dollar crashes – This indicator quantifies flight from the U.S. dollar toward a broad set of foreign currencies. With a floating exchange rate, depreciation of domestic currency

[45] The interbank cost of borrowing as a component of the banking sector is provided in extension of the precedent FSI model by Illing and Liu.

represents a loss to its holders, while unexpected volatility heightens uncertainty in foreign exchange markets. The researcher finds this a valuable signal of sharply decreasing foreign exchange transactions, which creates a sense of uncertainty as to the profitability of institutions system-wide. A reasonable implication of such uncertainty is increased demand for liquidity from the domestic financial system, requiring unanticipated, and potentially inefficient, lending. The calculation of this variable uses the formula

$$Weighted\ Dollar\ Crash_t = \frac{x_t}{max[x \in (x_{t-j} | j = 0,1,\dots,365]} \tag{6}$$

where x is the Trade-Weighted \$U.S. Exchange Index.

III—Credit markets

(6) Covered interest spread – This spread provides insight into uncertainty regarding government bond markets. Using U.K. government bonds as the counterpart to U.S. bills, arbitrage opportunities in efficiently functioning government-debt markets should drive the covered interest spread to zero. If the spread remains persistently non-zero, then the data signals arbitrageurs' unwillingness to hold a particular government's debt. This can imply a nation's difficulty in acquiring liquidity for governments, signaling the onset of stress. The calculation is

$$Covered\ Interest\ Spread_t = (1 + r_t^*) - \left(\frac{F_t}{S_t^*}\right)(1 + r_t) \tag{7}$$

where r* is the 90-day U.K. Treasury bill rate as of noon on day t; F is the 90-day forward rate for the U.K.–U.S. exchange rate as of noon on day t; S* is the spot U.K.–U.S. exchange rate as of noon on day t; and r is the 90-day U.S. Treasury bill rate as of noon on day t.

(7) Corporate bond spread – This, like the bank bond spread, measures medium- to long-term risk; it includes impressions of risk from corporations in all sectors. The researcher finds that the corporate bond spread is a useful stress indicator because, when the probability of losses increases, firms have trouble financing debt and may be less able to obtain liquidity, which implies still greater stress. The calculation is

$$Corporate\ Bond\ Spread_t = 10CB_t - 10TB_t \tag{8}$$

where 10CB is the 10-year Moody's Aaa-rated corporate bond yield; and 10TB is the 10-year Treasury yield.

(8) Liquidity spread – This spread measures changes in the short-term trend of differences in bid prices (BP) and ask prices (AP) on three-month Treasury bills. The bid–ask spread is a

component[46] of security transaction cost for concurrent sale (AP) and purchase (BP). As the spread (transaction cost) decreases, market security becomes more liquid. As instrument liquidity shrinks, the difference between its concurrent purchase price and sale price (the bid–ask spread) increases. Thus, market security's bid–ask spread measures its liquidity. The greater the spread, the more illiquid the market security and the greater the stress. The calculation is

$$Bid\ Ask\ Spread_t = \left(\frac{1}{30}\right) \sum_{i=0}^{29} \left[\frac{AP_{t-i} - BP_{t-i}}{\left(\frac{AP_{t-i} + BP_{t-i}}{2}\right)} \right] \tag{9}$$

where the moving average is calculated over the previous 30 trading days.

(9) 90-Day commercial paper–Treasury bill spread – This spread measures the short-term risk premium on financial companies' debt. During periods of rational expectations, this differential informs the researcher of changes in the underlying risk of financial companies' operations. The indicator is found by taking the difference between 90-day financial commercial paper rates ($90day\ CP$) and 90-day Treasury bill secondary market rates ($3mo\ TB$). The calculation is

$$90day\ Comm.Paper\ Treas.Spread_t = (90day\ CP_t) - (3mo\ TB_t) \tag{10}$$

Because data on 90-day financial and non-financial commercial paper rates only go back to 1Q: 1997, weekly 90-day commercial paper rates were used as a proxy for financial paper from 4Q: 1996 back to 1Q: 1985.

(10) Treasury–yield curve spread – The literature on the slope of the yield curve shows that the variable is a useful predictor of recessions.[47] Because of the combination of long-term uncertainty and short-term liquidity needs at the outset of—and during—recessionary times, including this variable as an indicator of financial stress is intuitive. This variable measures the 30-day moving average of the difference between three-month Treasury bill yields ($3mo$) on a bond-equivalent basis with 10-year constant maturity yields ($10yr$). The calculation is

$$Treasury\ Yield\ Curve_t = \left(\frac{1}{30}\right) \sum_{i=0}^{29} (10yr_{t-i} - 3mo_{t-i}) \tag{11}$$

---

46 In addition to dealer fees.

47 See Estrella and Hardouvelis (1991), Estrella and Mishkin (1996), Estrella and Trubin (2006), Haubrich (2006), and Haubrich and Bianco (2011).

IV—Equity markets

(11) Stock market crashes – This indicator captures the extent to which equity values in the S&P 500 have dropped since a year earlier. Moreover, by examining movements in the S&P 500 Financial Index, the researcher gains insight into expectations about the future of the banking industry. We calculate the variable in each time period (t) as

$$Stock\ Market\ Crash_t = \frac{x_t}{max[x \in (x_{t-j} | j = 0,1,\ldots,364]} \qquad (12)$$

where x refers to the S&P 500 Financials Index. To determine the demand for loanable funds, an examination of stock market crashes using the S&P 500 index is pursued, but the results are qualitatively similar.

3.3. Variable transformation – $\int_{-\infty}^{x_j} f(x_{jt})dx_{jt}$

Once the 11 indicators above are computed, the individual time series must be transformed to prepare for aggregation into the Financial Stress Index. The process involves generating a cumulative density function (CDF) for each indicator. This can be problematic insofar as each series spans different dates. To ensure commensurate percentiles, CDFs are generated using a common set of dates where data is fully populated for all indicators included in computing the FSI.[48]

The process of converting a given indicator into its CDF requires an intermediate step of computing a rank ordering of the data in the series. Once the corresponding rank series is generated for each indicator, the CDFs are computed as

$$CDF(x_{jt}) = \frac{Rank\ (x_{jt})}{number\ of\ daily\ observations} \qquad (13)$$

In most cases, the higher an indicator's computed value, the higher the rank associated with the value. For instance, a rank of 4,237 would be associated with the largest daily observation in the indicator's time series, while a rank of 1 would belong to the smallest daily observation. However, there are several series in which this convention is reversed: weighted dollar crashes,

[48] For example, in the initial construction, common dates spanned the period between September 26, 1991 and March 31, 2009, resulting in 4,237 daily observations.

stock market crashes, and the Treasury–yield curve spread. For these series, the largest daily observation of the indicator is assigned a rank of 1, while the smallest observation is ranked 4,237.

The reason for the reverse rank ordering of weighted dollar crashes and stock market crashes comes from their computation as outlined in section 3.2. Since the observed values of both variables are computed as the current value over the past year's high value, a larger output implies a smaller deviation of current value from recent past data and, therefore, a lower precedent. The reverse rank ordering of the Treasury–yield curve spread comes from literature that demonstrates the relationship between macroeconomic outcomes and the slope of the yield curve. Flat or inverted yield curves signal slow growth. Thus, the smaller the calculated yield curve spread, the more difficult financing becomes, which indicates a higher precedent. A summary of the rank assignments and variable names is provided in Table 14 in Appendix A.

3.4. Variable weighting – w_{jt}

Having appropriately computed the indicators' CDFs, all that remains is aggregation into the FSI. For weights, data from the Federal Reserve Board's Flow of Funds statistical release (Z.1) are used.[49] This data is separated into four sectors: bank loans (BNK), foreign exchange credit (FX), equity (EQ), and debt (DT). For any given quarter, total dollar flows through each sector are converted into a proportion of total dollar flows through all sectors, using the following equation:[50]

$$Z\ Proportion_t = \frac{Z_t}{BK_t + FX_t + EQ_t + DT_t}\ , Z \in \{BK, FX, EQ, DT\} \tag{14}$$

The Flow of Funds proportions are then used as weights for aggregating the 11 CDF functions generated above. Each indicator is identified as belonging to one of the four credit sectors and is weighted appropriately. When multiple indicators belong to a single credit sector, the sector

49 The FSI construction involved four competing aggregation methods: equal weights; equal variance weights; credit weights; and principal component weights.

50 In the data gathered, FX flows were missing until 3Q: 1997. To impute the data, the relative proportion of FX dollar flow to total flow of funds in each quarter prior to 3Q: 1997 was assumed to be constant, held to its known weight in 3Q: 1997 flow of funds data.

proportion is divided by the number of indicators in the sector, producing fractional weights. An equation representing the procedure follows:

$$Fractional\ Z\ Proportion_t = \frac{Z\ Proportion_t}{number\ of\ indicators\ in\ Z} \tag{15}$$

With the fractional proportions calculated, a daily FSI time series is computed by the equation

$$\begin{aligned} FSI_t = {} & (CDFSI1A_t + CDFSI2E_t + CDFSI10A_t + CDFSI11A_t) \times \frac{BK\ Proportion_t}{4} \\ & +(CDFSI3A_t) \times FX\ Proportion_t \\ & +(CDFSI4A_t + CDFSI5A_t + CDFSI6A_t + CDFSI7A_t + CDFSI8A_t) \times \frac{DT\ Proportion_t}{5} \\ & + (CDFSI9A_t) \times EQ\ Proportion_t \end{aligned} \tag{16}$$

Once the daily FSI series is generated, a quarterly FSI series is computed by taking a simple average of all daily FSI observations in the quarter.

3.4.1. Alternative variable weighting schemes

Having deduced the value of an indicator's precedent, the researcher must determine its contribution to overall financial stress during a given period of time. There is a rich literature on aggregation weights; this study has compared four alternative weighting schemes:

(i) Equal weights – In the absence of a priori knowledge of each component's importance in the aggregate index, a common weighting scheme gives all indicators equal importance. Two problems are apparent immediately. First, the equal-weighting scheme has no economic significance. It lacks intuition and is ultimately arbitrary. Second, the researcher implicitly assigns more importance to the economic sector with the highest number of observed variables.

(ii) Equal variance weights – Another method in this preliminary analysis fixes the variance of the weights over time by subtracting the indicator's sample mean from each observation and dividing by the indicator's sample standard deviation. As above the economic intuition of such a weighting technique is not readily clear.

(iii) Credit weights – To provide economic significance for the aggregation of FSI components, a credit weighting scheme is implemented. The four economic sectors under scrutiny in this project are weighted according to quarterly data on their share of total credit, including debt, equity, foreign exchange, and banking markets from the Federal Reserve Board's

statistical release Z.1 (Flow of Funds).[51] This lends some economic significance to the question of how much an indicator's precedent matters. However, the problem with this method is that where there are multiple indicators for a single sector, the sector's weight is divided evenly among them. This method seems to suffer from the same arbitrariness as that which plagues the equal weights approach, though less severely, given the weights' sensitivity to change in total credit composition.

(iv) Principal component – A useful approach for uncovering structural relationships between numerous time series is to identify orthogonal eigenvectors of the variance–covariance matrix of one's data. Each of these eigenvectors represents a linear combination of the data series employed by the researchers (called a factor) and is capable of tracking a certain percentage of the overall variability in the original data. In the forecast literature, it is not uncommon for a single factor to be responsible for the great majority of overall variability. Consequently, this eigenvector is selected as the appropriate weighting scheme. However, an examination of the variance–covariance matrix of the 11 series comprising the FSI in this study reveals that there is more than a single eigenvector. Instead, there are four eigenvectors, which together count for about 75–80 percent of variability in the data. Knowing this, the researchers create a weighting vector by taking the weighted sum of these four eigenvectors (using the proportion of variance explained earlier).

Two drawbacks to the principal-component method are readily apparent. First, many economists believe that the justification of these weights is not guided by a priori reasoning; therefore, the quality of the findings is increasingly subject to the peculiarities of the data. Further, without an a priori framework from which to consider the results, findings may simply be accepted as truth rather than challenged on logical grounds. Second, weighting based on a single component creates a fixed set of weights for all dates in the analysis, forcing market relationships to hold in the data when reality shows they may not.

[51] Federal Reserve Board statistical release E.16 (Country Exposure Lending Survey) is used to supplement the foreign exchange information in the Flow of Funds statistical releases Z.1 and Z.7.

4. Results and selection of CFSI weighting method

4.1. Comparison of candidate weighting methods

An agnostic selection process begins with the comparison of candidate stress series with documented stress episodes.

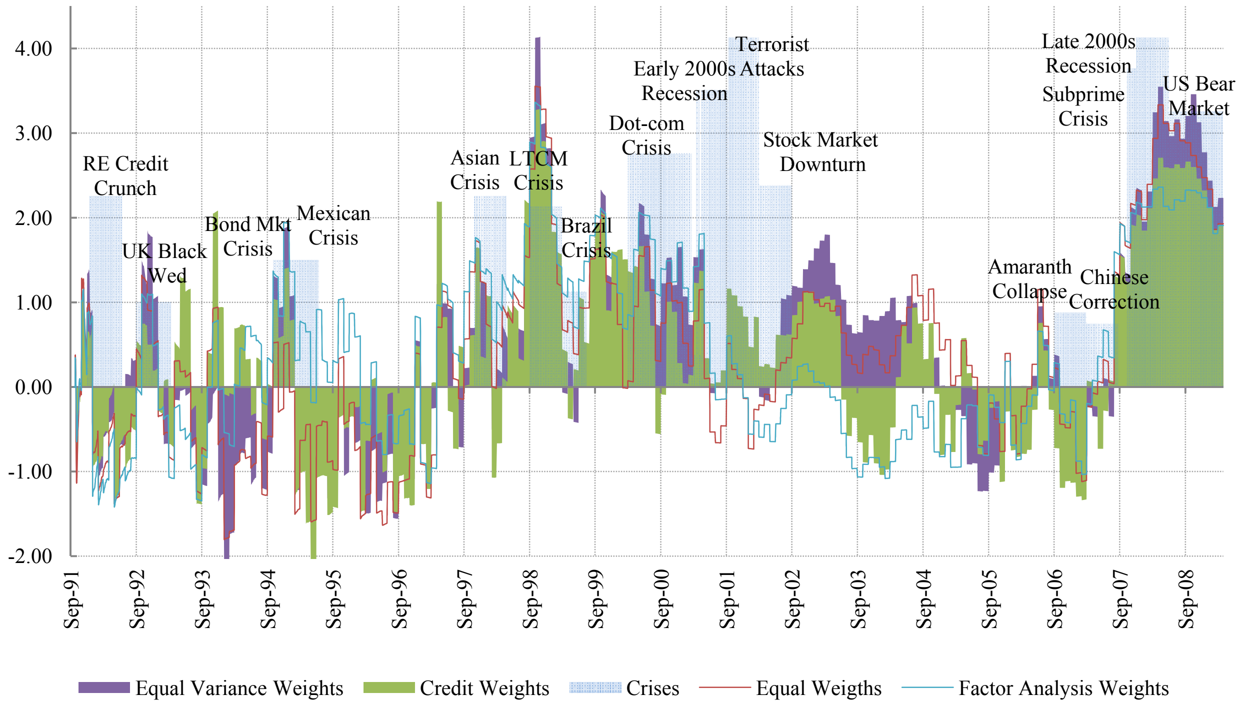


Fig. 1. Financial Stress Index: standardized monthly candidate series.

Note: Crisis bars are based on the Federal Reserve Bank of Cleveland's expert survey; larger bars reflect scaled judgment of greater significance.

Fig. 1 compares candidate standardized monthly financial stress series from 3Q: 1991 to 1Q: 2009. For reference, expert ranking assessment of literature-based series of financial crises during the same time frame is also shown. The figure suggests a number of comparative observations: All weighting methods show episodes of stress that are not documented in the literature. The concern here is that this particular set of stress measures is wrongly identifying periods of low stress as high stress, an error one may call a false positive. It is also possible, of course, that these same episodes simply represent periods in which stress was not large or severe enough to attract the attention of researchers or policymakers. The ambiguity is unsettling for identification of a specific episode: Is it a false positive or a period of stress unrecognized by the monitors? The ambiguity of the expert-based ranking of financial crises does not seem very helpful in distinguishing the relative accuracy of the variously weighted stress series. Furthermore, different weighting methods produce different interpretations of the same time periods. For example, the principal component weights and the credit weight series both indicate stress in 2Q–3Q: 1993, whereas both the equal weights and the equal variance weights series indicate that this period was expansionary (negative stress). The reverse is true for 1Q: 2003–4Q: 2004 , which is identified as a period of stress by both the equal weights and the equal variance series, while the principal component weights series indicates that this period is in expansion (negative stress). Interestingly, the credit weights series provides a more nuanced reading of this period. It identifies 1Q: 2003 as stressful; 2Q: 2003–1Q: 2004 as expansionary; and 2Q–4Q: 2004 as stressful.

To a policymaker concerned with ex ante action for systemic risk, it is critical to identify stressful episodes correctly. The prospect of conflicting identifications of stressful periods is particularly alarming. It raises the key question: What would be a sound basis for comparing and selecting the optimal financial stress series? In other words, how can the alternative series be reliably evaluated for accuracy?

4.2. Comparison of candidate weighting methods with expert survey

The literature on continuous index measures, including most recent studies, largely ignores the formidable problem of benchmarking. One recent exception is Hatzius et al. (2010), who select from the candidate indexes the one that performs best in forecasting economic growth through GDP. This approach is not applicable to the present study, whose objective is to forecast systemic banking risk rather than economic activity. Illing and Liu (2003) use an expert survey to select some alternative weighting schemes, based on a series' ability to capture the expert rank ordering of crisis episodes. For parity, we consider a similar comparison, but control the results for demonstrated expertise, tested separately. The light-blue background bars in Fig. 1 represent expert-ranked, literature-based financial crises. Comparison of different weighting methods reveals that each ranks high-stress episodes differently (see Table 1). This is not unexpected, given that different weighting methods vary in their emphasis on different financial markets and conflict in their prioritization of the aggregate stress indexes.

Applying a similar method for Canada, Illing and Liu (2003) showed that the credit weights series is optimal in minimizing type I/type II errors in rank ordering. By contrast, we consider alternative quantitative measures of both rank ordering and timing precedence, seeking the series that correlates most closely with the expert rank order, minimizes root mean square error, and usefully precedes the onset of financial crises (as shown by one-way Granger precedence). Consistent with the findings of Illing and Liu, Table 2 shows that the credit weights series is optimal when all three of these considerations are taken into account.

Table 1

Chronological rank ordering of survey stress episodes.

DATE	EVENT	EXPERT RANK	CREDIT WEIGHTS	PRINCIPAL COMPONENT	EQUAL VARIANCE	EQUAL WEIGHTS
12/2/1991	Scandinavian Crisis	18	16	13	10	10
12/16/1991	RE Credit Crunch	8	15	15	12	12
9/1/1992	ERM Crisis	14	18	18	16	16
9/16/1992	U.K. Black Wed	14	14	12	8	8
10/3/1994	Bond Market Crisis	11	8	6	7	13
12/15/1994	Mexican Crisis	11	7	7	6	14
10/27/1997	Asian Crisis	8	6	9	11	5
9/1/1998	LTCM Crisis	10	1	1	1	1
3/1/1999	Brazil Crisis	13	13	11	18	11
3/1/2000	Dot-com Crisis	6	10	14	17	19
2/20/2001	Turkish Crisis	18	9	8	9	9
3/15/2001	Early 2000s Recession	4	5	4	4	6
9/17/2001	Terrorist Attacks	1	12	16	13	15
3/1/2002	Stock Market Downturn	7	17	19	14	17
9/1/2006	Amaranth Collapse	16	19	17	19	18
2/27/2007	Chinese Correction	17	11	10	15	7
10/11/2007	Subprime Crisis	3	4	5	5	4
12/3/2007	Late 2000s Recession	1	2	2	2	2
6/2/2008	U.S. Bear Market	5	3	3	3	3

Table 2

Survey rank ordering and timing precedence results.

