Diagnosing the Financial System: Financial Conditions and Financial Stress*

Scott Brave[a] and R. Andrew Butters[b]
[a]Federal Reserve Bank of Chicago
[b]Northwestern University

We approach the task of monitoring financial stability within a framework that balances the costs and benefits of identifying future crisis-like conditions based on past U.S. financial crises. Our results indicate that the National Financial Conditions Index (NFCI) produced by the Federal Reserve Bank of Chicago is a highly predictive and robust indicator of financial stress at leading horizons of up to one year, with measures of leverage playing a crucial role in signaling financial imbalances. At longer forecast horizons, we propose an alternative sub-index of the NFCI that captures the relationship between non-financial leverage, financial stress, and economic activity.

JEL Codes: G01, G17, C43.

1. Introduction

Monitoring financial stability is not unlike a medical practitioner using a person's body temperature, blood pressure, and other vital signs to make a diagnosis of their health. Identifying the magnitude of financial stress at any given point in time can depend critically on

*This article is a revision of a previously published working paper (Scott Brave and R. Andrew Butters, 2010, "Gathering Insights on the Forest from the Trees: A New Metric for Financial Conditions," Federal Reserve Bank of Chicago Working Paper No. 2010-07). The views expressed herein are those of the authors and do not necessarily represent the views of the Federal Reserve Bank of Chicago or the Federal Reserve System. The authors thank the editor and two anonymous referees as well as Hesna Genay, Spencer Krane, Alejandro Justiniano, and Federal Reserve Bank of Chicago, 2011 Midwest Finance Association, and 2011 Computing in Economics and Finance seminar participants for their helpful comments and suggestions, and Max Lichtenstein for excellent research assistance.
Author e-mails: sbrave@frbchi.org and r-butters@kellogg.northwestern.edu.

the indicators examined and the reference points used. For instance, (i) what is a "normal" level of financial stress and subsequently a level that would warrant concern? (ii) are the risks associated with both extremely low values and high values the same? and (iii) how well does any one measure predict the true underlying state of the financial system?

What proves to be a "normal" level in medicine often tends to be a range rather than a particular value. For example, it is common to simply select some number of standard deviations from the population mean, or, equivalently, "build the range" by including everything that falls outside a desired percentage of the population. A possible source of bias in this kind of reasoning, however, might result if a priori we believe that some members of that population are in fact "sick" and their vital signs are skewed as a result. Including these members when calculating averages or percentiles will reduce the power of this metric to distinguish between states of the world, i.e., a "healthy state" and a "sick state."

In the context of financial stability, the "sick state" conforms with the notion of a financial crisis, where a number of financial indicators deviate substantially, and perhaps in a particular way, from their historical averages. If we envision every observation in time of an indicator as many different patients, some "sick" and some "healthy," and its value as their "vital sign," we can begin to build the intuition behind a statistical approach to monitoring financial stability. In fact, we can apply existing methods used in medical statistics, such as receiver operating characteristic (ROC) curve analysis, to approach this problem.

The ROC curve methodology is most commonly associated with the evaluation of the outcome of a medical test, which has at its base a Bayesian calculation of the following sort: Given a known incidence of a disease in a population, how likely is it that a positive test result is reflective of a true occurrence in sample? We take a similar approach, using known incidences of U.S. financial crises to inform our analysis of potential future stress on financial conditions. In this way, the past can be used as a guide to detecting future financial imbalances by evaluating the health of the financial system against relevant benchmarks from past financial crises.

The ROC methodology is rather flexible in that it can be used both to characterize the historical predictive ability of an indicator of financial conditions and to devise a rule based on that relationship with which to judge its future realizations. However, it leaves unanswered the question of how to appropriately weight potentially conflicting signals from a number of "vital signs." This is perhaps trivial when financial markets are operating smoothly, but it becomes more of an issue when markets become segmented. In this case, extracting a signal from a large number of indicators of financial conditions that reflects the systemic importance of each is likely to provide a more robust diagnosis.

This is the approach we take in applying the ROC methodology to a dynamic factor constructed from an unbalanced panel of 100 mixed-frequency indicators of financial activity. The resulting weekly index is the National Financial Conditions Index (NFCI) made publicly available by the Federal Reserve Bank of Chicago. Drawing on the work of Berge and Jordà (2011) in applying the ROC methodology to business-cycle dating, we then describe a statistical framework that balances the costs and benefits of identifying future crisis-like conditions based on the level of the index during past U.S. financial crises and explore the NFCI's robustness as both a contemporaneous and leading indicator of financial stress.

Our results indicate that the NFCI is 95 percent accurate in identifying historical crises contemporaneously, with a decline to 80 percent accuracy at a lead of up to one year. Furthermore, breaking down the index into subcomponents reflecting the themes of risk, credit, and leverage can enhance the nature of the signal provided by the NFCI as to the severity of the crisis, with leverage playing a crucial role in signaling financial imbalances. At horizons beyond one year, we show that a particular combination of household and non-financial business leverage measures in the NFCI proves to be a consistent leading indicator of financial stress and its impact on economic activity.

The rest of the paper proceeds as follows. In the next section, we describe the ROC methodology and its application to the NFCI and the risk, credit, and leverage sub-indexes. The following section then documents the robustness of the NFCI as a measure of financial stability. Finally, we compare its ability to capture financial stress relative to several alternatives. The concluding section summarizes our findings and discusses their potential consequences for policies aimed at promoting financial stability.

2. Monitoring Financial Stability

In this section, we outline a method of evaluating the state of the financial system based on a single composite indicator of financial conditions, the Federal Reserve Bank of Chicago's National Financial Conditions Index, or NFCI. The salient properties of the NFCI are first summarized. We then proceed to describe the ROC curve methodology that underlies our investigation of U.S. financial crises. Finally, we summarize the relevant features of the NFCI as a contemporaneous and leading indicator of financial stress.

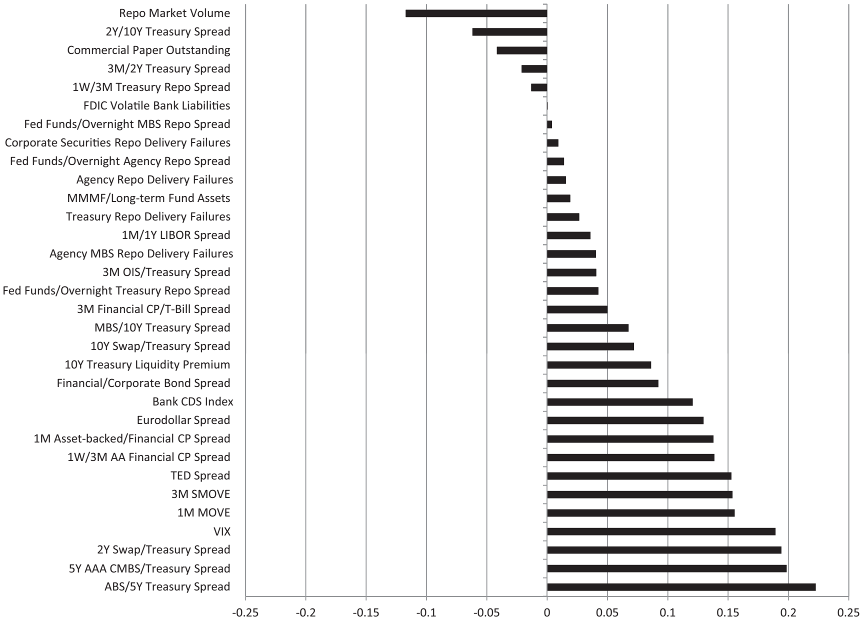
2.1 Features of the NFCI

Similar to other measures of financial conditions in the literature, such as the indexes of Illing and Liu (2006), Nelson and Perli (2007), Hakkio and Keeton (2009), Hatzius et al. (2010), and Matheson (2011), the NFCI is essentially a weighted average of a number of financial indicators where the weight given to each reflects the indicator's ability to explain the total variation among them. Indexes of this kind have the advantage of descriptively capturing the interconnectedness of their components. This attribute is desirable in the context of monitoring financial stability, as it means that more weight in the index is placed on financial indicators which have historically been systemically important.