		EXPERT RANK	CREDIT WEIGHTS RANK	PRINCIPAL COMPONENT RANK	EQUAL VARIANCE RANK	EQUAL WEIGHTS RANK
Rank Order	Correlation	1.00	0.52	0.28	0.39	0.22
	RMSE	-	5.38	6.54	6.04	6.83
Timing Precedence	Granger		0.12 / 0.44	0.26/0.72	0.81/0.52	0.35/0.68

4.2.1. Critique of the approach

Three observations can be made immediately:

1. The expert-based rank ordering of the episodes, subjective and qualitative in nature, differs substantially from all of the other candidate stress series in ranking several crisis episodes, such as “terrorist attacks,” “stock market downturn,” and “early 2000s recession,” among others.

2. The candidate stress series contain a number of episodes that are not identified in the literature as crisis episodes, for example, stress episodes from the 1993–94 period and 1996, among others.
3. The timing of stress episodes tends to lead the timing of financial crises. This relationship makes economic sense, insofar as financial series foretell the advent of financial conditions, whereas financial crisis simply registers conditions' eventual deterioration to a recognized critical state.

These observations raise serious conceptual concerns about the applicability of an expert-based ranking of crises in validating either the rank ordering or the timing identification of stress episodes. This conclusion, of course, is very consistent with our arguments in sections 2.2 and 2.3. These reasons are discussed below corresponding to the above observations. First, expert surveys are based on human emotions that trigger the adrenal "fight-or-flight" response to stress. Experts, like the rest of us, still experience greater stress in the events that are the most frightening, not necessarily those that are the most systemically stressful. This is apparent in the conflicting evaluations of the "terrorist attacks" stress episode by the objective, stress-based series and the subjective, expert-score-based Z series (see Fig. 2). Similarly, the systemic impact of an event that appears less threatening can be seriously underestimated by experts, as in the LTCM stress episode.

Moreover, the expert survey is subjective and inherently biased toward crisis, mixing shock and response. As Boyd, De Nicolò, and Loukoianova (2010) show, one corollary of this is that the crisis series and the shock series have different rank orders.[52] Lastly, as discussed earlier, the expert-based timing of stress episodes is inherently flawed; it is generally the product of professional consensus, with accidental, rather than predefined, timing identification. Thus, we find the survey-based selection method problematic and must find a more rigorous method.

[52] Boyd, De Nicolò, and Loukoianova (2010), p. 14.

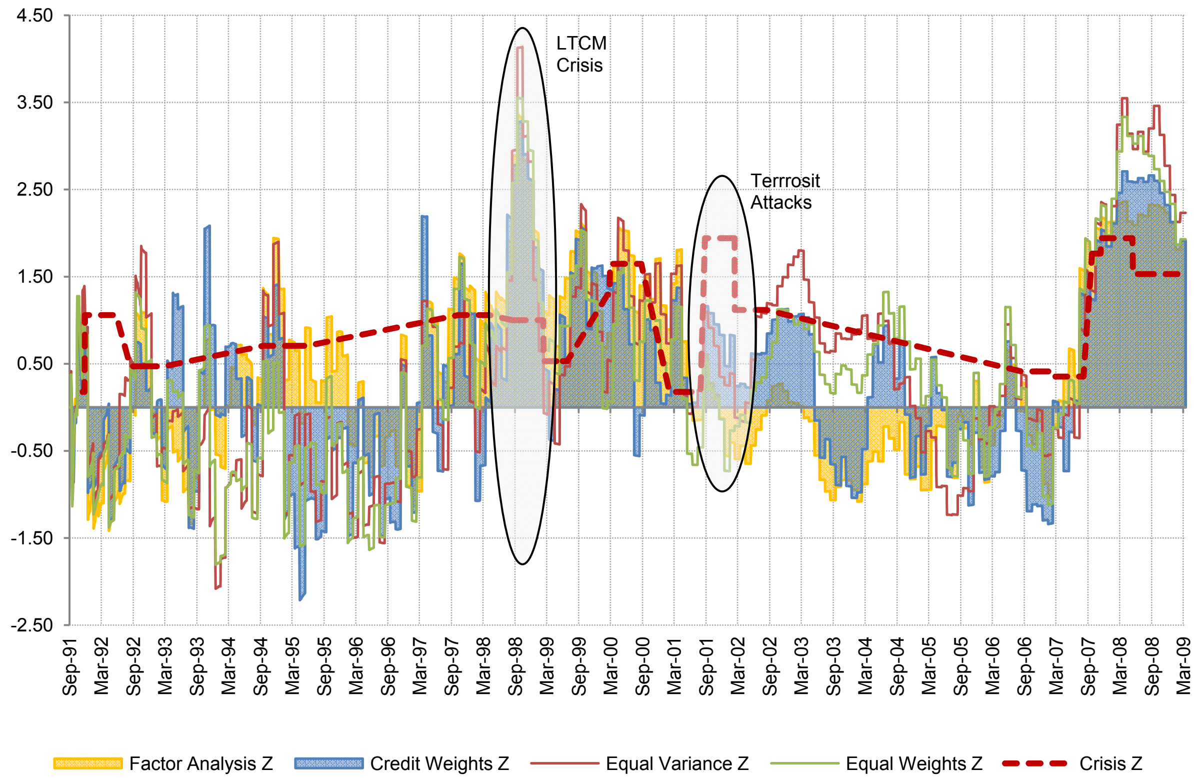


Fig. 2. Comparison of objective and subjective standardized monthly series.

4.3. Selection of CFSI from candidate weighting methods

Following a rigorous financial-stress-series selection process, we show that using the credit weights construction method for CFSI is optimal, both for use in a supervisory EWS and for monitoring systemic banking risk. This finding further confirms earlier survey-based results (section 4.2), as well as similar results by Illing and Liu (2003, 2006). Additional tests to optimize the monitoring frequency of the selected CFSI are described in section 5.5. The following selection method is used to determine which measure of financial stress best represents systemic stress.

1. Select volatility series to benchmark financial stress

The ability to create an early warning system (EWS) hinges on accurate identification of stressful and non-stressful periods. Stress is directly observable as a continuum difference between factors affecting entire market populations. Several types of stress are observable in each market, typically through measures that describe the relative risk/return differential of a product vis-à-vis a reference interest rate; product bid–ask spreads; and spreads that describe market liquidity. An alternative aggregate risk measure may be obtained using market volatility indexes, which estimate overall market risk statistically and are generally obtained directly from a sub-population of market participants.

An estimation error is inherent in volatility indexes, insofar as the market population differs from that entering the index. Given the supervisory objectives of ex ante action for systemic risk, the advantage of observing the factors underlying market distress is clear.[53] Nevertheless, volatility measures prove very useful in benchmarking candidate stress series. Comparing the candidate stress series to volatility benchmarks is valuable for determining which weighting scheme best tracks stress. To this end, we select benchmark market volatility series for each financial market.[54] Since the current version of CFSI addresses four distinct markets—interbank, FX, credit, and equity—we chose volatility indexes that describe these markets—MOVE, VDAX, LBOX,[55] and VIX, respectively.

2. Define systemic stress

Define systemic stress as two consecutive weeks of market volatility above the previous quarterly thresholds, or concurrent volatility in at least two distinct markets.

3. Extract a reference series of stress signals in each market

First, each benchmark is transformed into a weekly indicator of stress in the representative market (see Fig. 3). The calculation is as follows:

$$Z_{i,w,q} = \mathit{Index\ Weekly\ Stress\ Indicator}_{i,w,q} = \mu_{Index_{i,w,q}} - \left[\mu_{Index_{i,q-1}} + \sigma_{Index_{i,q-1}}\right] \quad (17)$$

---

53 We discussed the opacity of volatility indexes vis-à-vis spreads in section 2.4. It may be useful to elaborate the argument here as well. Market volatility indexes are constructed as weighted blends of prices of specific populations. From the standpoint of providing insight about why prices move, a volatility index is essentially a "black-box" series that allows only indirect insight into factors causing the movement of market prices. By contrast, a spread is a difference in prices between related securities and is observable directly. A change in spread communicates the relative activity of the two securities. In addition, the sources of a given change in spread are directly observable in the rise and fall of the two securities.

54 To avoid bias in the financial stress series, CFSI is careful to exclude these volatility series from the financial stress construction. In principle, it is also possible to construct a measure that directly utilizes markets' volatility benchmarks (e.g., use of VIX in the FRB St. Louis Financial Stress Index). CFSI avoids direct use of the volatility indexes in constructing financial stress for two reasons, one being precedent, and the other a preference for constructing the financial stress index from observable market elements (e.g., spreads) that are not indexes in themselves. This omission of the individual market volatility indexes makes them useful as a benchmark for independent monitoring and benchmarking of the CSFI.

55 LBOX (Lehman Swaptions Volatility Index) was used in the original benchmarking through September 4, 2009. LBOX was superseded by In July 2009 by BOX (Barclays Swaptions Volatility Index).

where, for each market i, w is the week number in the series and q is the quarter number in which the given week resides. The first term in the indicator equation is the mean of the benchmark index over the current week; the second term is the mean of the benchmark index over the previous quarter; and the third term is the standard deviation of the benchmark index over the previous quarter. If the weekly stress indicator is non-negative, the sector described by each index is considered in measuring the stress for the week. The logic here is that if present expectations are sufficiently greater than some medium-term trend in expectations, conditions in the represented market become stressful as the distribution of returns in the represented market becomes muddled to market participants.

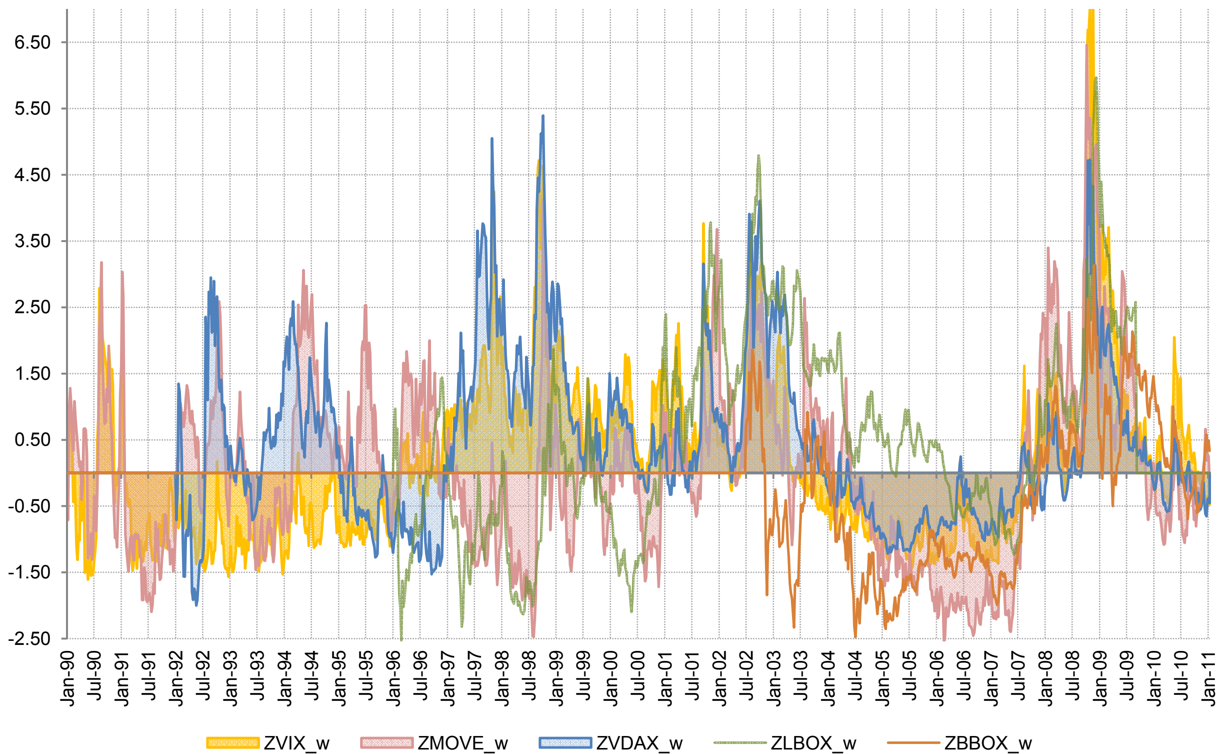


Fig. 3. Weekly reference market volatility series.

4. Establish a weekly series of systemic stress

As none of the four benchmark indexes represent a broad enough measure of stress to uniquely identify a period of systemic stress across financial markets, the indexes are combined into a single benchmark indicator. After the four benchmark weekly stress indicators are generated, a pattern for identifying a systemically stressful week is determined. The researchers test a trial operational definition that identifies a systemically stressful week as satisfying at least one of the following conditions:

- No less than two of the benchmark indicators signal stress for the same week.
- A single benchmark indicator signals stress for at least the second week in a row.

To this end, we establish a combined reference binary series of weekly systemic stress signals based on the definition of systemic stress adapted in (1), as follows:

$$S_{w,q} = \begin{cases} 1, [Z_{i,w,q} > \theta_i] \wedge \left([Z_{j,w,q} > \theta_j] \vee [Z_{k,w,q} > \theta_k] \vee [Z_{l,w,q} > \theta_l]\right) \\ 0, else \end{cases} \quad (18)$$

Here, θ from a set of $\{\theta_i, \theta_j, \theta_k, \theta_l\}$ is the signal threshold set to 1 std in the initial operational trial. Combined signals of systemic stress are found in 92 percent of the weeks from 3Q: 1991 to 1Q: 2009, as summarized in Table 3 (panel A), using this trial operational definition.

Table 3

Benchmarking summary of weekly signals of systemic stress.

PANEL A:	Benchmark	# Stress Weeks	% Stress Weeks	# Non-Stress Weeks	% Non-Stress Weeks	Max Z
Benchmark volatility series	VIX	615	67%	302	33%	7.27
	MOVE	538	59%	379	41%	6.46
	VDAX	544	61%	342	39%	5.39
	LBOX / BBOX	382	56%	295	44%	5.97 / 3.13
	Combined	848	92%	78	8%	7.27
PANEL B:	Benchmark	# Stress Weeks	% Stress Weeks	# Non-Stress Weeks	% Non-Stress Weeks	Max Z
Candidate FSI series	Credit weights	278	30%	639	70%	4.29
	Principal Component	320	35%	597	65%	4.55
	Equal variance	306	33%	611	67%	5.69
	Equal weights	265	29%	652	71%	4.98

5. Determine the timing and origins of systemic stress episodes

Clearly, the above trial of systemically stressful weeks is insufficient and flawed as an exclusive identification benchmark. To begin, there are simply too many benchmarked weekly volatility-based signals. This indicates that the trial operational definition above cannot sufficiently differentiate systemic conditions on a weekly basis. This is not altogether unexpected, given the intuition that weekly volatility indexes would be quite sensitive to daily volatility noise. One way to improve this identification is to raise the discrimination thresholds, for example, by increasing 1 std to 2 std; increasing the signal persistence level from two consecutive weeks to a greater number of weeks; or increasing the signal "resonance" from two concurrent markets to a greater number. Similarly, in a financial stress index series, a similar weekly signal series would have a different basis than that embedded in the volatility-based series. Since FSI is a combined CDF measure, there is no notion of magnitude when we compare a weekly mean to a previous quarter mean plus one standard deviation. At times, FSI movements will be more compact than the benchmark volatility series movements; elsewhere, FSI movements will be more pronounced. Evidence of this basis discrepancy can be found by comparing panels A and B in Table 3. A comparison of the overall duration of a stress episode, identified using either volatility- or stress-based methods, should be more meaningful.

Alternatively, the above operational results may still be used to identify systemic stress episodes, in the spirit of Hendricks, Kambhu, and Mosser (2007) and Kambhu, Weidman, and Krishnan (2007): from the time the movement in the stress/risk series crosses the positive signal threshold to the time this movement crosses the negative signal threshold. To proceed thus, we identify a systemic stress episode as a period of contiguous weeks for which each week in the period is deemed systemically stressful. Here, the data appears more forgiving because the number of episodes ranges from 27 in the credit markets (using the LBOX and BBOX volatility indexes) to 49 in the interbank markets (using the MOVE volatility index). Over 3Q: 1991 to 1Q: 2009 horizon, 27 such episodes are identified by the combined benchmark using 1 std discrimination threshold θ (see panel A in Table 4 and Table 18 in Appendix C). The basis difference between the volatility and stress series also disappears (see panel B in Table 4). The number of systemic stress episodes ranges from 28 for the equal weights series to 36 for the credit weights series.

With the contiguous systemically stressful weeks and the systemic stress episodes in hand, one can undertake a series of statistical procedures to determine which version of FSI best tracks our benchmark.

Table 4.

Benchmarking summary of systemic stress episodes.

Panel A:	Benchmark	# Stress Episodes
Benchmark volatility series	VIX	47
	MOVE	49
	VDAX	33
	LBOX/BBOX	27
	Combined	27
Panel B:	Benchmark	# Stress Episodes
Candidate FSI series	Credit weights	36
	Principal Component	27
	Equal variance	30
	Equal weights	28

6. Determine stress signals series from each candidate stress index

Transform each alternative financial stress series into a signal in the same manner as the volatility benchmarks in (3). Here, it is useful to repeat this transformation for each frequency of the financial stress time series that is intended to be tested for monitoring, as well as for the quarterly series used for the EWS.[56] Each intended frequency should result in a corresponding signal series. The formula below gives a transformation for the weekly signal:

$$FSI\ Weekly\ Stress\ Indicator_{t,w,q} = \mu_{FSI_{w,q}} - \left[\mu_{FSI_{q-1}} + \sigma_{FSI_{q-1}}\right] \qquad (19)$$

7. Conduct comparative error and noise/signal analysis

Conduct a type I/type II error analysis of alternative FSIs and compare their noise/signal ratios. The results of error and noise/signal analysis for weekly, FOMC, and quarterly frequencies are shown in Table 5 (panel A). The table shows classification results for two types of comparisons of the alternative weighting methods: error analysis and additional non-parametric tests (see step 8). Volatility series are used to create a reference time series of systemic stress signals (steps 1–5 above). Panel A compares an error analysis of the time series of systemic stress episodes produced by the alternative weighting methods (step 6 above) against the reference time series. In this table, type I errors indicate the “failure rate” to correctly identify the reference systemic stress episodes; type II errors indicate the “false positive rate” in incorrectly identifying a reference systemic stress episode. The noise/signal ratio is defined as the ratio of type II errors to one minus type I errors.