Unlike the other financial conditions indexes in the literature, however, the NFCI uses a very flexible estimation procedure for these weights that builds off the work of Stock and Watson (2002), Doz, Giannone, and Reichlin (2006), and Aruoba, Diebold, and Scotti (2009) and is described in detail in the appendix. By allowing for the inclusion of financial indicators of varying reported frequencies that start and end at different times, it can produce a high-frequency index of financial conditions with minimal restrictions. Furthermore, the estimation strategy makes use of both cross-sectional and time-series information, exploiting the historical cross-correlations and dynamic properties of the indicators to determine their weight in the index. In this fashion, the NFCI is able to consistently extract the common signal in 100 indicators of financial activity on a weekly basis since 1973.

Volatility and credit risk measures tend to receive positive weights in the NFCI, while measures of credit and leverage tend

Figure 1. Ranking of NFCI Indicators by Factor Loadings: Risk Indicators



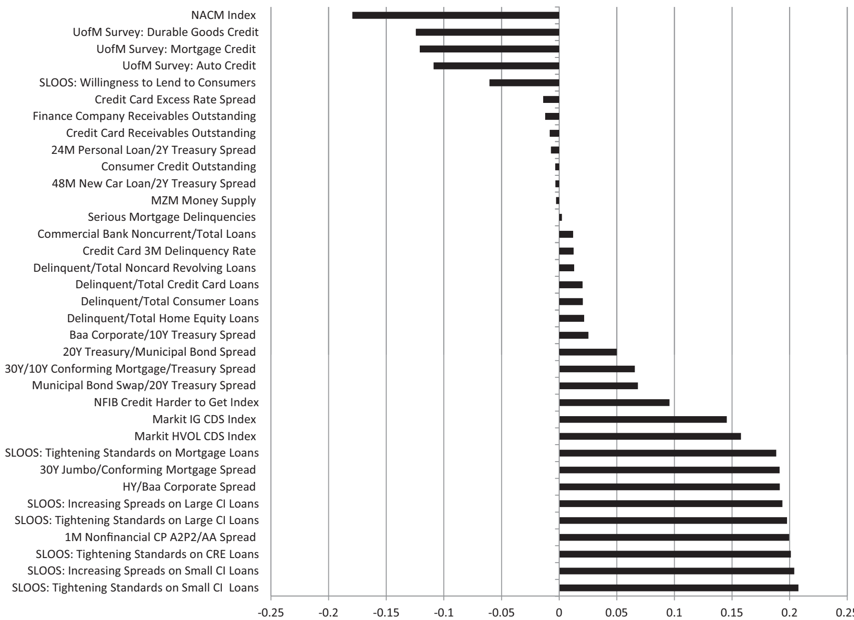
to be negatively correlated with the level of the index. This pattern of increasing volatility and risk premia and declining liquidity and leverage is consistent with tightening financial conditions and provides the basis for the NFCI's interpretation. Brave and Butters (2011) provide a detailed examination of the individual indicators in the NFCI. We summarize their findings in figures 1–3, which depict the weights, or factor loadings, for all 100 indicators classified into three types: risk, credit, and leverage.[1]

2.2 Receiver Operating Characteristics (ROC) Curve Analysis

We begin by constructing prior information on the incidence of U.S. financial crises. Ideally, we would like to have a professional

[1] Additional information on the NFCI, including data sources and complete data descriptions, can be found at www.chicagofed.org/nfci.

Figure 2. Ranking of NFCI Indicators by Factor Loadings: Credit Indicators

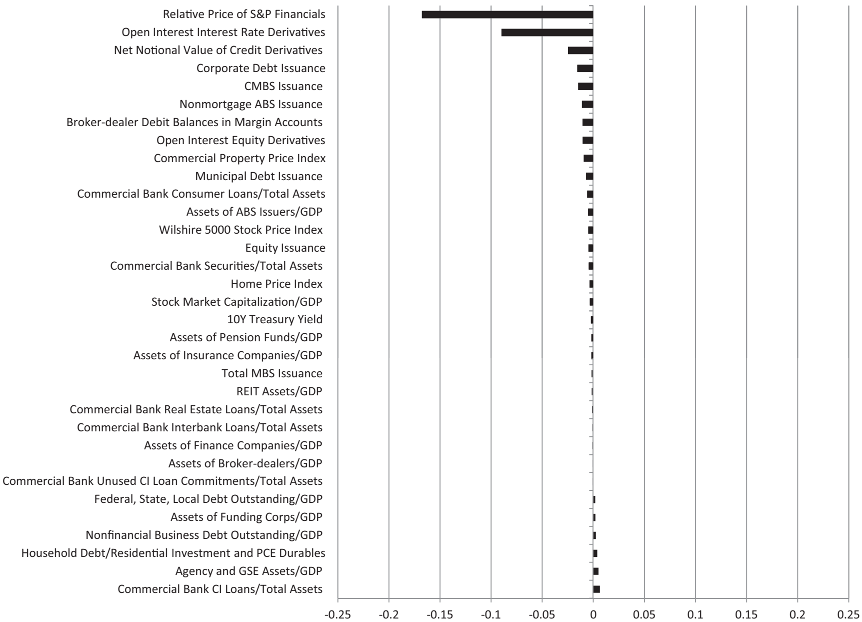


consensus much like the business-cycle dates that the National Bureau of Economic Research produces for U.S. recessions and expansions. Unfortunately, what we have to resort to are the historical accounts of events in U.S. financial history for which there is not always consensus. With this information in hand, we then follow the approach used by Berge and Jordà (2011) in evaluating indicators of the business cycle. By placing relative weights on the utility from correctly predicting crisis and non-crisis states and the disutility from making a false positive versus false negative evaluation, we develop a threshold decision rule for the transition into and out of crisis-like financial conditions.

2.2.1 Financial Crisis Dating

Table 1 provides a list of five financial crises in U.S. financial history over the last forty years, along with some of the major events that occurred during each. To arrive at these crisis episodes, we conducted

Figure 3. Ranking of NFCI Indicators by Factor Loadings: Leverage Indicators



a survey of the literature on U.S. financial crises.[2] Each decade contains at least one crisis episode, with the shortest episode extending for only two years and the longest seven years in length. The earlier episodes tend to be more concentrated in the commercial banking system, while the latter begin to take on a larger capital markets element. In determining these episodes, we paid close attention to similarities in the historical accounts on the triggers of the crisis and the policy responses that followed.

Laeven and Valencia (2008, 2010), Reinhart and Rogoff (2009), and López-Salido and Nelson (2010) go through similar exercises in producing crisis dates, albeit at a lower frequency.[3] While their start and end dates often differ significantly from those in table 1,

[2]Some examples include FDIC (1984, 1997), Schreft (1990), Spero (1999), Laeven and Valencia (2008), and Reinhart and Rogoff (2009).

[3]For the most recent crisis, the alternative dating conventions mostly conform with table 1.

Table 1. U.S. Financial Crises Since 1973

International Banking Crisis (1973–75)	1973w2	Jan. 11, 1973	DJIA peaks above 1000 and then begins to sharply decline.
	1973w19	May 11–13, 1973	Bilderberg meeting to discuss developing pressures in "petrodollars."
	1973w42	Oct. 17–18, 1974	U.S. National Bank of San Diego declared insolvent, first billion-dollar bank failure; Arab oil embargo begins.
	1974w11	Mar. 17, 1974	Oil embargo is lifted, but money-center banks and REITS continue to experience problems.
	1974w19	May 10, 1974	Regulatory agencies step in with financial assistance for Franklin National Bank.
	1974w41	Oct. 9, 1974	Franklin National Bank collapses and is acquired by European American Bank.
	1974w49	Dec. 6, 1974	DJIA bottoms out at 45% decline.
	1975w4	Jan. 19, 1975	Assisted merger of Security National Bank of New York with Chemical National Bank by regulatory agencies.
	1975w21	May 23, 1975	Regulatory agencies assist Bank of the Commonwealth to keep it afloat.
Dollar, Banking, and LDC Crises (1977–84)	1977w40	Oct. 5, 1977	Dollar peaks and then begins steep decline against major foreign currencies.
	1978w44	Nov. 1, 1978	Carter administration announcement of a dollar defense program.
	1978w51	Dec. 17, 1978	OPEC decides to keep its U.S. dollar reserves but increase oil prices in 1979.
	1980w11	Mar. 14, 1980	Carter administration announcement of imposition of credit controls.
	1980w13	Mar. 26, 1980	Regulatory agencies step in with financial assistance for First Pennsylvania National Bank.