8. Conduct additional non-parametric comparative analysis

To enable further non-parametric comparative tests[57] of the competing alternative FSIs, transform each FSI series into signaling grades, in which an FSI signal is graded by its z-score

[56] SAFE EWS uses quarterly financial stress series. However, we anticipate that a higher-frequency version of a financial stress time series may also be useful for monitoring: weekly, bi-weekly, monthly, and between FOMC meetings (irregular intervals of approximately six weeks). It is optimal to choose a monitoring frequency that minimizes the noise/signal ratio while avoiding identification of idiosyncratic (that is, political and non-financial) stress episodes. See section 5.5.

[57] Since FSI is essentially a relative rank-order signaling measure, selecting and benchmarking the best FSI among competing alternatives can be improved by non-parametric statistical tests. As a signaling measure, FSI is also similar to a two-dimensional rating system for systemic stress, in which a signal of a certain “grade” conveys information not only about the probability of systemic stress, but also about its severity.

(number of standard deviations from the mean). Conduct additional non-parametric tests of the alternative stress series as grade-based rating systems of systemic stress. Our benchmarking process involved two additional non-parametric tests: a receiver operating characteristic (ROC) analysis and a Somer's D analysis, a measure of association describing the difference of the conditional probabilities. Classification results for non-parametric comparisons against the reference series for weekly, FOMC, and quarterly frequencies are shown in Table 5 (panel B). Non-parametric results also include testing the null hypothesis (Ho) that the rating system constructed by a candidate FSI series is random. As can be seen in panel B, credit weights is the only candidate that can consistently reject H_0 of randomness for all three frequencies while providing the desirable rating power (ROC and Somer's D) and error performance (type I, type II, and noise/signal).[58]

Table 5

Results of FSI Selection Procedure.

		CW	PC	EV	EW	**CW**	PC	EV	EW	**CW**	PC	**EV**	EW
		Weekly				FOMC				Quarterly			
Panel A - Error Analysis	Type I	28.5%	27.4%	28.6%	29.3%	20.7%	21.5%	22.2%	21.5%	**36.6%**	43.7%	42.3%	45.1%
	Type II	10.6%	12.7%	14.4%	15.7%	18.5%	19.3%	19.3%	18.5%	9.9%	7.0%	**5.6%**	5.6%
	Noise/Signal	**14.9%**	17.5%	20.2%	22.1%	**23.4%**	24.5%	24.8%	23.6%	15.6%	12.5%	**9.8%**	10.3%
Panel B - Rating Analysis	ROC	**56.7%**	54.9%	50.6%	49.3%	**56.7%**	56.1%	51.7%	48.4%	**58.5%**	56.2%	**60.0%**	53.2%
	SOMER's D	**13.3%**	9.7%	1.2%	-1.4%	**13.3%**	12.3%	3.4%	-3.2%	**17.1%**	12.5%	**20.0%**	6.4%
	H_0	**Yes**	Yes	No	No	No	No	No	No	**Yes**	No	**Yes**	No

Note: CW – Credit Weights, PC – Principal Component, EV – Equal Variance Weights, EW – Equal Weights.

5. CFSI results

To summarize, CFSI is constructed as a weighted, relative rank-order signaling measure of observable financial stress components. Selection of the optimal CFSI weighting method satisfies the dual supervisory objectives of early warning system use and monitoring; it is driven by non-parametric statistical testing. Three tests support use of the credit-aggregates weighting method: type I/type II error analysis; receiver operating characteristic (ROC); and Somer's D analysis. The last of these is a measure of association describing the difference in the conditional probabilities of observing a systemic stress episode, given groups of standardized FSI distances

58 Results for quarterly frequency show that equal-variance weights may be slightly superior to credit weights in this frequency designed for EWS use. However, the choice is unclear: both have about the same rating power. The error analysis for this frequency shows that equal-variance weights offer better noise/signal discrimination, whereas credit weights offer the least failure to identify systemic stress episodes. The choice of credit weights for all horizons is maintained by considerations of operational efficiency for both monitoring and EWS.

to mean. The testing method is shown to be optimal under flexible use frequencies for monitoring and EWS: weekly, FOMC-meeting frequency, and quarterly.

Fig. 4 shows the results of the credit weights financial stress index series selected as CFSI. Daily, weekly, monthly, and quarterly results are shown from September 1991 to March 2011. First, the CFSI series is reported in "levels" scaled from zero—the smallest possible stress to 100—the largest possible stress, by contrast with the reporting of stress in standard deviations in Fig. 1–Fig. 3. Second, the CFSI values are continually updated (at a daily frequency) as the observed financial stress components evolve. This means that new observations of stress would change the components' cumulative density function (CDF). Therefore, re-adjustments in the component CDF values would cause corresponding adjustments in the value of CFSI level at a given point in time. These re-adjustments are made daily and captured in adjusted CFSI levels that can be observed at a frequency suited to the particular CFSI use objectives. Third, Fig. 4 provides useful insights about CFSI's ability to differentiate idiosyncratic (that is, political and non-financial stress) in the markets. As the figure shows, daily CFSI captures high-frequency market stress, including idiosyncratic stress. Weekly CFSI is only marginally better in reducing the impact on CFSI of very-short-lived idiosyncratic events. Monthly and quarterly CFSI are significantly better at sifting out episodes of idiosyncratic events that make no lasting fundamental impact on market stress.

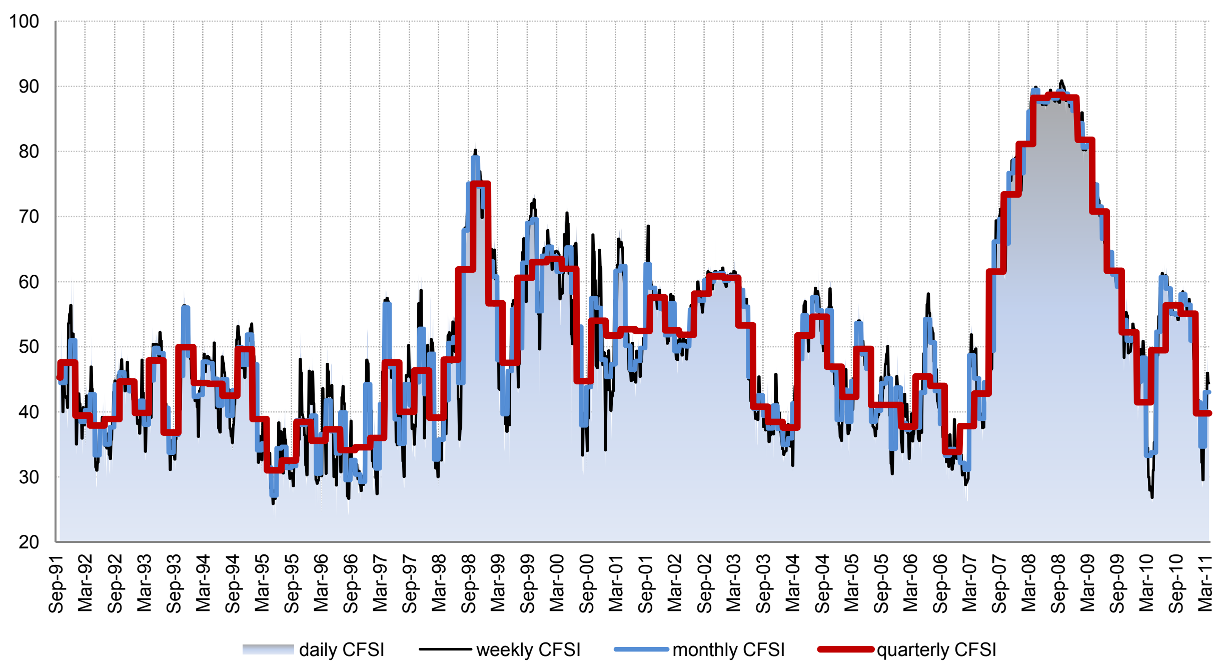


Fig 4. CFSI (September 1991–March 2011).

5.1. Dependability

Although a review of the literature shows that the use of a stress index as a dependent variable in EWSs is gaining wider acceptance,[59] a better case may be made to show that the dependent variable constructed in this manner is, indeed, "dependable" for both monitoring and EWS. The key challenge in this regard is to demonstrate that a particular set of FSI frequencies, constructed with the optimal weighting methodology used in this study, can efficiently filter out idiosyncratic noise events that affect the markets temporarily. Insofar as CFSI fails to do so, a reasonable objection may be made that an EWS model with CFSI as a dependent variable would aim to predict political or non-financial events, which is neither possible nor desirable from the viewpoint of estimating EWS regressors. This demonstration is straightforward and illustrates CFSI's ability to filter out idiosyncratic events that cause short-term market volatility (such as 9/11, Desert Storm, the Iraq War, etc.).

Fig. 18 and Fig. 19 in Appendix C show that the quarterly FSI index indeed possesses these desirable filtering characteristics, which are important for use in an EWS of systemic banking risk. The figures compare the filtering ability of daily, weekly, and quarterly CFSI. While higher-frequency financial stress indexes (daily and weekly) reflect concurrent volatility during systemic stress episodes, both financial and idiosyncratic, the quarterly CFSI eschews idiosyncratic volatility, reflecting a slower accumulation of financial imbalances.

5.2. CFSI robustness

The use of quarterly CFSI (CFSIq) as a dependent variable in an EWS is predicated on the confirmation of its econometric robustness and the analysis of its time-series properties. It is well known that regressions of one time series on another frequently lead to high goodness-of-fit, even if there is no meaningful relationship between variables; this problem is known as *spurious regression*. Insofar as the statistics of a stress series follow a random walk, explaining and forecasting stress statistics' future behavior becomes impossible. An EWS of systemic banking risk is essentially a forecasting model. Thus, the applicability of a financial stress series as a dependent variable in such an EWS depends strongly on verification of stationarity. If CFSIq is

[59] See section 2 and Gramlich, Miller, Oet, and Ong (2010).

found to be non-stationary, then the EWS researcher would need to verify cointegration before making further EWS model adjustments.

In generating an EWS model that predicts financial stress, a useful question to ponder is to what extent a CFSIq stress series self-predicts. The first step toward answering this question is to determine the underlying data generation process for stress. Fig. 17 – Correlogram of quarterly FSI (see Appendix A) provides a summary of CFSIq's autoregressive properties. As the CFSI correlogram, autocorrelation, and partial autocorrelation functions show, the effects of lagged levels of CFSIq tend to dissipate after six quarters. A fairly fast decline indicates that the time series' long-run development is not affected beyond a certain horizon. This aspect of the correlogram is more consistent with a deterministic time series that has a stationary AR(1) component than with a nonstationary series. Thus, the decay of the autocorrelation function suggests a significant autoregressive component to CFSIq in absence of a moving average component.

A possible reason for the autoregressive significance may be the nonstationarity of CFSIq time series. We conduct extensive graphical analysis—including the correlogram analysis and unit root tests of the quarterly CFSI series—and conclude that the quarterly series of the financial stress index can be considered stationary with a nonzero mean at a 5 percent critical level or weakly stationary around a deterministic trend at a 10 percent critical level. The results of CFSI stationarity testing are given Appendix B. This is a welcome finding because the process shows that CFSI can be used in level form as a dependent variable in a forecasting EWS.

5.3. Decomposition of CFSI

The dynamic weighting method by which CFSI is constructed allows for intriguing interpretations of economic conditions. The weights of CFSI's four market components fluctuate as the structure of the financial system evolves. In turn, as these weights change (see Fig. 5), some market sectors become more or less pertinent relative to others. For example, the weight for the credit markets increased from 0.3 to 0.4 during the subprime crisis, and this sector played an increasing role in the change in CFSI over the crisis. Conversely, the weight for the equity markets decreased from nearly 0.5 in the late 1990s to roughly 0.3 in 2010. Clearly, the increase in equity markets played a significant role in the decrease in CFSI after the subprime crisis; however, this effect would have been larger if the weight had been as large as in the late 1990s.

Fig. 6 shows the movements of specific components within the monthly CFSI, providing insight into the amount of stress that each of the four distinct markets contributed to the overall stress series. As the figure shows, over time the component from the foreign exchange market contributed substantially less to overall financial stress than other markets. One can also observe that measures from the credit, interbank, and equity markets tend to contribute significantly to overall financial stress. Their contributions in periods of financial stress tend to rise and fall together, amplifying overall changes in financial stress. The correlated behavior of stress components of the equity, interbank, and credit markets does have some exceptions. Consider, for example, the evolution of the subprime crisis of 2007–10. There was an observed initial stress increase in all four markets composing CFSI. As the crisis progressed and the Federal Reserve took extraordinary steps to mitigate this stress, CFSI shows a decrease in overall stress starting in April 2008. The most marked drop-offs in stress were first apparent in the CFIS's equity-market stress component of CFSI, followed by stress declines in interbank and credit markets in 2009. A similar, but less dramatic, pattern can be observed in the latent phase of the LTCM crisis of 1998, as the Federal Reserve put stabilizing measures in place, first reducing stress in the equity markets, then relieving stress in the interbank and credit markets.

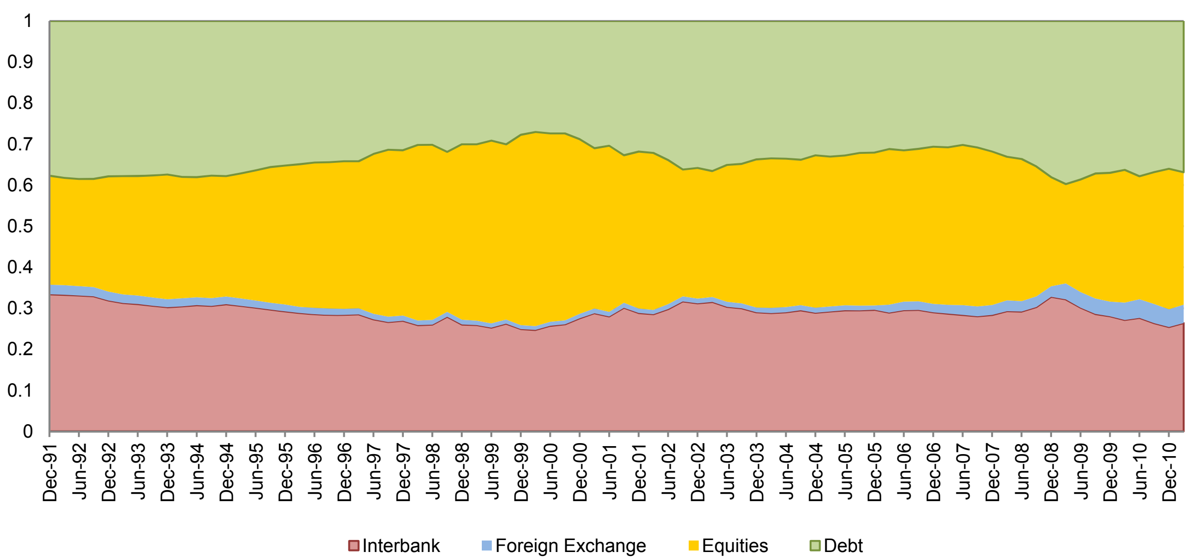


Fig. 5. Dynamic change in CFSI market component weights (4Q: 1991–1Q: 2011).

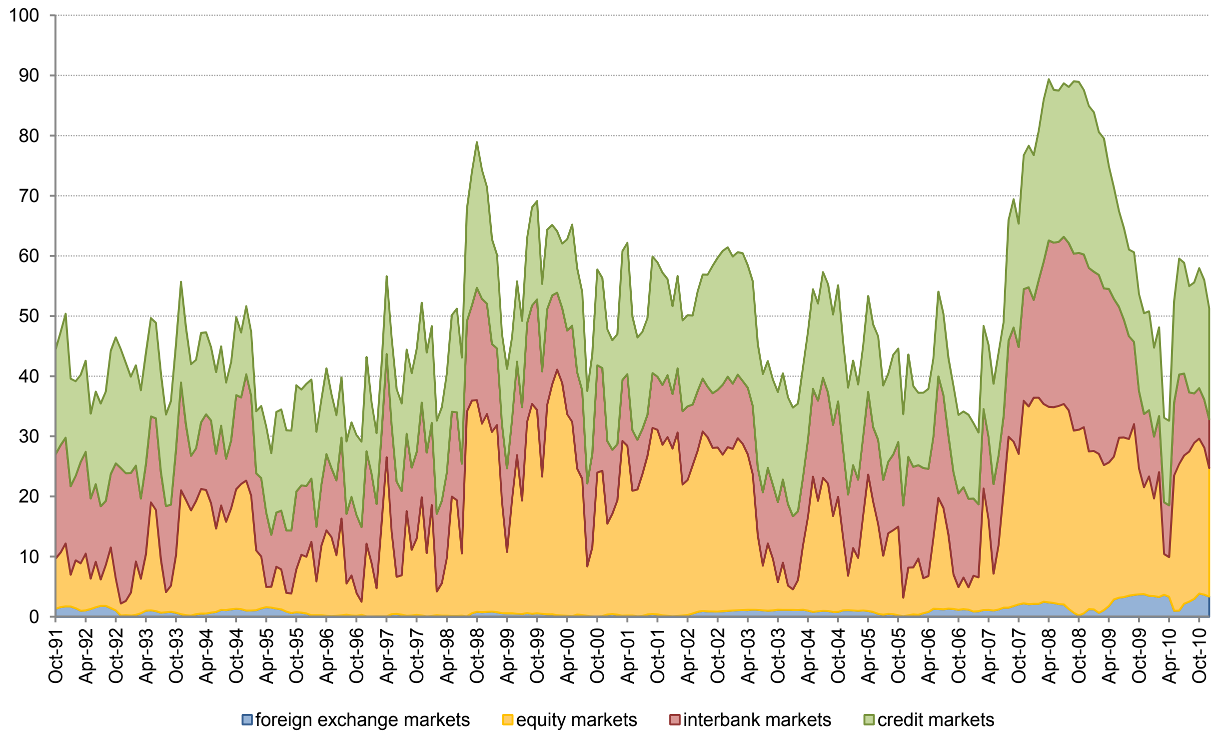


Fig. 6. CFSI components.

It is quite useful to consider the relative contributions of the four markets in tandem with their volatility benchmarks. Fig. 7 focuses on the four markets' volatility benchmarks, measured in terms of standard deviations from their means from 3Q: 2006 to 4Q: 2010. Three of the four[60] volatility measures are collectively positive from September 2007 to October 2009.[61] By contrast, the financial stress series identifies the subprime episode differently (see Fig. 2, Fig. 4, and Fig. 6): from March 2007 to December 2009. The episode is identified earlier, and some residual stress is shown beyond the episode's end implied by the volatility benchmarks. This pattern of identification is not unique to the subprime crisis; other episodes tend to be identified earlier and longer through stress measures rather than volatility.

Data from 2010 shows that equity markets experienced a period of increased volatility from January to September, whereas other markets' volatility diminished. In addition, credit-market volatility persisted from December 2009 to June 2010, while volatility in the other markets

[60] We use the LBOX volatility series as a benchmark for credit markets up to October 2009 and the BBOX volatility series thereafter.

[61] Note that the choice of distance threshold is critical for timing the stress episode. If we choose one standard deviation as the threshold, the stress episode is shortened to December 2007–September 2009.

subsided. These volatility observations are useful to a degree, but fail to provide insight as to the mechanism behind these volatility trends.

This is where observations from individual components of financial stress offer substantial benefit. Fig. 8 and Fig. 9 decompose stress in the interbank and credit markets, respectively, from 2Q: 2006 to 4Q: 2010. Fig. 8 shows that in the initial phase of the subprime crisis, from March to July 2007, interbank markets' stress was driven primarily by growth in the interbank liquidity spread and bank bond spread, later accentuated by the financial beta. The interbank cost of borrowing only became a factor at the height of the crisis, from March 2008 to May 2009, because interbank costs decreased as the Federal Reserve began decreasing the federal funds rate among other, less conventional tools. As the figure shows, stress in the interbank markets substantially subsided toward December 2009, with one exception: the interbank liquidity spread remained large.

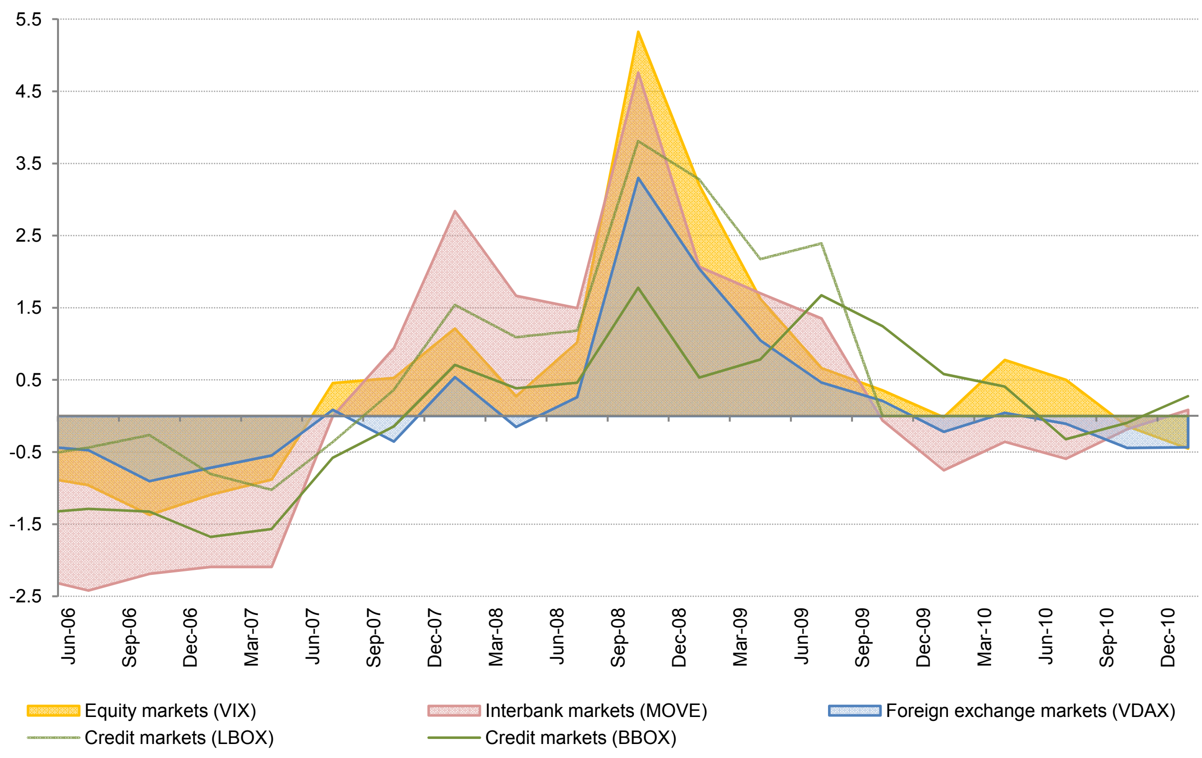


Fig. 7. Market volatility benchmarks (3Q: 2006–4Q: 2010).

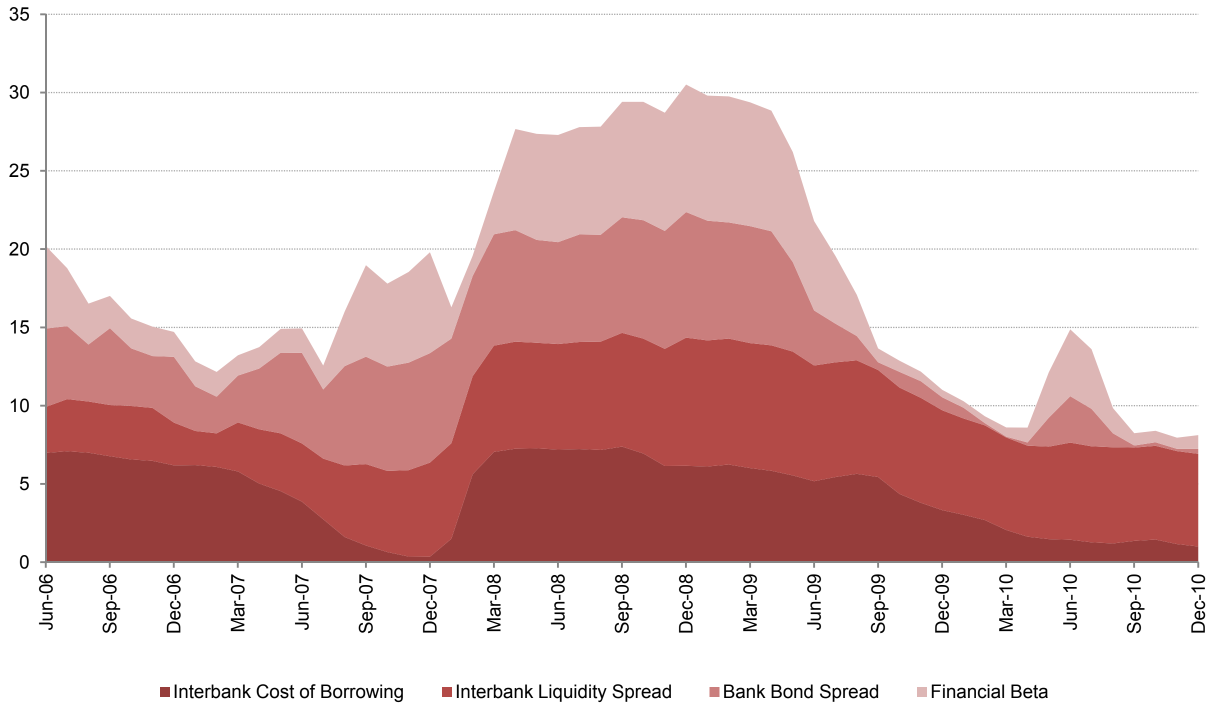


Fig. 8. Components of stress in Interbank markets (3Q: 2006–4Q: 2010).

Fig. 9 shows the components of stress in the credit markets. At the onset of the subprime crisis, from March to July 2007, credit markets' change in stress was mainly driven by increases in the covered interest spread and the commercial paper–T-bill spread, with other spreads remaining relatively steady. At the height of the crisis, from March 2008 to October 2009, increases in the covered interest spread, corporate bond spread, and commercial paper–T-bill spread were the most significant, accentuated by the liquidity spread. As stress in the credit markets subsided toward December 2009, only the corporate bond spread and the liquidity spread remained wide.

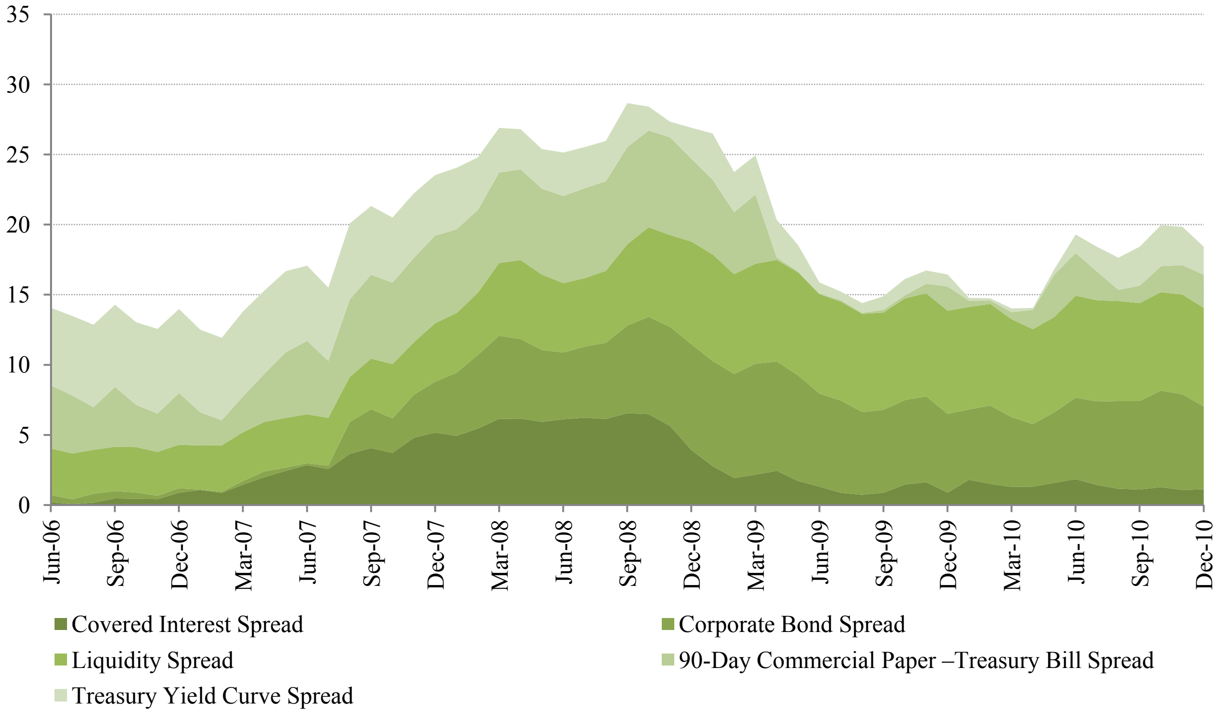


Fig. 9. Components of stress in credit markets (3Q: 2006–4Q: 2010).

6. Discussion of the results

6.1. Duration pattern of systemic stress episodes

In this section, we present evidence of a structural connection between financial deregulation and the pattern of systemic-stress episodes. Drawing on the literature and our data, we argue that there appears to be "no free lunch" in reducing risk. Greater growth and risk reduction for individual institutions in good times are followed by adverse systemic risk effects in bad times. Our evidence suggests that the frequency of systemic-stress episodes remains consistent before and after U.S. financial deregulation. We observe that after deregulation, the speed of systemic stress propagation slows (the benefit of risk diversification for individual institutions). However, the length of the recovery from systemic stress also increases substantially (the penalty of universal banking).

In comparing the candidate financial stress series of Fig. 1, the observer will note a pronounced difference in patterns of stress volatility before and after 1998. Although the frequency of stress episodes is generally similar among the different candidate series, the episodes' duration appears substantially different. Pre-1998 stress episodes tend to be short relative to post-1998 episodes, which tend to dissipate more slowly.

Additional insight into the apparent pattern in stress-episode duration can be obtained by considering a graph of the rate of change in various candidate financial stress series (delta Z_{FSI}) per unit of time (dZ_{FSI}/dt). The physical meaning of this is the velocity of financial stress (see Fig. 10). It may also be useful to consider this information as a view of volatility, in the sense that volatility describes the variation of price over time. Higher values of velocity (volatility) of stress at the episode's onset (respectively recovery) indicate faster evolution of critical states (respectively faster recovery). Lower values of velocity (volatility) of stress at onset (respectively recovery) indicate longer onset of stress (respectively slower recovery).

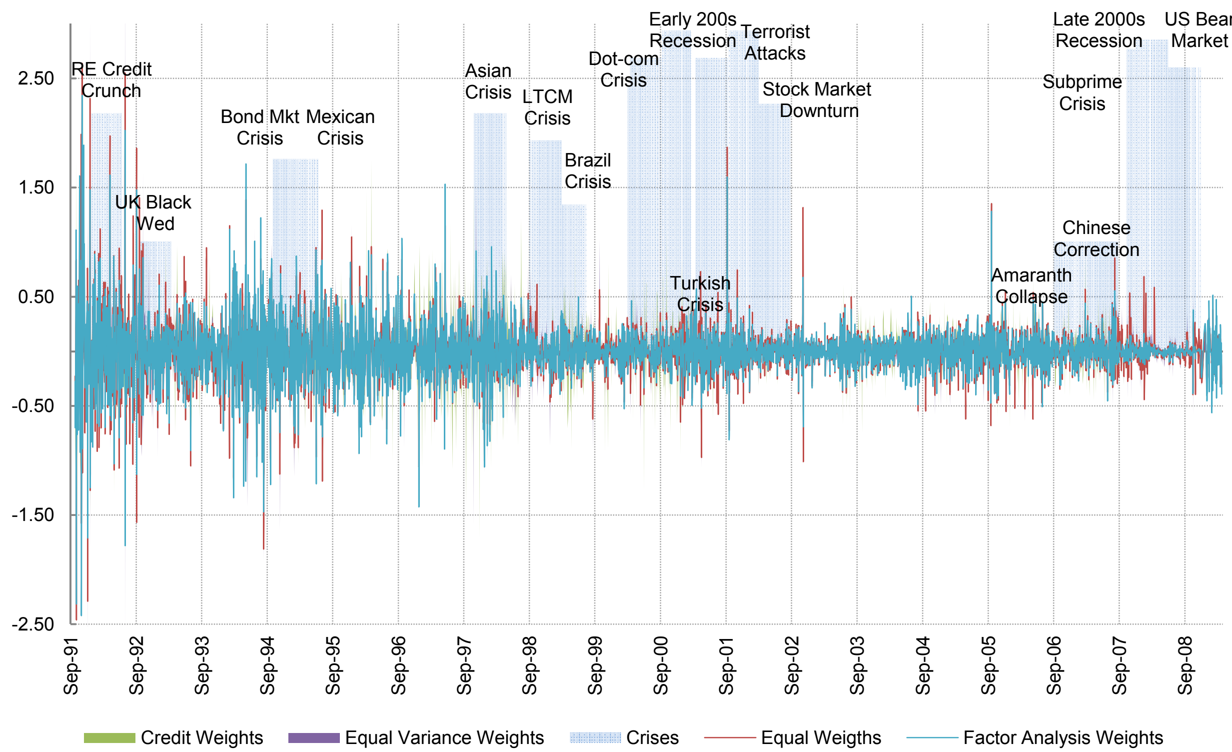


Fig. 10. Velocity of financial stress (4Q: 1991–4Q: 2008).
Note: Crisis bars are based on the Federal Reserve Bank of Cleveland's expert survey, in which larger bars reflect scaled judgment of greater significance.

Fig. 10 supports the observation from Fig. 1 that there may be a change in the stress pattern before- versus after 1998. The comparison of velocity of stress allows a view into crisis onset and post-crisis recovery. Pre-1998 stress velocity is characterized by sharp swings in Z_{CFSI} prior to crises episodes within a more volatile bandwidth: generally from -0.5 std to +0.5 std and 2 to 4 times higher at onset of crises. Post-1998 stress velocity bandwidth is roughly half: from -0.25 std to +0.25 std, while also amplifying something like two to four times at the onset of crises. Fig. 1 and Fig. 10 describe a slower evolution of crises (a welcome pattern) and slower recovery from crises (an unwelcome pattern) after 1998.

Further clarity can be obtained by directly considering distribution of the duration of stress episodes pre- and post-1998, shown in Fig. 11 and Fig. 12. Both CFSI and volatility benchmarks indicate a similar pattern of stress episodes' increased duration in the post-1998 period.

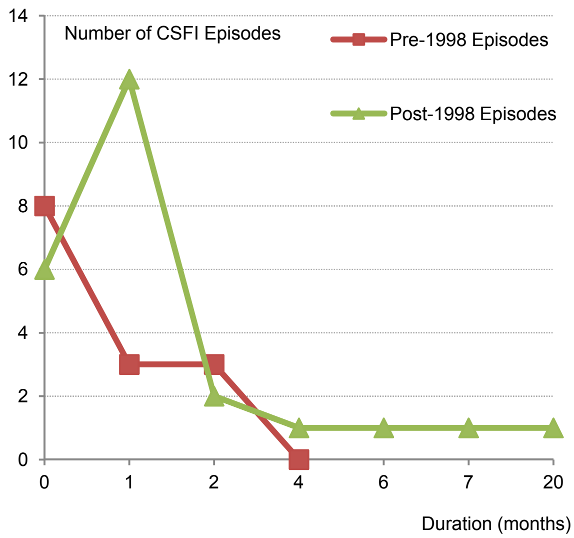


Fig. 11. Duration distribution of CFSI systemic-stress episodes (months).

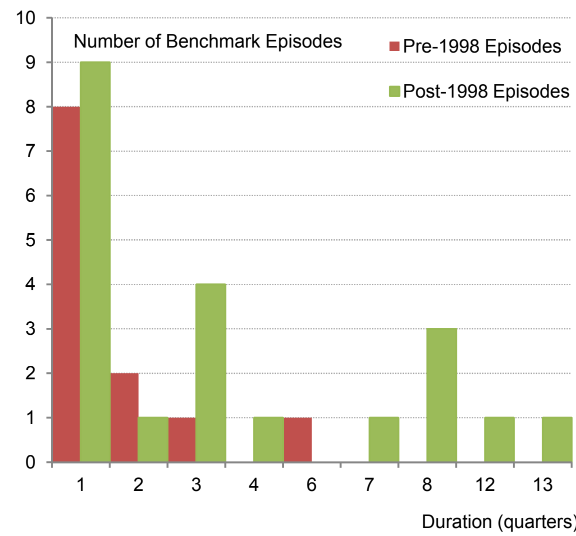


Fig. 12. Duration distribution of benchmark systemic-stress episodes (quarters).

Our evidence supports the idea that the change in pattern only affects the duration of stress episodes and not their frequency. Table 6 confirms the observation in Fig. 1 that the frequency of systemic-stress episodes remains generally consistent in pre- and post-1998 periods.

Table 6

Systemic-stress episode frequency, pre- and post-1998.