(continued)

Table 1. (Continued)

	1980w27	July 3, 1980	Federal Reserve announces phase-out of credit controls.
	1981w43	Nov. 4, 1981	FDIC assists merger of Greenwich Savings Bank, first in a series of mutual savings bank assisted mergers.
	1982w26	July 5, 1982	Penn Square Bank fails.
	1982w31	Aug. 12, 1982	Mexico defaults on its debt, beginning of LDC crisis.
	1982w44	Nov. 6, 1982	Mexico and IMF reach accord on loan plan.
	1984w19	May 9, 1984	Run on Continental Illinois begins, bank borrows $3.6 billion through discount window.
	1984w27	July 1, 1984	Regulators develop plan to take over Continental's bad loans.
	1984w39	Sept. 26, 1984	Resolution of Continental completed.
S&L Crisis, Black Monday, and LBO/Junk Bond Collapse (1987–91)	1986w53	Jan. 1, 1987	Federal Savings and Loan Insurance Corporation becomes insolvent.
	1987w41	Oct. 19, 1987	"Black Monday": DJIA −22.6%, S&P500 −20.4%.
	1989w32	Aug. 9, 1989	Financial Institutions Reform Recovery and Enforcement Act (FIRREA) signed into law.
	1989w41	Oct. 13, 1989	"Friday the 13th Mini Crash" helps to trigger junk bond market collapse.
	1990w3	Jan. 15, 1990	Campeau Corporation files for bankruptcy after junk bond defaults.
	1990w7	Feb. 13, 1990	Drexel Burnham Lambert files for bankruptcy, beginning of end to LBO boom.
	1990w31	Aug. 2, 1990	Iraqi invasion of Kuwait, DJIA declines 18% in three months.
	1990w52	Dec. 31, 1990	First phase of implementation of Basel I regulatory capital and leverage ratio requirements.
	1991w9	Feb. 28, 1991	Gulf War ends.

(continued)

Table 1. (Continued)

Asian Crisis, Russian Default and LTCM, Y2K, NASDAQ Bubble, 9/11, and Enron (1997–2002)	1997w43	Oct. 27, 1997	"Mini crash" brought on by Asian financial crisis.
	1998w32	Aug. 17, 1998	Russia defaults on its debt.
	1998w37	Sept. 23, 1998	Collapse of LTCM (Federal Reserve steps in with support).
	1999w29	July 27, 1999	IMF approves stand-by credit for Russian Federation.
	1999w39	Oct. 1, 1999	Fed establishes Century Date Change Special Liquidity Facility.
	1999w52	Jan. 1, 2000	Y2K passes.
	2000w35	Sept. 1, 2000	NASDAQ peaks above 4000, then begins to sharply decline.
	2001w37	Sept. 11, 2001	Terrorist attack on the World Trade Center.
	2001w50	Dec. 13, 2001	SEC enforcement action against Enron.
	2002w28	July 15, 2002	Arthur Anderson indicted.
	2002w31	July 30, 2002	Sarbanes-Oxley Act passed.
Subprime Mortgage Crisis and Aftermath (2007–Current)	2007w31	July 31, 2007	Bear Stearns liquidates two hedge funds investing in MBS.
	2008w11	Mar. 14, 2008	Bear Stearns sold to JP Morgan Chase with NY Fed support.
	2008w28	July 12–15, 2008	Fannie Mae and Freddie Mac receive government assistance, IndyMac fails.
	2008w37	Sept. 14–16, 2008	Lehman Brothers files for bankruptcy, AIG receives govt. support, Reserve Fund "breaks the buck."
	2008w39	Sept. 26, 2008	Washington Mutual Bank failure, largest failure in terms of assets to date.
	2008w40	Oct. 3, 2008	Emergency Economic Stabilization Act passed (TARP).
	2008w47	Nov. 23, 2008	Citigroup requires government assistance.
	2009w2	Jan. 16, 2009	Bank of America requires government assistance.
	2010w18	May 9, 2010	EU, ECB, and IMF announce $1 trillion aid package after Greek debt crisis.
	2011w32	Aug. 5, 2011	S&P downgrade of U.S. credit rating.

their crisis episodes are largely subsets of those that we consider, with the exception of the 1997–2002 period, which none of the others deems as evidence of crisis-like financial conditions. Our episodes share with those of López-Salido and Nelson (2010) the tumultuous events in financial markets that accompanied the 1974–75, 1981–82, and 2007–09 U.S. recessions, as well as the savings and loan crisis of the mid-1980s and early 1990s. The latter two crises are the singular focus of Laeven and Valencia (2008, 2010) and Reinhart and Rogoff (2009), as each is concerned primarily with periods characterized by substantial bank failures. We focus on the dating convention of Laeven and Valencia (2010) for the U.S. crises, as it differs the most from table 1 and López-Salido and Nelson (2010).

2.2.2 The Area Under the ROC Curve

The ROC method requires that we categorize each point in time as falling within a crisis or non-crisis period. Given the dating conventions in table 1, consider the derivation in Berge and Jordà (2011):

$$TP(c) = P[I_t \geq c || C_t = 1] \tag{1}$$

$$FP(c) = P[I_t \geq c || C_t = 0], \tag{2}$$

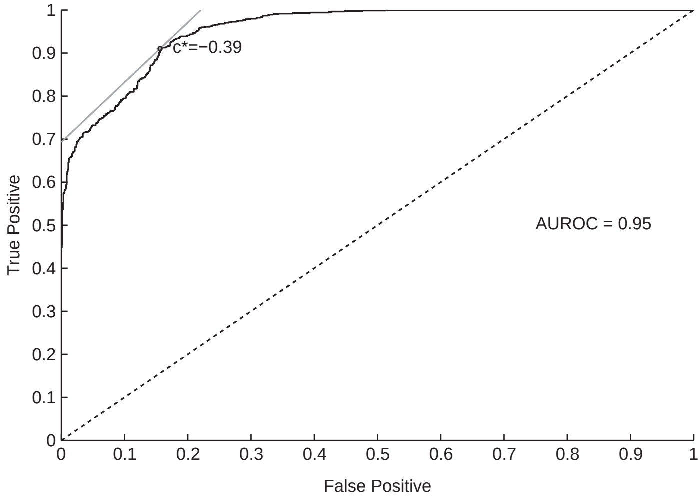
with $C_t \in \{0, 1\}$ indicating the non-crisis and crisis states of financial conditions, respectively. $TP(c)$ is typically referred to as the true positive, sensitivity, or recall rate, and $FP(c)$ is known as the false positive or 1-specificity rate for an indicator, I_t, observed with value c. The relationship between the two is described by the ROC curve. With the Cartesian convention, this curve is given by

$$\{ROC(r), r\}_{r=0}^{1}, \tag{3}$$

where $ROC(r) = TP(c)$ and $r = FP(c)$.

Figure 4 depicts this curve estimated non-parametrically for the weekly National Financial Conditions Index produced by the Federal Reserve Bank of Chicago. The closer to one or zero the area under the curve (AUROC)—depending on whether the index is positively or negatively correlated with the crisis episodes in table 1, respectively—the more predictive it is of crisis-like financial conditions, where statistical significance is judged relative to the area

Figure 4. NFCI ROC Curve with Utility Line



under the ray from the origin extending at a 45-degree angle.[4] The NFCI is highly predictive of the crisis episodes in table 1, with an AUROC of roughly 0.95 that is statistically significant at standard levels of confidence. Repeating this procedure with a lead of one and two years between the values of the NFCI and the crisis episodes produces AUROCs that remain statistically significant at 0.80 and 0.62, respectively.