Frequency (SSE/year)	CFSI SSE	Benchmark SSE
Pre-1998	2.49	1.64
Post-1998	2.06	1.65

6.1.1. Bank deregulation and structural change

One possible explanation for this pattern is a structural change in the U.S. financial system. Indeed, 1998 was marked by a rebuilding of the U.S. financial architecture and opened a new era of financial consolidation and universal banking.[62] The U.S. Financial Services Modernization Act (the Gramm–Leach–Bliley Act) became law in 1999, demolishing the structural separation that formerly existed between commercial banks, investment banks, securities firms, and insurance companies. One of the key beneficiaries early on was the large U.S. commercial bank Citicorp, whose merger with the Travelers Group insurance company, announced in 1998, was allowed to proceed because of the 1998 temporary rule exemption.[63]

There is a strong empirical link between regulation and systemic crises. Kemerrer (1910) states that between 1890 and 1908, there were 28 U.S. banking panics. Miron (1986) finds that prior to the creation of the Federal Reserve, banking panics in the United States were seasonal. Freixas and Rochet (2008) find that many financial crises worldwide have been partly initiated by a global movement toward financial deregulation, an argument supported by a large number of empirical studies of the relationship between crises and regulation. Kaminsky and Reinhart (1999)[64] suggest that "crises may have common origins in the deregulation of the financial system and the boom-bust cycles and asset bubbles that, all too often, accompany financial liberalization." Caprio and Klingebiel (1996)[65] provide cross-country evidence of a natural lag between financial liberalization and adjustment of regulatory structure relative to supervisory practices; this may partially explain the link between deregulation and banking crises. Mishkin (1997)[66] emphasizes this point in discussing the U.S. savings and loan crisis, asserting that "deregulation of a financial system and rapid credit growth can be disastrous if banking institutions and their regulators do not have sufficient expertise to keep risk taking in check." There are numerous empirical studies supporting this connection, for example McKinnon and Huw (1996), Sachs et al. (1996), and Weller (1999). In an extensive empirical review of U.S. bank deregulation, Calomiris (2000) finds that "the single most important factor in banking instability has been the organization of the banking industry."[67]

In addition to the lag between the financial deregulation and regulatory adjustments cited by many authors, another mechanism linking deregulation and systemic risk is risk diversification. Universal banking allows financial intermediaries to grow larger and more diverse, enabling

[62] See Calomiris (2000) and Wilmarth (2002).
[63] Travelers Group Inc. and Citicorp, 84 Federal Reserve Bulletin 985, 1013-14 (1998).
[64] P. 480.
[65] Pp. 24 and 30.
[66] P. 28.
[67] P. 3.

them to benefit from more efficient portfolio diversification to take larger risk. After the Glass–Steagall Act of 1993 and the Bank Holding Company Act of 1956 and preceding the financial deregulation of 1999, U.S. financial intermediaries were not allowed to become universal banks. Calomiris examines U.S. and German universal banking history during the 1870–1914 period and concludes, from the evidence, "that German industrial growth was helped, and American growth was hindered, by their respective financial systems."[68] Modern finance portfolio theory offers an intuitive explanation for this growth. From the viewpoint of an individual universal bank, a larger and more diverse bank is more insulated from the risk of failure and, thus, could be individually "safer." Paradoxically, as more institutions become larger and universally alike, once crisis sets in, contagion among institutions can be expected to persist longer and recovery can be expected to take more time. This is indeed the pattern observed above in Fig. 1, Fig. 10, Fig. 12, Fig. 14, and Table 6. The apparent safety of an individual large, diversified financial institution is also a source of moral hazard and an implicit too-big-to-fail subsidy. Critics of the Financial Modernization Act have argued that institution-specific benefit in portfolio diversification would be offset by the increase in systemic risk that would accompany the growth of universal banks. Reviewing studies of systemic risk in a post-deregulation era, Wilmarth (2002) writes that "doubts about the claimed advantages of universal banks are buttressed by concerns that financial conglomeration will aggravate the problem of systemic risk in financial markets."

We can deduce that the structural break occurred approximately during the announcement and implementation of the U.S. Financial Services Modernization Act. A formal test is appropriate to interpret the break empirically. We use the Quandt likelihood ratio statistic[69] to test for breaks at all dates within the 15-percent trimmed monthly time series. We consider the first-order difference equation with one lag of Z_{CFSI} to test for a structure break.

As Fig. 13 shows, the maximum Quandt likelihood-ratio statistic occurs in May 1998 (F-statistic = 9.007), which is statistically significant at a 1 percent critical value. This is a welcome result because the Financial Services Act was passed by the Senate precisely at this time, leading up to the U.S. Financial Services Modernization Act later in the year.

---

68 P. 265.

69 Quandt (1960).

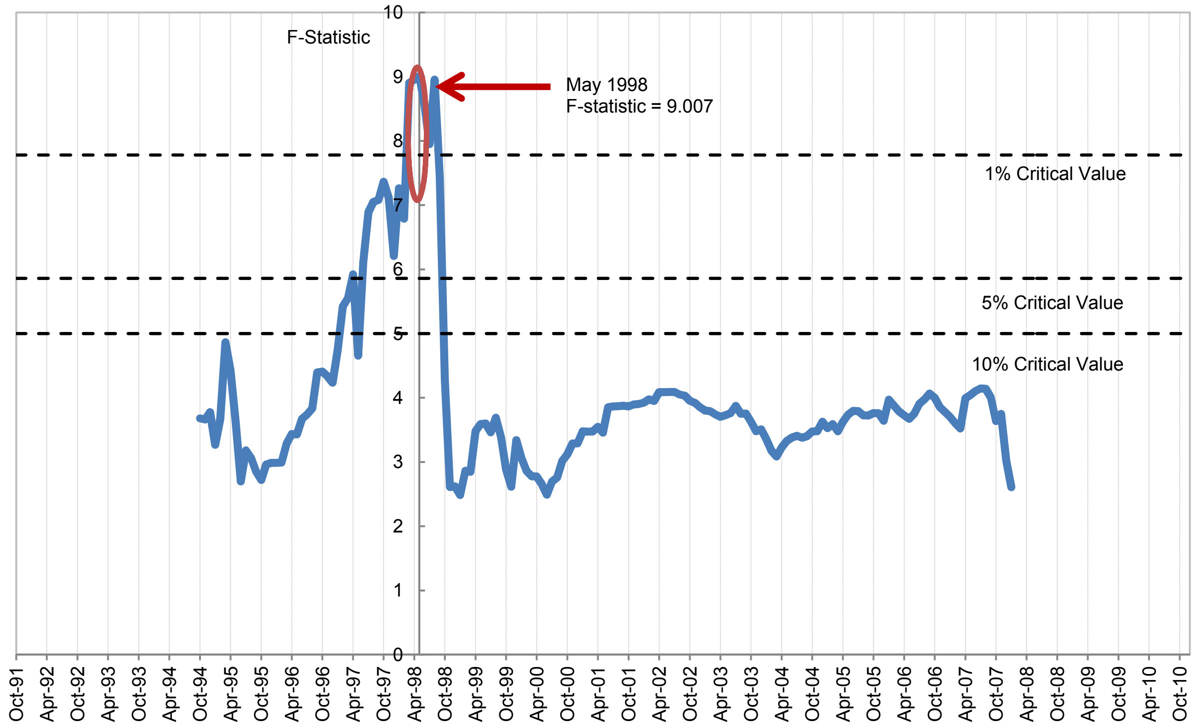


Fig. 13 Quandt likelihood-ratio testing for structural break in Z_{CFSI}.

6.2. Comparisons of CFSI with currently available alternates

Since 2009, development of U.S. financial stress measures has been an expanding area of research. Currently, as many as 12 new financial stress indexes of varying frequency are available. Table 7 and Fig. 14 summarize the currently available alternative series of financial stress, comparing their Z-scores from December 1991 through December 2009. A comparative visual assessment of CFSI against alternative series is promising for CFSI: it tends to distinctly identify stress episodes (see, for example, the period from 4Q: 1991 to 4Q: 1998) and seems to do so earlier than competing indexes (this is clear in the LTCM crisis of 1998 or the subprime crisis of 2007).

Table 7

2010 Summary of Alternative Financial Stress Series.

Index	Source	Frequency	Construction	Components
CFSI	FRBC	d, w, m, FOMC, q	U.S. credit weights	12
SLF FSI	FRBSL	w, m, FOMC, q	Principal components	18
KCFSINDX	FRBKC	m, q	Principal components	11
WRAISTRS	Bloomberg	d, w, m, FOMC, q	Equal weights	7
HSCLOG	Bloomberg	d, w, m, FOMC, q	Equal weights	4
BFCIUS	Bloomberg	d, w, m, FOMC, q	Equal weights	10
Deutsche Bank FCI	http://www.princeton.edu /~mwatson/	q	Weighted Σ of principal components	7
OECD FCI		q	U.S. GDP weighted Σ	6
MacAdv FCI		q	Δ U.S. GDP impulse response	5
GS FCI		m, q	U.S. macro model weights	4
Citi FCI		m, q	U.S. CB weights	6
Mishkin FCI		m, q	Principal components	45

Note: Frequencies are designated as follows: d - daily, w - weekly, m - monthly, q - quarterly. FOMC designation indicates that available data can match FOMC meeting schedule.

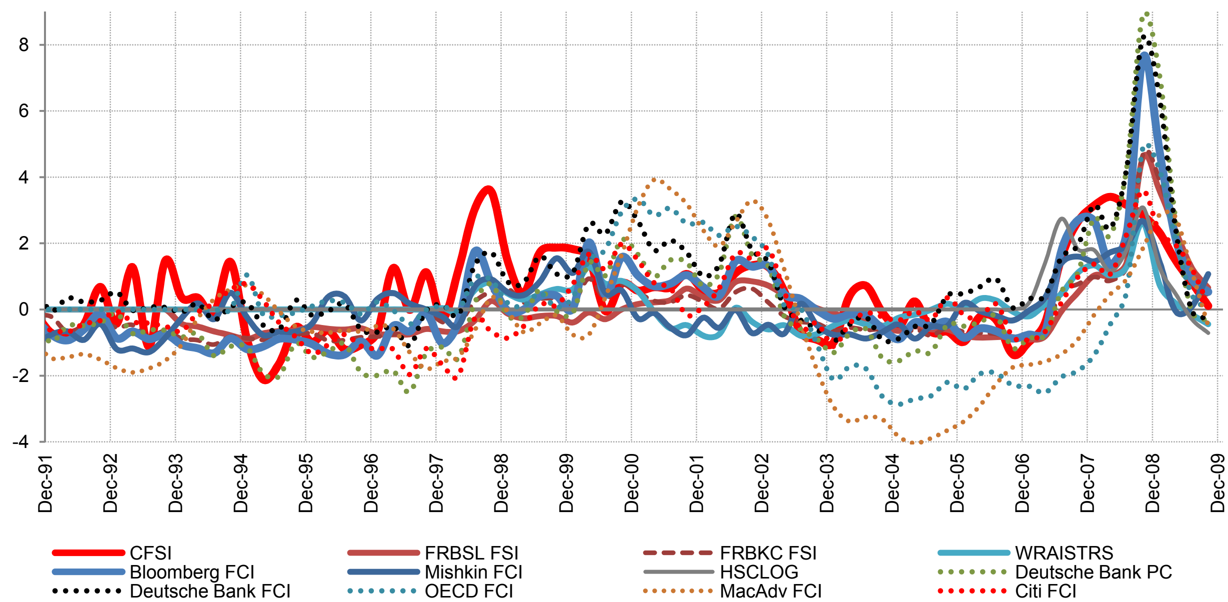


Fig. 14. Comparative summary of alternative financial stress series.

Fig. 15 briefly summarizes this comparison, considering CFSI only against alternative high-frequency financial stress series. It is useful to compare the alternative series, both to assess the relative quality of CFSI and to discover areas for possible future enhancement. In comparing the relative performances of the alternative indexes, we follow the formal process described in section 4.3. At the time of this writing, we had completed the first seven elements of this process but not the monthly error analysis. Table 8 and Table 9 show the results weekly and quarterly error analysis, respectively. Several comments should be made about these results. First, CFSI noise-signal statistics places them firmly in the middle of the competing indexes: perhaps not the best, but not the worst. Second, any observer of Fig. 15 would react to this result with some incredulity: After all, CFSI clearly picks historically relevant stress episodes that are not picked up by other indexes. How then could its noise-signal results not reflect this? Third, the key to explanation is in the component composition of CFSI and alternative indexes. The CFSI is a “true” stress index, only admitting stress components directly observable in the markets. By contrast, the construction of its competitors directly includes one or more market volatility indexes. The error-analysis against volatility-based episodes thus becomes a biased test and should not be seriously considered.

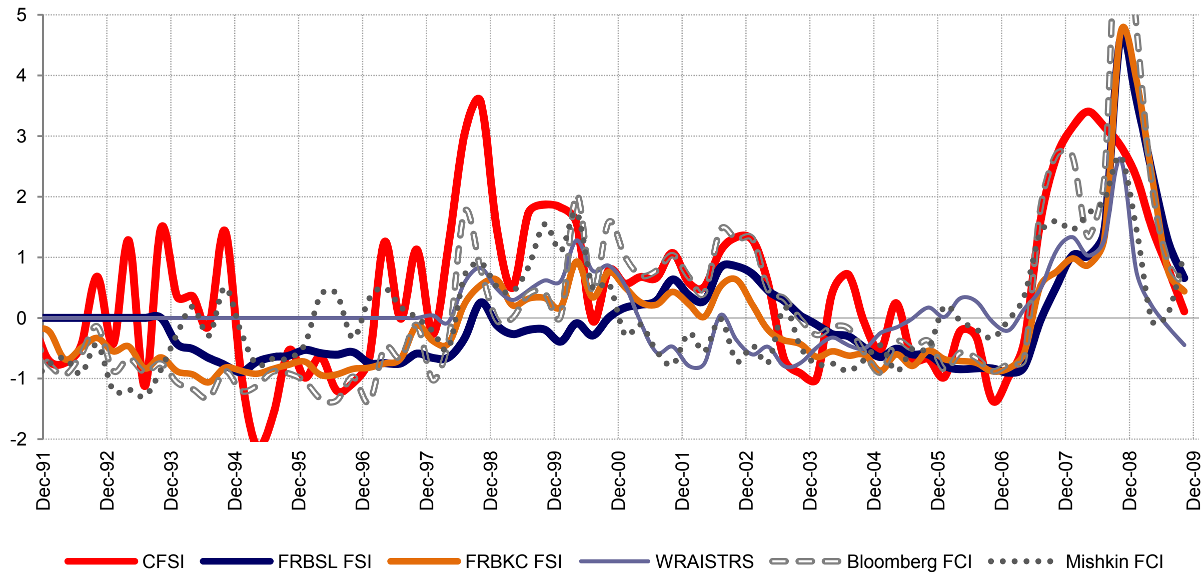


Fig. 15. Comparative summary of high-frequency alternative financial stress series.

Table 8

Error analysis - Monitoring weekly signals.

Method	Observations	Type I error (%) “failure rate”	Type II error (%) “false positive rate”	Noise/signal
CFSI	911	29%	11%	16%
FRBSL FSI	856	26%	6%	8%
WRAISTRS	648	26%	13%	18%
HSCLOG	178	12%	11%	12%
BFCIUS	1013	36%	26%	41%

Table 9

Error analysis - Monitoring quarterly signals.

Method	Observations	Type I error (%) "failure rate"	Type II error (%) "false positive rate"	Noise/signal
CFSI	70	19%	14%	18%
FRBSL FSI	66	18%	8%	9%
FRBKC FSI	74	20%	11%	14%
WRAISTRS	50	22%	20%	26%
HSCLOG	14	7%	7%	8%
BFCIUS	78	32%	28%	42%

7. Applications of CFSI

The two major supervisory applications of CSFI are monitoring and forecasting financial systemic stress. For both, it is reasonable to consider using CFSI in conjunction with alternative stress indexes. This is consistent with the EWS design principle of using different models in parallel.[70] This practice would certainly raise questions of interpretation insofar as different stress indexes disagree or result in different forecasts. To the supervisor, this disagreement is not unwelcome; it requires consideration of the reasons why a particular index indicates stress or leads to forecasts of systemic stress in an EWS setting, whereas another does not. We believe that instead of avoiding the difficult problem of interpreting such mixed signals, the supervisor should confront the possibility that a different type of systemic stress may be highlighted and consider it carefully.

7.1. Use as a dependent variable for early warning system

An important application of CFSI or similar stress indexes is as a dependent variable in an early warning system of systemic stress. We discuss our implementation of this idea in a parallel paper (Oet, Eiben, Gramlich, Miller, and Ong, 2011). In this approach, we build on existing micro- and macroprudential early warning systems to propose a hybrid set of models for systemic risk, incorporating the structural characteristics of the financial system and its feedback mechanisms. In this EWS setting, we explain CFSI by means of data from five large bank holding companies and regress institutional imbalances (constructed as Z-scores) using an optimal lag method. Our EWS utilizes both public and proprietary supervisory data and monitors

[70] For systemic-risk EWS design principles, see Gramlich, Miller, Oet, and Ong (2010).

microprudential information from systemically important institutions to anticipate build-up of systemic stress, captured by CFSI, in the financial markets at large. For the supervisor, this CSFI-based early warning system provides a toolkit of possible institutional supervisory actions that can be used to diffuse the build-up of systemic stress in the financial markets. As in our investigation of appropriate action thresholds for monitoring (see section 7.2), we investigate and suggest levels for action thresholds appropriate for this EWS.

7.1.1. Data Considerations

Data considerations for the dependent variable in an EWS include two key concerns: overcoming data-quality issues; and selecting criteria for possible inclusion of additional data indicators.

Data quality issues typically result from the limited availability of data indicators and the emergence of new data. Hatzius et al. (2010) make an important contribution to this problem by applying unbalanced panel-estimation techniques and enabling significant extension of the constructed time series.[71]

No clear theory exists for inclusion criteria regarding additional data indicators, although precedents and current research lend significant insights for further identification of relevant indicators.[72] Clearly, because CFSI intends to capture increases in systemwide stress, incorporating data from markets beyond the four already included (interbank, credit, equity, and foreign exchange) may be important and deserves further investigation. High-frequency additions from property markets and risk-transfer markets are the most obvious possibilities. Changes in property values reflect a wealth argument for increased stress. Insofar as demand shocks cause property values to fall, these markets become more illiquid in the short run, thus increasing demand for funds, which leads to increased stress. Disequilibria in risk-transfer markets may imply reluctance to bear risk that results from increased uncertainty, thus elevating stress as proportionately more risk is mispriced in the short run.

71 Application of the unbalanced panel technique should be even more critical in overcoming data-quality issues in independent variables.

72 Development of the CFSI series was the only option for SAFE EWS in early 2009. At the time of this writing, as many as 12 distinct financial stress indexes are available. Key among the recent research contributions are studies by Hakkio and Keeton (2009), Hatzius et al. (2010), and Kliesen and Smith (2010).

This study's use of financial beta may be refined by examining the volatility of the worst-performing quintile or decile relative to the S&P 500 index. Utilizing the so-called *tail beta* approach is appealing because it identifies the banks whose performance is most sensitive to broad-based declines in business profitability and which therefore signals stress more pronouncedly. It may also be useful to enhance the set of volatility benchmarks used to validate modifications of CFSI, for example, by finding a more suitable volatility proxy for foreign exchange markets or adding volatility benchmarks specific to a financial institution, for example, the Philadelphia Stock Exchange KBW Bank Index.[73]

7.1.2. Technical Considerations

A principal technical challenge for a dependent variable is correct identification of its intended objective. In particular, the EWS researcher seeks a financial stress series that is, indeed, "dependable" for measuring stress episodes and crisis conditions. Critically, the fact that some aggregate measure of financial stress may be constructed or found is not sufficient basis for using it as a dependent variable in a systemic stress EWS. One approach to this problem, established in practice, is to set a threshold for the series' noise/signal ratio[74] and select an alternative series that minimizes this ratio. Through CFSI, the Federal Reserve Bank of Cleveland SAFE EWS extends this approach by benchmarking against disjoint and aggregate signals of financial markets' volatility. Moreover, this technique enables timing identification of stress episodes and rigorous selection among alternative stress series.

There are several technical issues for continued refinement of CFSI. An important proposal would be to identify structural breaks in the relationships between components of CFSI in order to improve the weights' dynamic over time.

7.2. Monitoring financial stress

As a monitoring tool, CFSI's main benefit is to provide insights on stress in a number of financial markets. A meaningful reading of CFSI rides on the underlying requirement that there

---

73 If a volatility index is included as a sub-component of financial stress, the researcher should be careful to exclude it from the benchmarking schema to avoid bias.