Of the three crisis dating conventions, our baseline and the López-Salido and Nelson (2010) dates produce the most similar AUROC results for the NFCI, particularly at the one-year-ahead horizon. What is mostly responsible for the lower contemporaneous AUROC values of 0.76 and 0.62 when using the López-Salido and Nelson (2010) dates or those of Laeven and Valencia (2010) is the 1997–2002 period not being considered by either as a crisis and the omission of the early '70s and '80s by the latter. Still, for both alternative crisis dating conventions, the AUROC values for the NFCI remain statistically significant, and they remain so even at the three-year-ahead horizon as opposed to our baseline dates, which tend to

[4]The procedure for evaluating statistical significance is described in DeLong, DeLong, and Clarke-Pearson (1988).

already be more liberal in their definition of the beginning and ends of crises.

For the sake of comparison, consider table 2, which depicts the AUROC value for each of the indicators in the NFCI contemporaneously and at leads of up to three years.[5] Values in darkened and boldface fonts denote statistical significance at standard confidence levels based on the crisis episodes in table 1. The table categorizes each indicator into one of three types: risk, credit, or leverage. Risk indicators capture volatility and funding risk in the financial sector, while credit indicators are composed of measures of credit conditions, and leverage indicators consist of debt and equity measures.

In general, risk and credit measures tend to have very significant contemporaneous and one-year-ahead AUROC values. However, the AUROCs for risk measures at longer forecast horizons tend to be more significant than for credit measures, although a few of the latter remain significant predictors out to three years ahead. In contrast, there are considerably fewer significant leverage indicators at a contemporaneous forecast horizon, although several do appear to have a leading feature from two to three years ahead. Only a handful of indicators display a sense of being a superior contemporaneous or leading indicator to the NFCI, and many of these are neither observed weekly nor do they span all five crisis episodes in table 1, providing our motivation for the NFCI as the relevant measure for monitoring financial stability.

It is also possible to construct sub-indexes of the NFCI that reflect the different classes of indicators in table 2. Table 3 displays the AUROC for the NFCI and its three sub-indexes using our baseline and alternative crisis dates contemporaneously and up to a lead of three years. The predictive ability of the overall index is superior to any of the three sub-indexes from a year ahead to the contemporaneous horizon, although of the three sub-indexes, risk measures dominate. Two years and beyond, measures of leverage increasingly account for the ability of the NFCI to predict crisis conditions. In fact, relative to the Laeven and Valencia (2010) dates, the leverage

[5]One indicator in the NFCI, the net notional value of credit derivatives, did not have a sufficient history for the AUROC calculations.

Table 2. AUROC for Financial Indicators at Various Horizons

Indicator	Type	Freq.	# of Crises	Current	−1 Year	−2 Years	−3 Years
Baa Corporate/10Y Treasury Spread	Credit	W	5	**0.60**	0.50	**0.45**	**0.38**
1M Non-Financial CP A2P2/AA Spread	Credit	W	2	**0.82**	**0.64**	0.53	**0.40**
20Y Treasury/Municipal Bond Spread	Credit	W	5	**0.58**	**0.55**	**0.54**	**0.56**
24M Personal Loan/2Y Treasury Spread	Credit	Q	4	**0.38**	**0.35**	**0.29**	**0.32**
30Y Jumbo/Conforming Mortgage Spread	Credit	W	2	**0.90**	**0.88**	**0.70**	**0.57**
30Y/10Y Conforming Mort./Treasury Spread	Credit	W	2	**0.68**	**0.67**	**0.56**	**0.44**
48M New Car Loan/2Y Treasury Spread	Credit	Q	4	0.44	0.44	**0.33**	**0.34**
Commercial Bank Non-Current/Total Loans	Credit	Q	4	**0.80**	**0.76**	0.61	0.50
Consumer Credit Outstanding	Credit	M	5	0.45	0.50	0.54	**0.60**
Credit Card 3M Delinquency Rate	Credit	M	2	0.50	0.55	0.56	0.57
Credit Card Excess Rate Spread	Credit	M	2	**0.63**	0.53	0.49	0.52
Credit Card Receivables Outstanding	Credit	M	2	**0.42**	0.46	0.50	0.52
Delinquent/Total Consumer Loans	Credit	M	2	0.55	**0.64**	0.60	0.51
Delinquent/Total Credit Card Loans	Credit	M	2	0.50	0.56	0.54	0.50
Delinquent/Total Home Equity Loans	Credit	M	2	0.50	0.61	**0.70**	**0.68**
Delinquent/Total Non-Card Revolving Loans	Credit	M	2	0.51	**0.63**	0.60	0.50

(continued)

Table 2. (Continued)

Indicator	Type	Freq.	# of Crises	Current	−1 Year	−2 Years	−3 Years
Finance Company Receivables Outstanding	Credit	M	3	0.45	0.51	0.54	0.55
HY/Baa Corporate Spread	Credit	W	2	**0.79**	**0.56**	**0.39**	**0.17**
Markit HVOL CDS Index	Credit	W	1	**1.00**	**0.80**	**0.68**	NaN
Markit IG CDS Index	Credit	W	1	**0.99**	**0.77**	**0.60**	NaN
Municipal Bond Swap/20Y Treasury Spread	Credit	W	3	**0.63**	**0.77**	**0.79**	**0.78**
MZM Money Supply	Credit	M	5	0.48	**0.56**	0.54	0.52
NACM Index	Credit	M	2	**0.14**	**0.16**	**0.30**	0.49
NFIB Credit Harder to Get Index	Credit	M	5	**0.65**	0.53	**0.37**	**0.40**
Serious Mortgage Delinquencies	Credit	Q	5	0.56	**0.62**	0.56	0.54
SLOOS: Increasing Spreads on Large CI Loans	Credit	Q	3	**0.83**	0.71	0.54	0.28
SLOOS: Increasing Spreads on Small CI Loans	Credit	Q	3	**0.81**	0.65	0.47	**0.23**
SLOOS: Tightening Standards on CRE Loans	Credit	Q	3	**0.81**	0.71	0.60	0.39
SLOOS: Tightening Standards on Large CI Loans	Credit	Q	3	**0.83**	**0.71**	0.56	0.34
SLOOS: Tightening Standards on Mortgage Loans	Credit	Q	3	0.69	0.68	0.59	0.53
SLOOS: Tightening Standards on Small CI Loans	Credit	Q	3	**0.84**	**0.74**	0.59	0.37

(continued)

Table 2. (Continued)

Indicator	Type	Freq.	# of Crises	Current	−1 Year	−2 Years	−3 Years
SLOOS: Willingness to Lend to Consumers	Credit	Q	5	**0.31**	**0.34**	**0.38**	0.50
UofM Survey: Auto Credit	Credit	M	4	**0.24**	**0.30**	**0.33**	**0.41**
UofM Survey: Durable Goods Credit	Credit	M	4	**0.17**	**0.27**	**0.36**	0.47
UofM Survey: Mortgage Credit	Credit	M	4	**0.22**	**0.29**	**0.34**	**0.40**
10Y Treasury Yield	Leverage	W	5	**0.54**	0.52	0.52	0.51
Agency and GSE Assets/GDP	Leverage	Q	5	0.60	0.57	0.54	0.47
Assets of ABS Issuers/GDP	Leverage	Q	4	**0.29**	**0.30**	0.39	0.60
Assets of Broker-Dealers/GDP	Leverage	Q	5	**0.37**	0.45	0.48	0.52
Assets of Finance Companies/GDP	Leverage	Q	5	0.49	0.59	0.56	0.50
Assets of Funding Corps/GDP	Leverage	Q	5	0.54	**0.62**	**0.63**	0.52
Assets of Insurance Companies/GDP	Leverage	Q	5	0.42	0.51	0.50	0.52
Assets of Pension Funds/GDP	Leverage	Q	5	0.43	0.48	0.48	0.56
Broker-Dealer Debit Balances in Margin Accounts	Leverage	M	5	**0.37**	0.49	0.54	**0.59**
CMBS Issuance	Leverage	M	3	**0.42**	0.44	0.44	0.49
Commercial Bank CI Loans/Total Assets	Leverage	M	5	**0.56**	**0.59**	**0.65**	**0.66**
Commercial Bank Consumer Loans/Total Assets	Leverage	M	5	**0.43**	**0.42**	**0.43**	0.47
Commercial Bank Interbank Loans/Total Assets	Leverage	M	5	0.50	0.51	0.52	0.51
Commercial Bank Real Estate Loans/Total Assets	Leverage	M	5	0.49	0.49	0.49	0.47