74 Kaminsky and Reinhart (1999), Borio and Lowe (2002, Asset), Borio and Lowe (2002, Crises), Borio (2003), Borio and Drehmann (2009). Alternative benchmarking approaches can be found in Illing and Liu (2003, 2006) and Hatzius et al. (2010). Illing and Liu (2003, 2006) use the rank-ordering results of an expert survey. Hatzius et al. (2010) use efficacy in tracking future GDP growth.

is no aggregate noise in the index. Yet, noise is always present in the financial markets. Therefore, the "no aggregate noise" rule entails the assumption that uncorrelated simultaneous stress in the normally functioning markets tends to be arbitraged away through interconnections and transfers between these markets. When this assumption holds, any remaining stress may indicate a structural ailment in the markets that may not be quickly or easily arbitraged away through normal markets' function.

The problem of index noise is also known as *spurious correlation*. An index change can result from a purely accidental co-movement of components or from some underlying cause. An increase in the index is seen as an indication of a probable increase in underlying stress, although coincidence is always an underlying explanation. For example, the same stress index, taken on two consecutive days, may represent two vastly different levels of stress. One day's stress may be fueled by spurious coincidences in different markets, while another's day's is driven by correlated events arising from a common structural cause. Thus, using CFSI at a very high frequency, such as daily, entails the hazard of including too much noise. On the other hand, using CFSI at a very low frequency, such as annual, becomes questionable because of the interim evolution and transformation of the financial markets. To summarize, in using and interpreting CFSI, it is important to keep in mind that the aggregate stress index is, at best, only a relative weathervane of directional change in aggregate stress. Knowledge of such directional change in CFSI sheds light on monitoring systemic stress and helps us estimate the probability that a given event is economically meaningful.

7.2.1. Seeking optimal index frequency

The choice of an optimal monitoring frequency using CFSI is sensitive to several considerations, the most significant being its ability to filter out idiosyncratic stress episodes. An unbiased evaluation of CFSI's filtering properties requires appropriate independent benchmarks. Because CFSI encompasses high-frequency data from four markets, benchmarks that match these markets and frequency are preferred. The volatility benchmarks can be used to create a set of binary stress signals of varying frequency. In these signals, the existence of stress is indicated by volatility indexes that mirror the four markets in CFSI construction. The following volatility indexes are employed in constructing the binary stress signals: VIX, MOVE, VDAX, LBOX, and BBOX.

We consider an individual market to be in stress if the level of its volatility index surpasses a predetermined threshold. For example, a stress signal may be indicated by the difference in the Z-scores of the individual volatility indexes exceeding ¼ standard deviation. Alternatively, use of ½ standard deviations as a signaling threshold would lead to a different binary series. Which is better? Theoretical precedent for answering this question has been established by earlier research that used the signaling method.[75] By comparing signaling outcomes for various volatility thresholds, we can establish a set of thresholds that minimizes idiosyncratic episodes in a given time period.

This method also allows accommodation of structural breaks in a given time period. For example, as discussed in section 6.1, empirical data on frequency and duration of stress episodes as well as historical evidence of U.S. financial deregulation suggest that the 1998 period may constitute a structural break that has affected the duration—but not the frequency—of systemic stress episodes. For this reason, it is appropriate to explicitly include a structure break in the binary stress series derived from the four markets. A more restrictive threshold of ½ standard deviation can be used for the period 4Q: 1991–Q1: 1998; a less restrictive threshold of ¼ standard deviation can be used for the period 2Q: 1998–4Q: 2010.

We apply the signaling technique consistently, using the process described in section 4.3 (4). Stress in one market does not necessarily indicate widespread stress. We consider the system to be in stress if more than one market sends a signal based on these restrictions or if any market signals stress for consecutive periods. These criteria can be applied at various frequencies to optimally filter out idiosyncratic episodes that do not warrant policy actions. According to these guidelines, daily measures are far too impulsive to permit drawing any cogent conclusions; consequently, we test filtering capability at weekly, biweekly, monthly, and quarterly frequencies for both measures of stress; the results are shown in Table 10.

Table 10

Benchmarked stress episodes as a function of monitoring frequency.

Frequency	Stress Episodes	Non-Stress Episodes
Quarterly	8	69
Monthly	16	215
Biweekly	6	497
Weekly	12	995

[75] Kaminsky, Lizondo, and Reinhart (1998), Kaminsky and Reinhart (1996, 1999), Borio and Lowe (2002, Asset), Borio and Lowe (2002, Crises), Borio and Drehmann (2009).

It is not sufficient to simply create a benchmark series of stress signals. To determine the optimal monitoring regime using CFSI, we proceed to establish and test alternative CFSI-based systems for rating systemic stress. The optimal system can facilitate monitoring and guide the interpretation of systemic stress, as in Bordo's use of five grade ranges in the Financial Conditions Index. Furthermore, the rating approach can help determine the optimal CFSI monitoring frequency.

7.2.2. CFSI as a rating system: Classifying grades of stress thresholds

The logic of a CFSI-based rating system is as follows: When the CFSI has a low Z-score, it is unlikely that we are experiencing a stress episode; when the Z-score is high, it is more likely; and when it is moderate, the diagnosis is unclear. As a result, a CFSI rating system that effectively differentiates Z-score ranges vis-à-vis frequency of observation can be ideal for selecting optimal monitoring frequencies and for policy making.

To construct such a system, we divide the range of the Z-score of CFSI into grades, determine how many observations fall into each grade, and compare those observations to the benchmark binary stress series. We use two metrics for the effectiveness of the rating system: Somer's D and the area under the receiver operating characteristic (ROC) curve. Somer's D is a broad metric that shows the degree to which a low rating within the system contains more stress events.

$$Somer's\ D = 2\left[\frac{P(Rating\ Grade_{stress} > Rating\ Grade_{no\ stress}) + P(Rating\ Grade_{stress} = Rating\ Grade_{no\ stress})}{2}\right] - 1 \qquad (20)$$

The area under the ROC curve is a measure of the rating system's differentiating power. For a perfect rating system, the ROC statistic measures 100, while ROC for a rating system that is no better than random measures 50.

The results of testing are shown in Table 11. The optimal number of grades depends on the frequency chosen. Overall, it is clear that the rating system at a monthly frequency with four grades is optimal: It has an ROC of 85.6 and a Somer's D of 77.1; it is not equivalent to a random rating system at 5 percent significance.

Table 11

Results of non-parametric testing for optimal CFSI-based rating system.

		# of grades=2	# of grades=3	# of grades=4	# of grades=5
Quarterly	Somer's D	31.2	55.4*	49.8*	48*
	ROC	66	77.7	74.9*	74*
Monthly	Somer's D	59.6*	52.2*	77.1*	54.5*
	ROC	79.8*	76.1*	85.6*	77.3*
Biweekly	Somer's D	17.4	40.2	39.4	33.5
	ROC	58.7	71.1	69.7	66.7
Weekly	Somer's D	21.9	44.1*	42*	51.4*
	ROC	60.9	72*	71*	75.7*

*Indicates rating system is not equivalent to a random rating system

7.2.3. Application of the CFSI rating system to quantifying the probability of stress episodes

Because we have determined that the optimal number of grades for the rating system is four and the optimal threshold is ½ standard deviation through 1997 and ¼ standard deviation thereafter, we can look toward to applying these specifications. The continuous CFSI series can predict a stress episode based on the external stress series that has been constructed. To do so, we have used a probit model to obtain the implied probability of a stress episode. This could be beneficial for policy as well as future modeling. This model takes the form of

$$Stress = \alpha_0 + \alpha_1 Z_{CFSI} + u \qquad (21)$$

$$Stress = -1.9929115 + 0.524282 Z_{CFSI} + u \qquad (22)$$

The regression results are shown in tables 16 and 17 in Appendix C. On the basis of these parameters, we can create a continuous series of the implied probability of a stress episode as measured by the external benchmark series of stress (see Table 12 and Fig. 16). As of 4Q: 2010, the range of CFSI Z-scores was divided into grades; the minimum thresholds were 2.38 (grade 4); 0.82 (grade 3); –0.73 (grade 2); and –0.74 (grade 1) A higher grade implies a greater probability of stress; it is clear that the probability of a systemic stress episode has decreased significantly since the peak of the subprime crisis.

In comparison with Bordo, Dueker, and Wheelock's (2000) conjecture for severely and moderately distressed thresholds, our results indicate the need to revise the recommended thresholds. Moderate distress threshold should increase from 0.75 standard deviations to 0.82 standard deviations. Severe distress threshold should increase from 1.5 standard deviations to 2.38 standard deviations, based on the monthly CFSI series in the period between 4Q: 1991and 4Q: 2010.

Table 12

Probability of systemic stress episode by CFSI grade.

CFSI rating grades	Range	Probability of systemic stress at grade threshold
Grade 1	$Z_{CFSI} \leq -0.74$	0.07%
Grade 2	$-0.74 < Z_{CFSI} < 0.82$	0.87%
Grade 3	$0.82 < Z_{CFSI} < 2.38$	5.92%
Grade 4	$Z_{CFSI} \geq 2.38$	22.84%

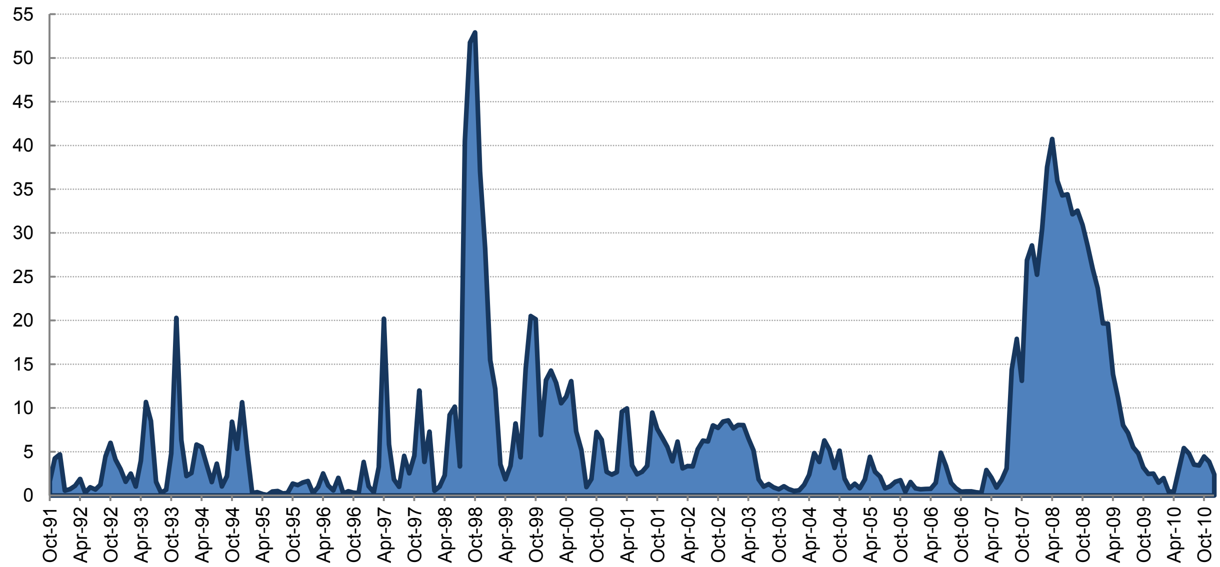


Fig. 16. Implied probability of systemic stress episode.

7.3. Crisis dating

The signaling process described above can be used to establish an objective method of crisis dating. The population of stress episodes determined by this method will vary according to the choices of signaling thresholds and frequency. Table 18 in Appendix C lists the stress episodes determined in the initial selection of credit weights by the CFSI construction method. During this process, a threshold of 1 standard deviation at weekly frequency was used. As discussed in section 7.2, a more efficient stress episode filtering occurs at a monthly frequency using ½ standard deviation during the period from 4Q:1991 to 1Q: 1998 and ¼ standard deviation thereafter.

References

Bernanke, B., Gertler, M., 1990. Financial fragility and economic performance. *Quarterly Journal of Economics* 105 (1): 87–114.

Bernanke, B., Gertler, M., 1995. Inside the black box: The credit channel of monetary policy transmission. *Journal of Economic Perspectives* 9 (4): 27–48.

Bolton, P., Freixas, X., 2000. Equity, bonds, and bank debt: Capital structure and financial market equilibrium under asymmetric information. *Journal of Political Economy* 108 (2): 324–351.

Bordo, M.D., Dueker, M., Wheelock, D., 2000. Aggregate price shocks and financial instability: An historical analysis. Federal Reserve Bank of St. Louis, Working Paper 2000-005B.

Borio, C., 2003. Toward a macroprudential framework for financial supervision and regulation? Bank for International Settlements, Working Paper, No. 128, Basel.

Borio, C., Drehmann, M., 2009. Assessing the risk of banking crises – revisited. Bank for International Settlements *Quarterly Review* (March): 29–46.

Borio, C., Lowe, P., 2002, Asset. Asset prices, financial and monetary stability: Exploring the nexus. BIS – Bank for International Settlements Working Paper, No. 114, Basel.

Borio, C., Lowe, P., 2002, Crises. Assessing the risk of banking crises. BIS – Bank for International Settlements, *Quarterly Review* (December): 43–54.

Boyd, J., De Nicolò, G., Loukoianova, E., 2010. Banking crises and crisis dating: Theory and evidence. CESifo Working Paper No. 3134, July.

Brave, S., Butters, R.A., 2011. Monitoring financial stability: A financial conditions index approach. Federal Reserve Bank of Chicago, *Economic Perspectives* Q1: 22–43.

Bussière, M., Fratzscher, M., 2002. Towards a new early warning system of financial crises. European Central Bank Working Paper No. 145, Frankfurt.

Calomiris, C.W., 2000. U.S. bank deregulation in historical perspective. Cambridge University Press.

Caprio, G., Klingebiel, D., 1996. Bank insolvencies: Cross-country experiences. World Bank Policy Research Working Paper No. 1620, Washington.

De Bandt, O., Hartmann, P., 2000. Systemic risk: A survey. European Central Bank Working Paper No. 35, Frankfurt, November.

Demirgüç-Kunt, A., Detragiache, E., 1998. The determinants of banking crises in developing and developed countries. IMF Staff Papers, 45 (1): 81–109.

Demirgüç-Kunt, A., Detragiache, E., 2005. Cross-country empirical studies of systemic bank distress: A survey. IMF Working Paper No. 96/05, Washington.

English, W., Tsatsaronis, K., Zoli, E., 2005. Assessing the predictive power of measures of financial conditions for macroeconomic variables. In: Bank for International Settlements, Investigating the relationship between the financial and real economy, BIS papers No. 22: 228–252 Basel.

Estrella, A., Hardouvelis, G., 1991. The term structure as a predictor of real economic activity. *Journal of Finance* 46 (2): 555–576.

Estrella, A., Mishkin, F., 1996. The yield curve as a predictor of U.S. recessions. Federal Reserve Bank of New York, *Current Issues in Economics and Finance* 2 (7): 1–6

Estrella, A., Trubin, M., 2006. The yield curve as a leading indicator: Some practical issues. Federal Reserve Bank of New York, *Current Issues in Economics and Finance*, 12 (5).

Freixas, X., Rochet, J.-C., 2008. *Microeconomics of Banking*. 2nd ed. Cambridge, Mass.: The MIT Press.

Gramlich, D., Miller, G., Oet, M., Ong, S., 2010. Early warning systems for systemic banking risk: Critical review and modeling implications. *Banks and Bank Systems*, 5(2): 199–211.

Granger, C.W.J., 1969. Investigating causal relations by econometric models and cross-spectral methods. *Econometrica* 37 (3): 424–438.

Hakkio, C.S., Keeton, W.R., 2009. Financial stress: What is it, how can it be measured, and why does it matter? Federal Reserve Bank of Kansas City *Economic Review*, 2Q: 5–50.

Hanschel, E., Monnin, P., 2005. Measuring and forecasting stress in the banking sector: Evidence from Switzerland. Bank for International Settlements, Working Paper No. 22, Basel.

Hatzius, J. Hooper, P., Mishkin, F.S., Schoenholtz, K.L., Watson, M.W., 2010. Financial conditions indexes: A fresh look after the financial crisis. NBER Working Paper Series, w16150.

Haubrich, J., 2006. Does the yield curve signal recession? Federal Reserve Bank of Cleveland, *Economic Commentary,* April.

Haubrich, J., Bianco, T., 2011. The yield curve as a predictor of economic growth. Federal Reserve Bank of Cleveland, April. <http://www.clevelandfed.org/research/data/yield_curve/2011/0511/index.cfm>.

Hendricks, D., Kambhu, J., Mosser, P. 2007. Systemic risk and the financial system, in Federal Reserve Bank of New York *Economic Policy Review*, 13 (2): 65–80.

Holmström, B., Tirole, J., 1997. Financial intermediation, loanable funds, and the real sector. *Quarterly Journal of Economics* 112 (3): 663–691.

Honohan, P., Klingebiel, D., 2003. The fiscal cost implications of an accommodating approach to banking crises. *Journal of Banking and Finance,* 27 (8):1539–1560.

Illing, M., Liu, Y., 2003. An index of financial stress for Canada. Bank of Canada Working Paper No. 2003-14, Ottawa, June.

Illing, M., Liu, Y., 2006. Measuring financial stress in a developed country: An application to Canada. *Journal of Financial Stability*, 2 (4):243–265.

IMF - International Monetary Fund, 2009, Responding. Global Financial Stability Report 04/2009, Responding to the financial crisis and measuring systemic risks. Washington, April.

Ishihara, Y. 2005. Quantitative analysis of crisis: Crisis identification and causality. World Bank Policy Research Working Paper No. 3598. Washington, May.

Kambhu, J., Weidman, S., Krishnan, N., 2007. Introduction: New directions for understanding systemic risk. Federal Reserve Bank of New York, *Economic Policy Review*, 13 (2): 3–14.

Kaminsky, G.L., Lizondo, S., Reinhart, C.M., 1998. Leading Indicators of Currency Crises. IMF Staff Paper, 45 (1):1–48.

Kaminsky, G.L., Reinhart, C.M., 1996. The twin crises: The causes of banking and balance-of-payments problems. Unpublished manuscript, Federal Reserve Board and International Monetary Fund, Washington.

Kaminsky, G.L., Reinhart, C.M., 1999. The twin crises: The causes of banking and balance-of-payments problems. *American Economic Review*, 89 (3): 473–500.

Kantor, L., Caglayan, M., 2002. Using equities to trade FX: Introducing the LCVI. JP Morgan Global Foreign Exchange Research, Investment Strategies: No. 7.

Kemmerer, E.W., 1910. Seasonal Variations in the Relative Demand for Money and Capital in the United States. Washington: Government Printing Office.

Kliesen, K.L., Smith, D.C., 2010. Measuring financial market stress. Federal Reserve Bank of St. Louis Economic Synopses, No.2.

Laeven, L., Valencia, F., 2008. Systemic banking crises: A new database. IMF Working Paper 08/224, September.

McKinnon, R.I., Pill, H.,1996. Credible liberalizations and international capital flows: The overborrowing syndrome. In: Takatoshi, I., Krueger, A.O., (Eds.) Financial Deregulation and Integration in East Asia. Chicago: University of Chicago Press.

Miron, J.A., 1986. Financial panics, the seasonality of the nominal interest rate, and the founding of the Fed. *American Economic Review* 76 (1): 125–140.

Mishkin, F.S., 1997. Understanding financial crises: A developing country perspective (June). NBER Working Paper Series, Vol. w5600.

Oet, M., Eiben, R., 2009. Financial stress index. Mimeo, Federal Reserve Bank of Cleveland.

Oet, M., Eiben, R., Bianco, T., Gramlich, D., Ong, S., Wang, J., 2011. SAFE: An early warning system for systemic banking risk. Federal Reserve Bank of Cleveland Working Paper 11-29.

Quandt, R.E., 1960.Tests of the hypothesis that a linear regression system obeys two separate regimes. *Journal of the American Statistical Association* 55 (290): 24–330.

Reinhart, C., Rogoff, K., 2008. This time is different: A panoramic view of eight centuries of financial crises. NBER Working Paper No. 13882, March.

research.stlouisfed.org, 2010. Appendix. Federal Reserve Bank of St. Louis National Economic Trends, January.

Rosenberg, M.R., 2008. Bloomberg, Financial Conditions Watch, 1 (1), August.

Sachs, J., Tornell, A., Velasco, A., 1996. Financial crises in emerging markets: The lessons from 1995. *Brookings Papers on Economic Activity*, Vol. 1, 147–215.

Song, F., Thakor, A.V., 2010. Financial system architecture and the co-evolution of banks and capital markets. *The Economic Journal* 120:1021–1055.

Swiston, A., 2008. A U.S. financial conditions index. International Monetary Fund, Working Paper WP/08/161, June.

Weller, C. 2001. Financial crises after financial liberalization: Exceptional circumstances or structural weakness? *Journal of Development Studies*, October, 98–127.

Appendix A. Description of data

Table 13

Financial stress index data sources.

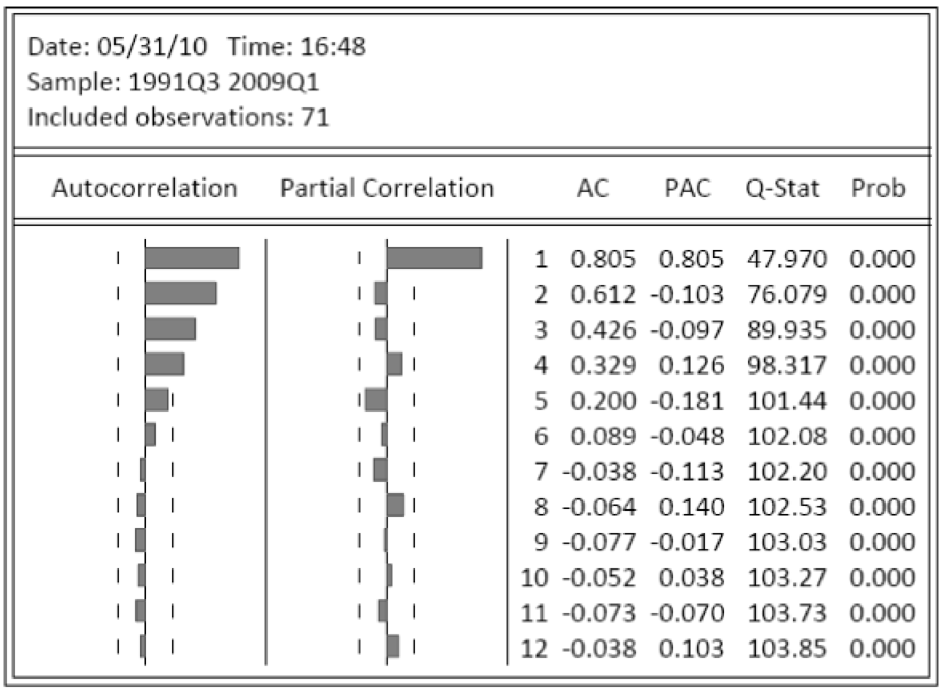
INDICATOR	DATA	SOURCE	VARIABLE	START DATE
INTERBANK MARKETS				
Financial Beta	Beta (S&P Financial & S&P 500)	Calculated: BLOOMBERG data	BETA	09/13/1990
	S&P 500 Financials Total Return Index	BLOOMBERG (SPTRFINL Index)	SP500_F	09/13/1990
	S&P 500 Total Return Index	BLOOMBERG (SPXT Index)	SP500	02/9/1988
Bank Bond Spread	(10 Year A Bank Bond Index)-(10YT-Notes)	Calculated: BLOOMBERG & FRED data	A_TN_S	09/26/1991
	10 Year A Bank Bond Index	BLOOMBERG (C07010Y Index)	A_10Y	09/26/1991
	10-year constant maturity Treasury Rate	FRED (DGS10)	T_10Y	01/02/1962
Interbank Liquidity Spread	TED Spread (3mo LIBOR - 3mo Tbill)	Calculated: BLOOMBERG (US0003m), FRED (DTB3)	TED_S	01/02/1985*
	3 mo LIBOR rate	BLOOMBERG (US0003m)	3moL	01/02/1985*
	US 90-day Treasury bill: Secondary Market Rate	FRED (DTB3)	3moTB	01/04/1954
Interbank Cost of Borrowing	3mo LIBOR-FedFundsTargetRate Spread	Calculated: BLOOMBERG (US0003m, FDTR)	L_FFR_S	01/02/1985*
	3 mo LIBOR rate	BLOOMBERG (US0003m)	3moL	01/02/1985*
	Fed Funds Target rate	BLOOMBERG (FDTR)	FFR	01/02/1985*
FX MARKETS				
Weighted Dollar Crashes	Weighted Dollar Crashes vs. Major Currency FX	Calculated, FRED (DTWEXM)	WTD_DCR	01/02/1973
	Trade Weighted $US Exchange Index: Major Currencies	FRED (DTWEXM)	DX_IND	01/02/1973
CREDIT MARKETS				
Covered Interest Spread	US-UK covered interest rate differential	BOE and FRED (DTB3)	UKUS_90S	01/02/1979
	UK 90-day Treasury bill rate	BOE Website	UK_90	01/31/1975
	90-day forward rate for the UK-US exchange rate	BOE Website	UKUS_90F	01/02/1979
	Spot rate for the UK-US dollar exchange rate	BOE Website	UKUS_S	01/02/1975
	U.S. Government 90-day Treasury bill rate (secondary market rate)	FRED (DTB3)	US_90	01/04/1954
Corporate Bond Spread	Moody's Seasoned Aaa Corporate Bond Yield - 10-Year T Note	FRED (DAAA - DGS10)	AAA10Y_S	01/03/1983
	Moody's Aaa Corporate Bond Yield	FRED (DAAA)	AAA10Y	01/03/1983
	10-year constant maturity Treasury Rate	FRED (DGS10)	T_10Y	01/02/1962
Liquidity Spread	30-day Moving Average of 3 Month U.S. Treasuries Bid-Ask Spread	Calculated, BLOOMBERG (USGG3M)	LIQ_S	01/02/1985*
	3 Month U.S. Treasuries Generic Index Bid Price	BLOOMBERG (USGG3M)	US90_B	06/17/1983
	3 Month U.S. Treasuries Generic Index Ask Price	BLOOMBERG (USGG3M)	US90_A	06/17/1983
	Bid-Ask spread on 90-day U.S. government Treasury bills	Calculated, BLOOMBERG (USGG3M)	TB_BA_S	06/17/1983
Commercial Paper - T-Bill Spread	(AA Commercial Paper) - (3 Month Treasury Bill Secondary Market Rate)	Calculated, FRED (DCPN3M U WCP3M - DTB3)	AACPTB_S	01/02/1985*
	90-day commercial paper AA (patched)	Patched data, FRED (DCPN3M and WCP3M)	AACP_90	04/09/1971
	US 90-day Treasury bill: Secondary Market Rate	FRED (DTB3)	US_90	01/04/1954
Treasury Yield Curve Spread	30-day Moving Average of 10 Year Treasuries - 3-month T-Bills	Calculated, FRED (DGS10, DTB3)	TREAS_S	02/01/1984*
	10-Year Treasury Note Yield at Constant Maturity (Avg. % p.a.)	FRED (DGS10)	T_10Y	01/02/1962
	U.S. Government 90-day Treasury bill rate (secondary market rate)	FRED (DTB3)	US_90	01/04/1954
	3-Month Treasury: Sec. Mkt Rate, Bond Equivalent Yield	Calculated, FRED (DTB3)	US_90_BE	01/03/1984*
EQUITY MARKETS				
Stock Market Crashes	Stock Market Crashes - S&P 500 Financials	Calculated, BLOOMBERG (S5FINL Index)	STMC_SPF	09/10/1990
	S&P 500 Financials Price Index	BLOOMBERG (S5FINL Index)	STPI_SPF	09/11/1989

* Start data set by data request specification.

Table 14

FSI construction: risk assignments and variable names.

INDICATOR	RANK NAME	RANK OF HIGHEST OBSERVED VALUE	CDF NAME
Financial beta	RKFSI1A	4,237	CDFIS1A
Bank bond spread	RKFSI2E	4,237	CDFIS2E
Interbank liquidity spread	RKFSI10A	4,237	CDFIS10A
Interbank cost of borrowing	RKFSI11A	4,237	CDFIS11A
Weighted dollar crashes	RKFSI3A	1	CDFIS3A
Covered interest spread	RKFSI4A	4,237	CDFIS4A
Corporate bond spread	RKFSI5A	4,237	CDFIS5A
Liquidity spread	RKFSI6A	4,237	CDFIS6A
Commercial paper/T-Bill spread	RKFSI7A	4,237	CDFIS7A
Treasury yield curve spread	RKFSI8A	1	CDFIS8A
Stock market crashes	RKFSI9A	1	CDFIS9A


Date: 05/31/10 Time: 16:48
Sample: 1991Q3 2009Q1
Included observations: 71

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.805	0.805	47.970	0.000
		2	0.612	-0.103	76.079	0.000
		3	0.426	-0.097	89.935	0.000
		4	0.329	0.126	98.317	0.000
		5	0.200	-0.181	101.44	0.000
		6	0.089	-0.048	102.08	0.000
		7	-0.038	-0.113	102.20	0.000
		8	-0.064	0.140	102.53	0.000
		9	-0.077	-0.017	103.03	0.000
		10	-0.052	0.038	103.27	0.000
		11	-0.073	-0.070	103.73	0.000
		12	-0.038	0.103	103.85	0.000

Fig. 17. Correlogram of quarterly FSI.

Appendix B. Stationarity of quarterly CFSI

Since nonstationary process may be due to a random walk, random walk with drift, or random walk with drift around a stochastic trend. We conduct several econometric tests for the three different forms under three different null hypotheses:

- Case 1. Test quarterly $CFSI_t$ as a random walk:

$$CFSI_t = \delta CFSI_{t-1} + u_t$$

- Case 2. Test quarterly $CFSI_t$ as a random walk with drift

$$CFSI_t = \beta_o + \delta CFSI_{t-1} + u_t$$

- Case 3. Test quarterly $CFSI_t$ as a random walk with drift around a stochastic trend.

$$CFSI_t = \beta_o + \beta_1 t + \delta CFSI_{t-1} + u_t$$

In each case, the null hypothesis is that $\delta = 0$, that is there is a unit root and time series is nonstationary:

$$\begin{cases} H_0: \delta = 0 \mid time\ series\ is\ nonstationary \\ H_A: \delta < 0 \mid time\ series\ is\ stationary \end{cases}$$

If the null hypothesis is rejected for case 1, then CFSIt is stationary with a zero mean. If the null hypothesis is rejected for case 2, then CFSIt is stationary with a nonzero mean. If the null hypothesis is rejected for case 3, then CFSIt is stationary around a deterministic trend. As Table 15 shows, quarterly CFSIt can be considered stationary with a nonzero mean at 5% critical level, or weakly stationary around a deterministic trend at 10% critical level.

Table 15

Unit Root tests of quarterly $CFSI_t$

Unit Root tests			DF	ADF	PP	KPSS	ERS	NP MZa	MZt	MSB	MPT
$CFSI_t$ as a random walk	Test statistic			-0.62	-0.56						
	critical values	1% level		-0.60	-2.60						
		5% level		-1.95	-1.95						
		10% level		-1.61	-1.61						
$CFSI_t$ as a random walk with drift	Test statistic		-2.55	-0.62	-2.72	0.40	2.23	-11.1	-2.36	0.21	2.20
	critical values	1% level	-2.60	-3.52	-3.52	0.74	1.91	-13.80	-2.58	0.10	1.78
		5% level	-1.95	-2.90	-2.90	0.46	3.04	-8.10	-1.98	0.23	3.17
		10% level	-1.61	-2.59	-2.59	0.35	4.05	-5.70	-1.62	0.28	4.45
$CFSI_t$ as a random walk with drift around a stochastic trend	Test statistic		-2.91	-2.89	-3.11	0.06	6.56	-14.00	-2.62	0.19	6.66
	critical values	1% level	-3.68	-4.09	-4.09	0.22	4.24	-23.80	-3.42	0.14	4.03
		5% level	-3.11	-3.47	-3.47	0.15	5.68	-17.30	-0.91	0.10	5.48
		10% level	-2.82	3.16	-3.16	0.12	6.78	-14.20	-2.62	0.19	6.67

Note: DF – Dickey Fuller test; ADF – Augmented Dickey Fuller test; ERS – Elliott-Rothenberg-Stock test; PP – Phillips-Perron test; KPSS – Kwiatkowski-Phillips-Schmidt-Shin test; NP – Ng-Perron test.

Appendix C. CFSI properties

Table 16.

Regression results of monthly CFSI-based rating system.

Variable	Coefficient	Std. Error	z-Statistic	Prob.
CFSI_Z	0.524282	0.117728	4.453346	0.0000
C	-1.992915	0.209251	-9.524046	0.0000
McFadden R-squared	0.208572	Mean dependent var		0.069264
S.D. dependent var	0.254454	S.E. of regression		0.232642
Akaike info criterion	0.415770	Sum squared resid		12.39400
Schwarz criterion	0.445575	Log likelihood		-46.02146
Hannan-Quinn criter.	0.427791	Deviance		92.04291
Restr. deviance	116.2998	Restr. log likelihood		-58.14990
LR statistic	24.25688	Avg. log likelihood		-0.199227
Prob(LR statistic)	0.000001			
Obs with Dep=0	215	Total obs		231
Obs with Dep=1	16			

Table 17

Granger test precedence results of alternative stress indexes vs. volatility series.

	CFSI = FSICW			**FSIEVW**			**FSIEW**			**FSIFA**		
	Obs	**F-Statistic**	**Prob.**	**Obs**	**F-Statistic**	**Prob.**	**Obs**	**F-Statistic**	**Prob.**	**Obs**	**F-Statistic**	**Prob.**
QUARTERLY (lags=2)												
FSI --> BBOX	**32**	**3.56471**	**0.0423††**	24	3.70077	0.0439	**24**	**5.07458**	**0.0171††**	**24**	**4.63041**	**0.023††**
FSI --> LBOX	52	1.1284	0.3322	**50**	**2.50094**	**0.0933††**	50	1.87745	0.1648	50	1.80398	0.1763
FSI --> MOVE	75	1.39051	0.2557	67	0.9235	0.4025	67	0.89577	0.4135†	67	1.27585	0.2864†
FSI --> VDAX	73	0.39496	0.6752	65	0.00417	0.9958	65	0.01729	0.9829	65	0.09867	0.9062
FSI --> VIX	75	0.01732	0.9828	67	0.06357	0.9385	67	0.12878	0.8794	67	0.59341	0.5555
BBOX --> FSI	**32**	**0.0978**	**0.9072††**	24	1.48745	0.2511	**24**	**0.44086**	**0.6499††**	**24**	**0.08663**	**0.9174††**
LBOX --> FSI	52	0.91645	0.407	**50**	**0.66771**	**0.5179††**	50	1.08596	0.3463	50	5.13669	0.0098
MOVE --> FSI	75	0.82354	0.4431	67	0.08121	0.9221	67	0.18482	0.8317†	67	0.52872	0.592†
VDAX --> FSI	73	1.72123	0.1865	65	0.0836	0.9199	65	0.06992	0.9325	65	0.07724	0.9258
VIX --> FSI	75	0.60951	0.5465	67	0.52715	0.5929	67	0.53806	0.5866	67	0.14379	0.8664
MONTHLY (lags=2)												
FSI --> BBOX	**102**	**2.44857**	**0.0917††**	**78**	**3.59052**	**0.0326††**	**78**	**3.96121**	**0.0233††**	**78**	**4.82196**	**0.0108††**
FSI --> LBOX	162	1.83975	0.1623	156	2.7154	0.0694†	156	2.00492	0.1382	156	1.36427	0.2587
FSI --> MOVE	231	1.23157	0.2938†	207	1.06766	0.3457†	207	0.19455	0.8234†	207	0.49647	0.6094
FSI --> VDAX	227	0.93624	0.3936	203	0.14794	0.8626	203	0.64809	0.5241	203	0.48729	0.615
FSI --> VIX	231	1.10481	0.3331	207	0.18631	0.8302	207	0.80026	0.4506	207	1.29159	0.2771
BBOX --> FSI	**102**	**0.0832**	**0.9202††**	**78**	**0.6996**	**0.5001††**	**78**	**0.2398**	**0.7874††**	**78**	**0.60011**	**0.5514††**
LBOX --> FSI	162	2.12791	0.1225	156	1.0253	0.3612†	156	2.44035	0.0906	156	5.33357	0.0058
MOVE --> FSI	231	0.31635	0.7291†	207	0.27221	0.762†	207	0.35783	0.6996	207	1.308	0.2726
VDAX --> FSI	227	1.80085	0.1676	203	4.53028	0.0119	203	1.80141	0.1678	203	2.80737	0.0628
VIX --> FSI	231	3.094	0.0472	207	6.27212	0.0023	207	3.34824	0.0371	207	4.32203	0.0145
WEEKLY (lags=2)												
FSI --> BBOX	**454**	**1.58082**	**0.2069††**	351	3.79732	0.0234	351	2.7948	0.0625†	**351**	**3.27128**	**0.0391††**
FSI --> LBOX	**711**	**6.93859**	**0.001††**	689	3.97751	0.0192†	689	7.75419	0.0005	689	3.93679	0.02
FSI --> MOVE	**1013**	**4.2798**	**0.0141††**	**910**	**2.17712**	**0.114††**	910	1.02608	0.3588	910	1.57649	0.2073†
FSI --> VDAX	1000	0.00941	0.9906	897	1.04032	0.3538	897	0.6002	0.5489	897	0.11188	0.8942
FSI --> VIX	1013	0.04918	0.952	910	0.16007	0.8521	910	0.20765	0.8125	910	0.71705	0.4885
BBOX --> FSI	**454**	**0.03664**	**0.964††**	351	1.69379	0.1853	351	0.81867	0.4419†	**351**	**0.37883**	**0.6849††**
LBOX --> FSI	**711**	**0.25426**	**0.7756††**	689	1.0077	0.3656†	689	2.42864	0.0889	689	3.99666	0.0188
MOVE --> FSI	1013	0.62025	0.538	**910**	**0.19263**	**0.8248††**	910	0.68708	0.5033	910	0.36572	0.6938†
VDAX --> FSI	1000	2.4502	0.0868	897	2.22256	0.1089	897	1.57922	0.2067	897	1.6318	0.1962
VIX --> FSI	1013	3.45097	0.0321	910	3.21946	0.0404	910	2.16228	0.1157	910	3.11983	0.0446
DAILY (lags=2)												
FSI --> BBOX	**2286**	**3.03456**	**0.0483††**	1768	8.3364	0.0002	1768	4.78798	0.0084	**1768**	**5.22381**	**0.0055††**
FSI --> LBOX	**3567**	**2.97617**	**0.0511††**	3454	2.40363	**0.0905†**	3454	1.36663	0.2551	3454	0.7728	0.4618
FSI --> MOVE	**5081**	**7.97978**	**0.0003††**	4563	2.93107	0.0534	4563	0.49054	0.6123	4563	0.85684	0.4246
FSI --> VDAX	5013	41.704	0.0000	4495	2.86819	0.0569	4495	2.45547	0.0859	4495	2.59337	0.0749
FSI --> VIX	5081	3.74891	0.0236	4563	2.46105	0.0855	4563	2.01197	0.1338	4563	4.07116	0.0171
BBOX --> FSI	**2286**	**0.19844**	**0.82††**	1768	3.98716	0.0187	1768	1.47854	0.2283	**1768**	**0.77944**	**0.4588††**
LBOX --> FSI	**3567**	**0.74201**	**0.4762††**	3454	1.01153	0.3638†	3454	1.60827	0.2004	3454	2.89532	0.0554
MOVE --> FSI	5081	3.8091	0.0222	4563	10.2206	0.00004	4563	10.6479	0.00002	4563	5.59973	0.0037
VDAX --> FSI	5013	3.50828	0.03	4495	3.11038	0.0447	4495	1.91592	0.1473	4495	2.34411	0.096
VIX --> FSI	5081	5.97135	0.0026	4563	11.4408	0.00001	4563	9.52589	0.00007	4563	7.19877	0.0008

†† – indicates one-way Granger causality with 79 percent or better confidence. † – indicates consistent one-way Granger precedence

Table 18

Systemic stress episode identification via volatility benchmarks (discrimination threshold θ = 1std).