(continued)

Table 2. (Continued)

Indicator	Type	Freq.	# of Crises	Current	−1 Year	−2 Years	−3 Years
Commercial Bank Securities/Total Assets	Leverage	M	5	0.47	**0.44**	**0.39**	**0.39**
Commercial Bank Unused CI Commitments/Total Assets	Leverage	Q	4	0.44	0.40	0.39	0.40
Commercial Property Price Index	Leverage	Q	4	0.37	0.48	0.56	0.60
Corporate Debt Issuance	Leverage	M	3	**0.39**	**0.41**	0.44	0.51
Equity Issuance	Leverage	M	3	0.44	0.46	0.54	0.52
Federal, State, Local Debt Outstanding/GDP	Leverage	Q	5	0.45	**0.38**	0.41	0.46
Home Price Index	Leverage	M	4	0.51	**0.56**	**0.56**	**0.57**
Household Debt/Residential Investment and PCE Durables	Leverage	Q	5	0.59	0.57	0.55	0.52
Municipal Debt Issuance	Leverage	M	1	0.40	0.48	0.43	0.47
Non-Financial Business Debt Outstanding/GDP	Leverage	Q	5	**0.62**	**0.70**	**0.71**	**0.67**
Non-Mortgage ABS Issuance	Leverage	M	2	0.41	0.44	0.45	0.48
Open Interest Equity Derivatives	Leverage	W	2	0.46	**0.43**	**0.42**	0.47
Open Interest Interest Rate Derivatives	Leverage	W	2	**0.35**	**0.32**	**0.39**	0.46
REIT Assets/GDP	Leverage	Q	5	0.43	0.44	0.40	0.46
Relative Price of S&P Financials	Leverage	W	3	**0.27**	**0.33**	**0.31**	**0.36**

(continued)

Table 2. (Continued)

Indicator	Type	Freq.	# of Crises	Current	−1 Year	−2 Years	−3 Years
Stock Market Capitalization/GDP	Leverage	Q	5	0.40	0.50	0.53	0.59
Total MBS Issuance	Leverage	M	2	0.49	0.40	**0.35**	**0.29**
Wilshire 5000 Stock Price Index	Leverage	M	5	0.47	0.51	0.54	**0.58**
MBS/10Y Treasury Spread	Risk	M	4	**0.75**	**0.79**	**0.69**	0.56
10Y Swap/Treasury Spread	Risk	W	3	**0.75**	**0.69**	**0.57**	**0.41**
10Y Treasury Liquidity Premium	Risk	W	3	**0.61**	**0.58**	0.47	**0.34**
1M Asset-Backed/Financial CP Spread	Risk	W	2	**0.91**	**0.81**	**0.73**	**0.61**
1M MOVE	Risk	W	3	**0.62**	0.49	**0.46**	**0.36**
1M/1Y LIBOR Spread	Risk	W	3	0.51	**0.43**	0.49	**0.43**
1W/3M AA Financial CP Spread	Risk	W	2	0.51	0.46	**0.65**	**0.65**
1W/3M Treasury Repo Spread	Risk	W	2	**0.39**	**0.45**	**0.55**	**0.57**
2Y Swap/Treasury Spread	Risk	W	3	**0.85**	**0.81**	**0.67**	0.48
2Y/10Y Treasury Spread	Risk	W	5	**0.34**	**0.28**	**0.33**	**0.37**
3M Financial CP/T-Bill Spread	Risk	W	5	**0.65**	**0.69**	**0.70**	**0.58**
3M OIS/Treasury Spread	Risk	W	1	**0.31**	0.48	**0.73**	**0.94**
3M SMOVE	Risk	W	2	**0.46**	**0.36**	**0.31**	**0.20**
3M/2Y Treasury Spread	Risk	W	5	**0.38**	**0.34**	**0.46**	**0.46**
5Y AAA CMBS/Treasury Spread	Risk	W	2	**0.93**	**0.71**	0.51	**0.28**
ABS/5Y Treasury Spread	Risk	M	3	**0.88**	**0.88**	**0.82**	**0.65**
Agency MBS Repo Delivery Failures	Risk	W	2	**0.55**	**0.55**	0.51	**0.44**
Agency Repo Delivery Failures	Risk	W	2	**0.54**	**0.55**	0.49	**0.44**
Bank CDS Index	Risk	W	1	**1.00**	**0.80**	0.57	NaN

(continued)

Table 2. (Continued)

Indicator	Type	Freq.	# of Crises	Current	−1 Year	−2 Years	−3 Years
Commercial Paper Outstanding	Risk	W	2	**0.45**	0.54	**0.60**	**0.59**
Corporate Securities Repo Delivery Failures	Risk	W	2	0.48	0.54	0.48	0.45
Eurodollar Spread	Risk	W	5	**0.82**	**0.79**	**0.71**	**0.62**
FDIC Volatile Bank Liabilities	Risk	Q	4	0.45	0.54	0.62	0.63
Fed. Funds/Overnight Agency Repo Spread	Risk	W	2	0.51	0.53	0.51	**0.45**
Fed. Funds/Overnight MBS Repo Spread	Risk	W	2	0.52	0.53	0.48	**0.45**
Fed. Funds/Overnight Treasury Repo Spread	Risk	W	2	0.51	**0.55**	**0.54**	0.51
Financial/Corporate Bond Spread	Risk	M	4	**0.81**	**0.84**	**0.79**	**0.67**
MMMF/Long-Term Fund Assets	Risk	M	5	**0.74**	**0.66**	**0.63**	0.55
Repo Market Volume	Risk	W	2	**0.42**	**0.45**	**0.39**	**0.41**
TED Spread	Risk	W	4	**0.75**	**0.79**	**0.75**	**0.62**
Treasury Repo Delivery Failures	Risk	W	2	0.51	0.53	0.47	0.50
VIX	Risk	W	3	**0.91**	**0.79**	**0.62**	**0.39**

Notes: The table displays the area under the ROC curve relative to the five crisis periods in table 1 contemporaneously and at leads of up to three years for each financial indicator in the NFCI. Areas significant from 50 percent are denoted by bold (95 percent significance level) and darkened fonts (90 percent significance level).

Table 3. AUROC at Contemporaneous and Leading Horizons

Indicator	Baseline				López-Salido and Nelson (2010)				Laeven and Valencia (2010)			
	0 Yr	**−1 Yr**	**−2 Yr**	**−3 Yr**	**0 Yr**	**−1 Yr**	**−2 Yr**	**−3 Yr**	**0 Yr**	**−1 Yr**	**−2 Yr**	**−3 Yr**
NFCI	**0.95**	**0.80**	**0.62**	0.48	**0.76**	**0.79**	**0.73**	**0.59**	**0.62**	**0.56**	**0.45**	**0.32**
Risk	**0.93**	0.80	**0.64**	0.50	**0.73**	**0.77**	**0.74**	**0.60**	**0.58**	**0.53**	**0.45**	**0.32**
Credit	**0.90**	**0.73**	**0.55**	**0.45**	**0.72**	**0.73**	**0.62**	**0.55**	**0.63**	**0.55**	**0.43**	**0.39**
Leverage	**0.78**	**0.72**	**0.66**	0.53	**0.69**	**0.79**	**0.79**	**0.71**	**0.77**	**0.74**	**0.68**	**0.55**

Notes: The table displays the area under the ROC curve relative to the crisis periods denoted in the column headings contemporaneously and at leads of up to three years for the NFCI and its three sub-indexes. Areas significant from 50 percent are denoted by bold (95 percent significance level) and darkened fonts (90 percent significance level).

sub-index is the dominant signal from the contemporaneous to the two-year-ahead horizon. We explore this result further below.

2.2.3 Crisis Thresholds

The utility function depicted in figure 4 is expressed as in Baker and Kramer (2007),

$$\begin{aligned} U = {} & U_{11} ROC(r)\pi + U_{01}(1 - ROC(r))\pi + U_{10} r(1-\pi) \\ & + U_{00}(1-r)(1-\pi), \end{aligned} \tag{4}$$

where U_{ij} is the utility (or disutility) associated with the prediction i, given that the true state is j, $i, j \in \{0, 1\}$ and π is the unconditional probability of observing a crisis episode during the sample period. Utility maximization implies that the optimal threshold value c^* is given by the solution to

$$\frac{\partial ROC}{\partial r} = \frac{U_{00} - U_{10}}{U_{11} - U_{01}} \frac{1-\pi}{\pi}, \tag{5}$$

that is, the point where the slope of the ROC curve equals the expected marginal rate of substitution between the net utility of accurate crisis and non-crisis episode prediction.

Essentially, one is weighing the costs of a type I versus type II error relative to the benefits of correctly predicting the true state when attempting to separate a mixture distribution into its unique components. This intuitively amounts to deciding on the emphasis (in utility terms) one wants to put on correctly identifying either state. An example of assigning equal weight to correctly identifying both crisis and non-crisis episodes would be assigning $U_{00} = U_{11} = 1$, $U_{01} = U_{10} = -1$. In contrast, placing all the emphasis on correctly identifying financial crises, and subsequently no emphasis on the likely error of identifying the other state as a crisis, the utilities could be assigned this way: $U_{00} = 0$, $U_{11} = 1$, $U_{01} = -1$, $U_{10} = -\epsilon$, where ϵ needs to be small but non-zero in order to prevent the utility function from being degenerate. Finally, a threshold rule that puts more emphasis on identifying non-crisis periods corresponds with a utility function like $U_{00} = 1$, $U_{11} = 0$, $U_{01} = -\epsilon$, $U_{10} = -1$.

These alternative approaches are consistent with three parameterizations of the level sets of the utility function. Graphically, each rule attempts to find the unique intersection of the linear utility function with the convex ROC curve. A rule placing a very steep penalty on missing early on an occurrence of a financial crisis thus looks to intersect the upward-sloping portion of the ROC curve. A rule that places a relatively larger penalty on missing an occurrence of a non-crisis period does the opposite and instead intersects the flatter portion of the ROC curve. The equal weight, or "unbiased," rule falls somewhere in between the other two on the ROC curve.

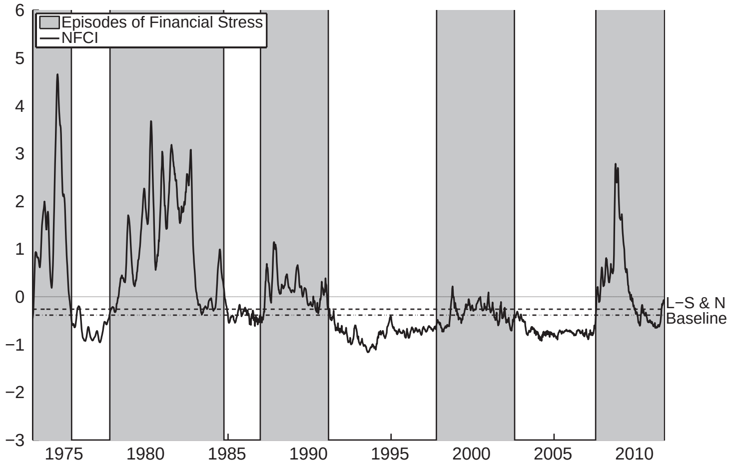
2.3 Making a Diagnosis

With equal weight on all four utilities (disutilities), the unbiased threshold rule balances the need to detect financial imbalances in advance of a crisis with the desire to avoid assigning movements in the NFCI at low levels to financial stress versus simple mean reversion. In essence, this derivation "rebases" the index based on the historical financial crises we have identified. However, as in any Bayesian analysis, it is necessary to investigate the sensitivity of the decision rule to the prior information used. For the NFCI, the unbiased threshold rule turns out to be rather robust to the dating of past financial crises.

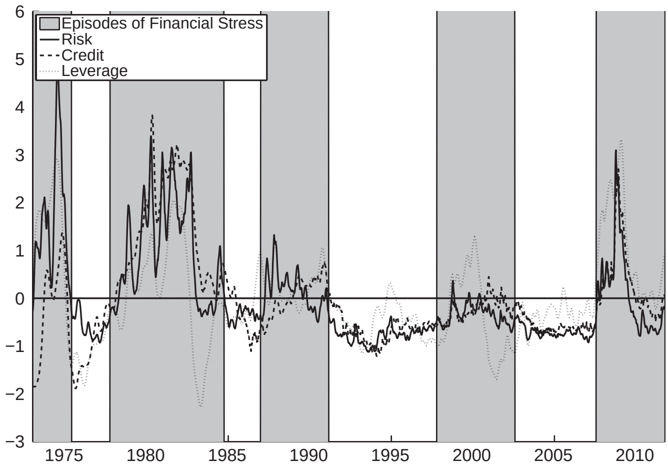
To see this, consider the top panel in figure 5, which depicts the history of the NFCI relative to its historical average and in standard deviation units, shading the crisis episodes in table 1 and plotting the unbiased contemporaneous thresholds based on those episodes and the alternative dates suggested by López-Salido and Nelson (2010). Interestingly, the resulting threshold for the index is a slightly negative number in both cases. This result is intuitive, as it is very apparent in figure 5 that the transition into and out of a crisis is often characterized by a sharp deviation from below and above the index mean, respectively.

The sub-indexes provide further evidence on the source of the initial impetus into and out of a financial crisis. The bottom panel of figure 5 plots each relative to its own mean and in its own standard deviation units. This scaling makes for easier viewing but does

Figure 5. The NFCI and Sub-Indexes



Crisis Dating Convention	Baseline	L-S & N
Contemporaneous Threshold	–0.39	–0.26



NFCI Sub-Index	Risk	Credit	Leverage
Contemporaneous Threshold	–0.39	–0.49	–0.13

obscure the relative contribution of each sub-index to the NFCI. However, it should be readily apparent from the figure that the risk measures are the dominant source of variation in the overall index, as they are nearly identical in quality to the NFCI. The most severe crises in our sample are characterized by above-average values of all three sub-indexes. Leverage is the most cyclical of the three, often leading the others into and out of a crisis. Credit instead tends to follow the more persistent risk sub-index over the course of a crisis with a slight lag.

The contemporaneous threshold values for each of the three, shown in the tables accompanying figure 5, reflect the above facts, with risk demonstrating a threshold largely in line with the NFCI, credit slightly below this, and leverage nearer its long-run average of zero. The behavior of leverage mirrors the results of Adrian and Shin (2010) on the procyclical nature of financial leverage, while the level of the threshold suggests a possible role for financial stability policy in stabilizing measures of financial leverage near their long-run averages. This remains the case even at policy-relevant horizons, where the leverage threshold is mostly unchanged and its AUROC values remain significant.

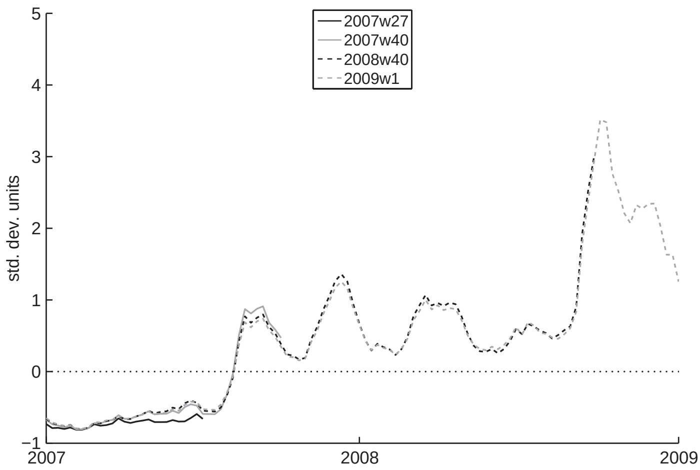
2.4 The Recent Crisis

It is helpful to consider what an unbiased threshold rule would have meant in hindsight for the most recent crisis. Such a rule based on our crisis episodes first signals the development of crisis conditions according to the NFCI in August 2007, nearly in step with our dating. Using the López-Salido and Nelson (2010) dates does little to alter this inference.[6] In fact, this is a feature common to the beginning and end of nearly all of the crises we consider. The leverage sub-index was the first to cross its threshold, as suggested by the results above. However, its leading signal was rather weak given its volatility, crossing the threshold in mid-2005 and again in early 2007, but each time falling back below it shortly thereafter.