START	END	SSE	SYSTEMIC STRESS IDENTIFICATION	TYPE	VIX θ	MOVE θ	VDAX θ	BOX θ
8/19/1991	9/16/1991	1	Eastern European shocks	FOR	1.8	2.3		
9/30/1991	11/18/1991	2	Bond Markets Shock	FIN	1.5	2.2		
11/25/1991	1/6/1992	2	Scandinavian Crisis	FOR	4.5	1.8		
1/13/1992	6/1/1992	2	RE Credit Crunch	FIN	1.9	6.0		
7/6/1992	12/21/1992	3	ERM Crisis	FIN	1.5		1.4	
1/4/1993	4/12/1993	4	N/A	N/A	1.6	1.5	-	
4/26/1993	6/14/1993	5	N/A	N/A	2.2		1.5	
7/5/1993	9/20/1993	6	IMF Warning (Russian crisis/Global Bond Markets Reversal)	FIN	1.5		1.8	
9/27/1993	12/26/1994	6	Bond Market Crisis	FIN	2.3	2.3	2.5	
1/9/1995	1/30/1995	7	Bond Market Crisis	FIN		1.2		
3/6/1995	3/13/1995	8	Mexican Crisis	FIN		4.4		
4/3/1995	6/26/1995	9	Mexican Crisis	FIN	3.2	1.8	2.9	
7/3/1995	9/11/1995	9	Japanese Bank Runs / Withdrawal from U.S.D Assets	FIN	1.3	3.9	1.2	
9/18/1995	12/25/1995	9	Daiwa Bank Bond Trading Loss / U.S. Budget Congressional Standoff	FIN	2.4	1.4		
1/1/1996	7/1/1996	9	Inflation worries / Fed signals end of interest rate cuts	FIN	2.0	1.7		
7/15/1996	7/7/1997	10	"Irrational exuberance" volatility	FIN	4.1	2.2	3.8	2.2
7/14/1997	9/22/1997	10	Concurrent U.S. & German equity shocks	FIN	2.2		4.7	3.2
9/29/1997	12/8/1997	10	Asian Crisis	FIN	1.9	1.9	1.4	2.1
12/15/1997	3/2/1998	10	Asian Crisis	FIN	4.4	2.2	2.7	2.7
3/30/1998	7/13/1998	11	Asian Crisis	N/A	2.8	-	2.0	
7/20/1998	3/29/1999	11	LTCM Crisis	FIN	1.5			1.2
4/19/1999	4/19/1999	12	LTCM Crisis	N/A				1.3
5/3/1999	7/19/1999	13	Brazil Crisis	FIN	1.3	1.6		
7/26/1999	8/23/1999	13	International aftershocks (Latin America and Asia)	FIN	1.8	2.9	2.7	3.8
8/30/1999	11/8/1999	13	Y2K anxiety (Credit insurance shocks)	FIN	1.5	2.9	1.6	4.4
11/22/1999	3/27/2000	14	Dot-com Crisis	FIN	-		1.2	
4/10/2000	6/12/2000	15	Dot-com Crisis	FIN	5.5	3.4	1.8	
7/3/2000	2/19/2001	16	Dot-com Crisis	FIN		1.4		2.4
3/12/2001	8/27/2001	17	Early 2000s Recession	FOR	3.4	3.7	2.3	3.8
9/3/2001	2/18/2002	17	Terrorist attacks	NONFIN	3.1	2.0	4.6	4.0
3/11/2002	12/2/2002	18	Stock Market Downturn	FIN		1.7		
12/16/2002	3/3/2003	19	Stock Market Downturn	FIN		-	1.6	
3/10/2003	4/7/2003	19	Iraq War	NONFIN	2.4		2.3	1.4
6/23/2003	11/17/2003	20	Treasury Correction	FIN		-		
2/2/2004	2/2/2004	21	US Presidential Campaign Stress	NONFIN	1.8	2.5		-
3/15/2004	4/5/2004	22	Terrorist Attacks in Spain	NONFIN	3.7		2.1	
3/29/2004	6/21/2004	22	Interest Rate Shock	FIN	1.8	2.5	2.3	1.8
7/5/2004	8/30/2004	23	Energy Expectations Shock	FIN	1.9		1.2	
9/27/2004	11/1/2004	24	Fannie Mae Crisis (Capital Injection)	FIN		1.2		3.7
1/3/2005	1/24/2005	25	Contraction in Foreign Capital Flows / Republicans Win Election	FIN	2.1	2.2	-	
2/28/2005	8/15/2005	26	Combined Equity and Energy Price Shocks	FIN	-	2.7		
8/22/2005	8/29/2005	26	Hurricane Katrina / Energy Price Shock	NONFIN	2.2		3.4	
9/5/2005	11/14/2005	26	Equity Expectations Shock	FIN	1.5	2.1	2.9	3.2
11/21/2005	8/21/2006	26	Housing Correction Fears / Inflation Expectations	FIN			1.8	5.2
8/28/2006	2/19/2007	26	Amaranth Collapse	FIN	1.2		1.4	1.6
1/1/2007	7/16/2007	26	Subprime Industry collapse domino	FIN	3.3	3.4	2.2	
2/26/2007	4/2/2007	26	Chinese Correction	FIN	9.4	3.6	4.2	
7/23/2007	3/24/2008	26	Subprime Crisis	FIN	6.5	4.3	6.0	7.6
12/3/2007	4/21/2008	26	Late 2000s Recession	FIN	2.2	4.4		4.3
6/9/2008	3/30/2009	27	US Bear Market	FIN	1.4	1.9	1.4	2.1

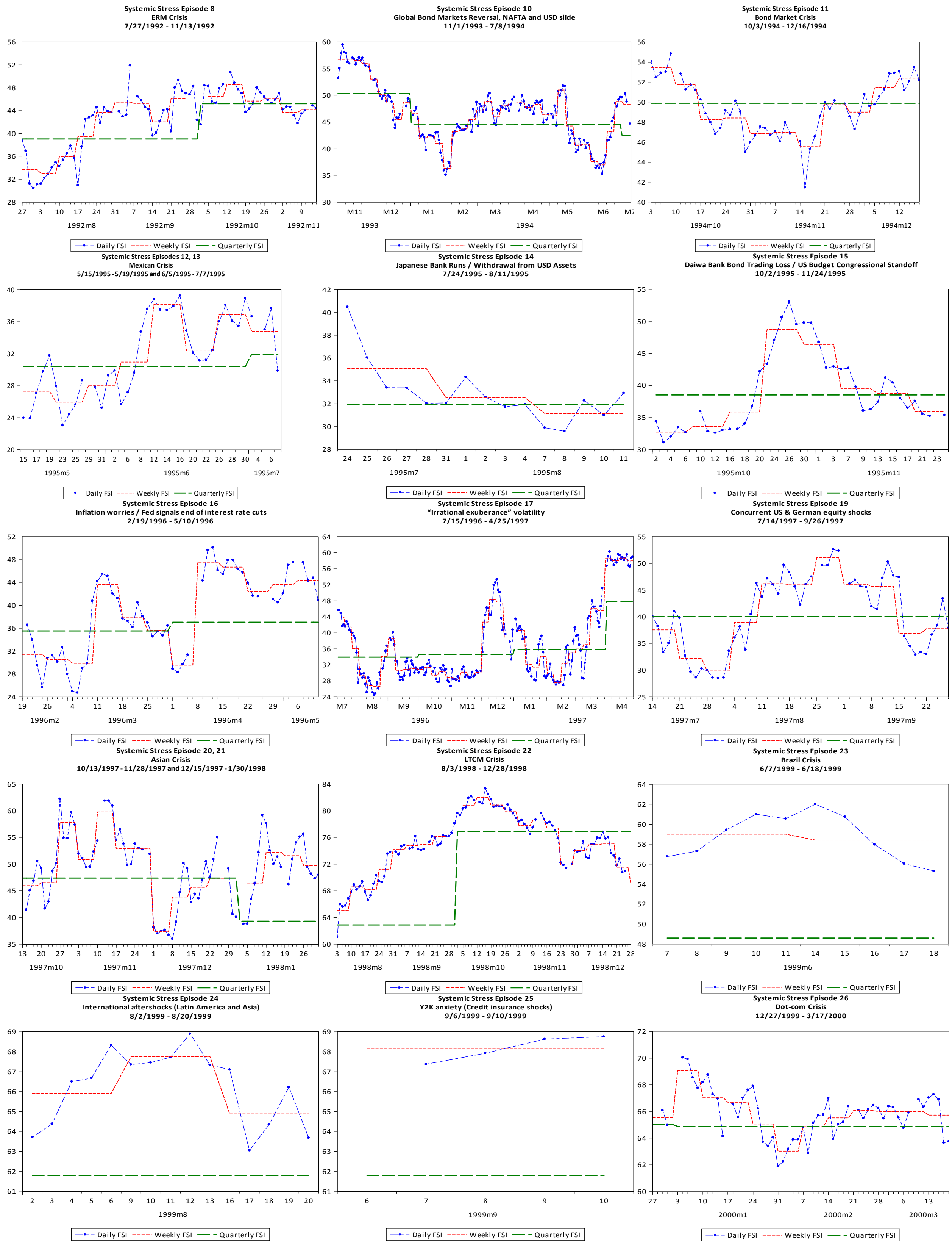


Fig. 18. Financial Systemic Stress Episodes: Quarterly Dependent Variable vs. High-Frequency FSI.

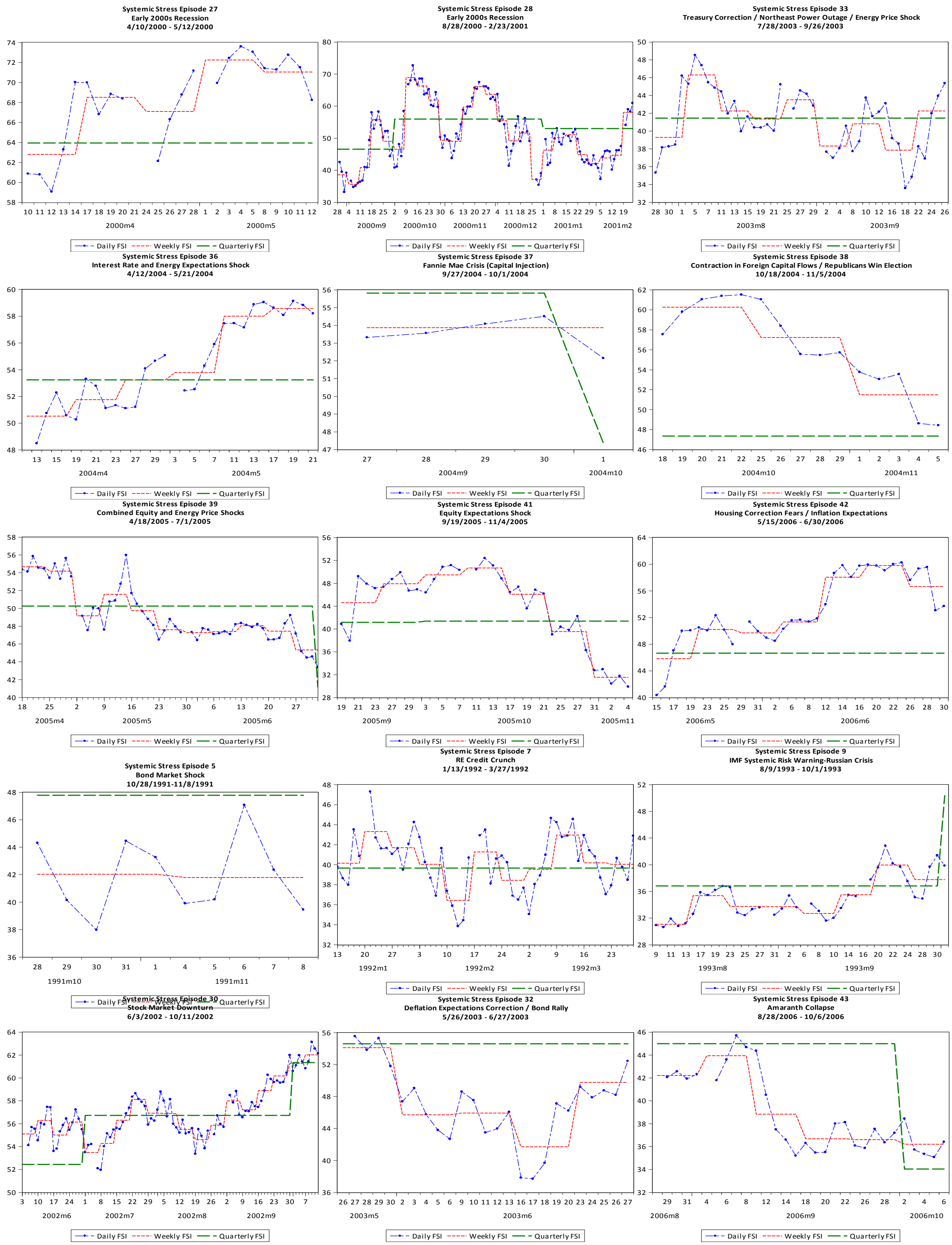


Fig. 18 (cont'd). Financial Systemic Stress Episodes: Quarterly Dependent Variable vs. High-Frequency FSI.

Fig. 18 (cont'd). Financial Systemic Stress Episodes: Quarterly Dependent Variable vs. High-Frequency FSI.

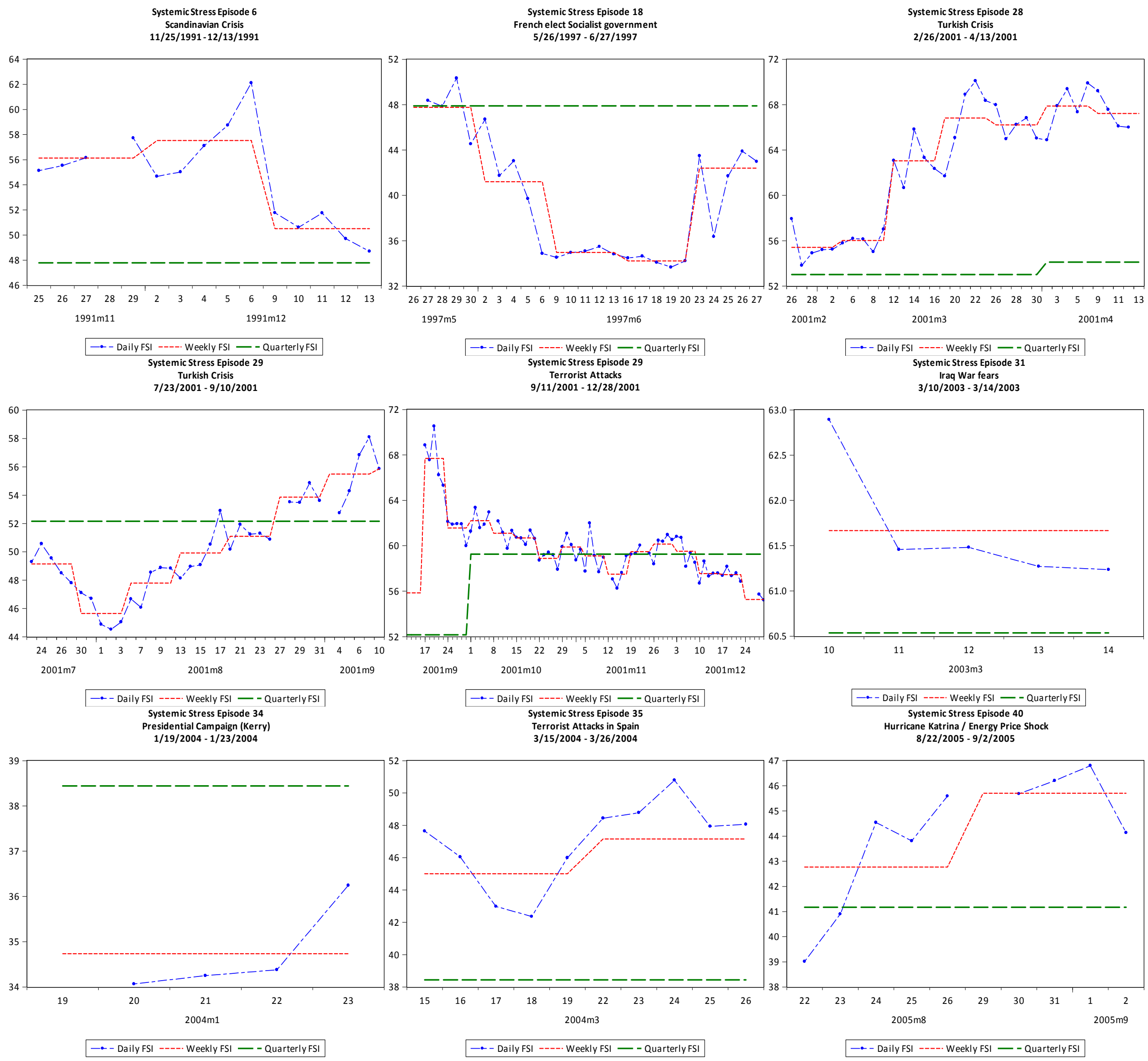


Fig. 19. Idiosyncratic Systemic Stress Episodes: Quarterly Dependent Variable vs. High-Frequency FSI.