However, using the additional information in the NFCI sub-indexes would have considerably qualified the diagnosis of the severity of the recent crisis on several occasions. For instance, the leverage sub-index began to spike first in June 2007, quickly reaching its highest point since the early 1970s. This was the case before the failure

[6]The Laeven and Valencia (2010) dates do not produce a reliable contemporaneous threshold, as they are heavily influenced by the omission of the 1973–75 period.

Figure 6. NFCI Revisions



of Lehman Brothers in the fall of 2008, which only exacerbated the increase until it exceeded its prior high from the early ’70s. Furthermore, after the collapse of Lehman, credit conditions tightened more so than at any point since the early 1980s. Both credit conditions and leverage were then subsequently much slower to return to their average levels than were measures of risk, with the leverage sub-index consistently above its threshold for much of the last two years.

Such an analysis, however, omits a role to be played by the nature of the construction of the NFCI, which is potentially nontrivial for a real-time evaluation of the state of financial conditions. We can address the out-of-sample properties of the NFCI by simulating its production in the period leading up to the most recent crisis. We do so holding fixed the availability of financial indicators at the end of each quarter and using revised data to focus solely on the impact that estimating the scale of the index and the weight assigned to each financial indicator has over this period. Our simulation runs from the second quarter of 2007 through the fourth quarter of 2008, thus capturing the run-up and height of the recent crisis.

Figure 6 is reassuring in the fact that despite very large movements in many of the financial indicators, revisions to the NFCI are small over this period. This suggests very little loss in efficiency in the ROC framework from our method of index construction. The effect of the crisis is initially characterized by a one-time jump in the level of the index of about 0.2 standard deviations between the second and third quarters of 2007. Revisions afterwards are smaller but do occur between the third and fourth quarters of 2007 and 2008. Furthermore, incorporating both data revisions and staggered data availability, the publicly available history of the NFCI since April 2011 seldom demonstrates revisions larger than 0.1 standard deviations in size.

Keep in mind that this analysis still remains subject to the Lucas critique, as it holds fixed both the reaction of financial markets to past policy and policy to past financial market events. It is not intended to be a substitute for a fully specified dynamic model of the interaction between the two. At the very least, it provides a historical basis for judging the current state of the financial system and provides a sense in which measures of leverage may signal the development of financial imbalances. In the conclusion, we discuss potential ways in which our results could be used to inform a policy aimed at financial stability.

3. Robustness of the NFCI

There are several properties of the NFCI that interact closely with the ROC methodology and warrant closer inspection. For instance, our motivation in using the NFCI as a measure of financial stability is closely tied to its systemic interpretation of movements in a number of financial indicators. However, if the nature of those interactions is subject to exogenous breaks over time, the inference provided by the NFCI may in fact be biased. On the other hand, by focusing only on a subsample of its history where these breaks are not likely to have occured, some information is likely to be lost from the omission of past crises.

In what follows, we examine the robustness of the NFCI as a measure of financial stability. We focus our analysis on several key assumptions underlying its construction, namely (i) the lack of structural breaks in the correlation properties of the underlying financial indicators, (ii) the ability of these indicators to span the breadth of financial activity, and (iii) a consistent relationship with economic conditions.

3.1 Structural Breaks

To test for structural breaks in financial activity, we constructed an alternative version of the NFCI using only data from the post-1984 period. Panel A of figure 7 plots both the post-1984 and full-sample NFCI relative to their sample means and scaled by their sample standard deviations.[7] For the period of time in which they overlap, most of the difference between the two indexes appears in their levels, as the variance of the post-1984 index by which it is scaled is considerably smaller than the same measure for the NFCI. This suggests that it is primarily the lower volatility of the post-1984 period that is driving what differences we do see, and is in line with broader findings on the "Great Moderation."

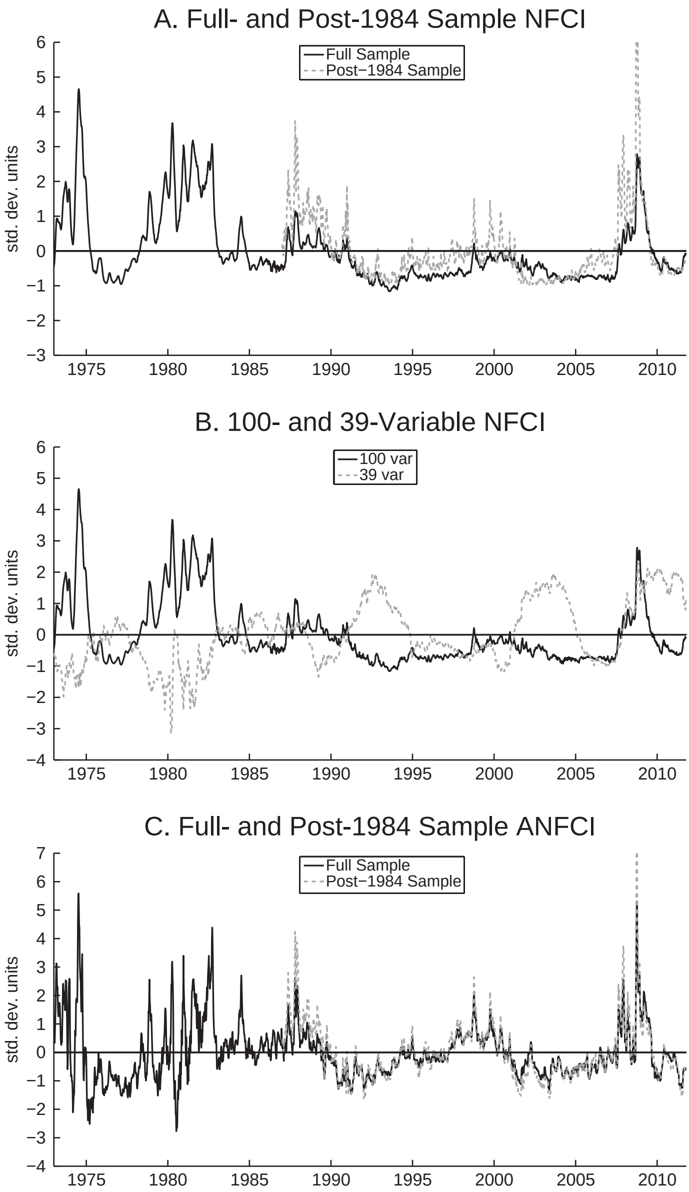
That said, there are noticeable differences in the patterns of the two indexes, particularly over the last decade or so. The factor loadings provide some evidence of where the differences are coming from.[8] For the most part, they reflect only small differences between the indexes. However, it is apparent that the post-1984 index does shift weight away from the more traditional banking indicators and towards more money- and capital-market-centric indicators. We conclude from this that focusing on a more recent time period demonstrates that these indicators have become more important to overall financial conditions.

There are several reasons why this is the case. First, many of these measures reflect the growth of the "shadow" banking system over the last decade at the expense of the commercial banking system. In this sense, there does seem to be some evidence of a structural break in financial activity. Second, many of these measures also played a large part in the most recent crisis. Given how unique the recent crisis was to the period post-1984, they not surprisingly explain considerably more of the co-movement during this period.

[7] Just as with the full-sample index, we do not consider the first two years of estimates, so that the shorter sample index begins in 1987.

[8] The factor loadings for all of the alternative indexes considered in this section are available from the authors upon request.

Figure 7. Stability of the NFCI over Time, Indicators, and Economic Conditions



3.2 Coverage of the Financial System

An alternative to using a shorter sample period is to instead focus only on the subset of financial indicators whose history extends back over most of the sample. In this way, we can further judge whether it is possible to consistently capture financial conditions over an extended period without incorporating information from more recently developed financial markets. Panel B of figure 7 plots the NFCI computed from its thirty-nine financial indicators that extend back to 1978 against the 100-variable index over the same time period.

One can see from panel B of figure 7 that the smaller index is capturing something very different than the larger one. Except for the most recent period where both indexes demonstrate large positive values, the two are highly negatively correlated. Well-known periods of deterioration in financial conditions, such as the late 1970s and early 1980s, appear in the narrower index as very loose periods for financial conditions. In fact, many of the factor loadings are of opposite sign to those in the NFCI.

There are several explanations for the differences between the NFCI and the smaller index. First, the NFCI spans a larger cross-section of financial activity, largely due to its greater inclusion of money- and capital-market-centric measures. Second, the subset of indicators in the smaller index also contains a bias towards those more likely to be affected by the change in volatility post-1984. This can be seen in the fact that many of the same indicators in the thirty-nine-variable index also show large changes in their factor loadings in the post-1984 index.

As an example of where the smaller index seems to be diverging, consider the Treasury yield-curve indicators. Measures from both the short and long end of the curve receive large positive factor loadings in the thirty-nine-variable index, meaning that as the yield curve steepens, the index rises. In contrast, the factor loadings they receive in the NFCI are much smaller and negative, meaning that as the yield curve steepens, financial conditions tend to improve. However, even this relationship is not stable over time. In the post-1984 index, the long end of the curve receives a large positive factor loading, while the short end receives a smaller negative factor loading.

3.3 Adjusting for Economic Conditions

The results above suggest to us that at least some of the instability in the NFCI over time and indicators may be due in part to changes in the level and volatility of economic growth and inflation. To examine this hypothesis, we followed Hatzius et al. (2010) in also constructing both full- and post-1984 sample indexes where each of our 100 financial indicators was first regressed on current and lagged values of a measure of the business cycle, the three-month moving-average Chicago Fed National Activity Index (CFNAI-MA3), and inflation, three-month total PCE inflation.[9]

Panel C of figure 7 plots both the post-1984 and full-sample histories for what we call the adjusted NFCI (ANFCI). For the period of time in which they overlap, the peaks and valleys in the shorter sample index are slightly more pronounced, but the general pattern is very similar to the full-sample index. Their factor loadings confirm that the differences between the two indexes tend to be much smaller than for the NFCI and its post-1984 counterpart. This suggests that after we have accounted for the decline in the volatility of economic growth and inflation, the resulting index is much more stable.

Interestingly, during the recent crisis this was not always the case. Here, because of the large movements in several of the indicators—particularly, the money- and capital-market-centric ones—the small differences in factor loadings across indexes have a more prominent effect. Reinforcing this is the fact that the shorter sample ANFCI also shifts some weight away from the traditional banking system and towards the money and capital markets. Thus, even after accounting for changes in the volatility of economic growth and inflation, it appears that there remains some instability in the NFCI over time attributable to the growth of the shadow banking system.

[9]The number of current and lagged values was chosen for each variable using the Bayesian information criterion (BIC), with the independent variables transformed to match the frequency of observation of the financial indicator. For weekly indicators, we assumed that only lagged values enter the regression and that these values were constant over the weeks of the month.

4. Other Indicators of Financial Stress

In this section, we compare the ability of the NFCI to capture financial stress relative to other measures suggested in the literature and the reformulations of the index described above. We focus on both the contemporaneous and leading abilities of these measures. Then, we put forth alternative leading indexes derived from the NFCI that draw on the leading properties of some of the indicators observed above and compare them against the credit-to-GDP-based and bank conditions measures that have recently received attention in the literature as measures of systemic risk.

4.1 Other Indexes

Table 4 computes the AUROC for each of the alternative versions of the NFCI discussed above. The values for the NFCI are repeated at the top of the table for the sake of comparison. None of the alternative indexes proves to be a superior contemporaneous indicator to the NFCI using our baseline crisis dates. Furthermore, while the post-1984 NFCI and both the shorter and longer sample ANFCI perform slightly better at leading horizons using our baseline dates, the differences are small using the López-Salido and Nelson (2010) crisis dates. The pattern is reversed using the Laeven and Valencia (2010) crisis dates so that the NFCI is much more informative at leading horizons and less informative at the contemporaneous horizon. Interestingly, the thirty-nine-variable index is the most predictive at the contemporaneous horizon using these crisis dates.

There are several reasons for these results. First, the ANFCI, by nature of its adjustment of measures of credit by economic growth and inflation, results in a number of additional measures that exhibit similar leading characteristics to the NFCI leverage sub-index. Second, by focusing on the post-1984 period, the factor loadings for several of these measures become considerably larger. Some factor loadings even change signs relative to those in the NFCI, reflecting a different correlation pattern in the shorter sample. This is true for both risk, credit, and leverage measures, several of which demonstrate leading qualities in table 2. The Laeven and Valencia (2010) dates attenuate these concerns, as they are primarily reflective of the recent crisis where leverage measures played a large role as seen above.

Table 4. AUROC at Contemporaneous and Leading Horizons

Indicator	Baseline				López-Salido and Nelson (2010)				Laeven and Valencia (2010)			
	0 Yr	**−1 Yr**	**−2 Yr**	**−3 Yr**	**0 Yr**	**−1 Yr**	**−2 Yr**	**−3 Yr**	**0 Yr**	**−1 Yr**	**−2 Yr**	**−3 Yr**
NFCI	**0.95**	**0.80**	**0.62**	0.48	**0.76**	**0.79**	**0.73**	**0.59**	**0.62**	**0.56**	**0.45**	**0.32**
Post-1984 NFCI	**0.60**	**0.71**	**0.70**	**0.58**	**0.75**	**0.84**	**0.83**	**0.67**	**0.70**	**0.73**	**0.67**	0.51
39-Variable NFCI	**0.34**	**0.28**	**0.33**	**0.37**	0.51	**0.40**	**0.37**	**0.46**	**0.78**	**0.66**	**0.56**	**0.59**
ANFCI	**0.73**	**0.74**	**0.69**	**0.53**	**0.60**	**0.71**	**0.74**	**0.58**	**0.62**	**0.63**	**0.60**	**0.46**
Post-1984 ANFCI	**0.63**	**0.73**	**0.76**	**0.64**	**0.60**	**0.73**	**0.76**	**0.64**	**0.68**	**0.70**	**0.62**	0.47
STLFSI	**0.86**	**0.63**	**0.43**	**0.19**	—	—	—	—	—	—	—	—
KCFSI	**0.93**	**0.76**	0.56	**0.31**	—	—	—	—	—	—	—	—
IMF	**0.89**	**0.77**	0.58	**0.28**	—	—	—	—	—	—	—	—
Credit-to-GDP	0.59	0.54	0.48	0.40	**0.76**	**0.81**	**0.83**	**0.76**	0.75	**0.81**	**0.82**	**0.74**
Non-Financial Leverage	**0.68**	**0.73**	**0.72**	**0.63**	0.50	**0.70**	**0.83**	**0.87**	**0.58**	**0.73**	**0.86**	**0.83**
Bank Conditions	**0.69**	**0.63**	**0.55**	**0.46**	**0.66**	**0.77**	**0.79**	**0.68**	**0.65**	**0.76**	**0.77**	**0.59**
Adj. Bank Conditions	**0.64**	**0.70**	**0.73**	**0.63**	**0.58**	**0.72**	**0.77**	**0.65**	**0.62**	**0.76**	**0.78**	**0.60**

Notes: The table displays the area under the ROC curve relative to the crisis periods denoted in the column headings contemporaneously and at leads of up to three years for the NFCI and its three sub-indexes. Areas significant from 50 percent are denoted by bold (95 percent significance level) and darkened fonts (90 percent significance level).

Table 4 also computes the AUROC for several competing indexes of financial stress produced by the Federal Reserve Banks of Kansas City and St. Louis—the KCFSI and STLFSI, respectively—and the monthly variant of Hatzius et al. (2010) produced by the International Monetary Fund (IMF) and described in Matheson (2011).[10] Of the three, the IMF measure comes the closest to demonstrating the breadth of coverage of the financial system in the NFCI. In contrast, the KCFSI and STLFSI contain a smaller number of primarily what we term in the NFCI as measures of risk.

All three competing indexes have much shorter histories than the NFCI, each beginning in the 1990s, so that AUROC values in the table are omitted for the alternative crisis dating conventions given their diminished relevance. In addition, other than the STLFSI, the remaining indexes are observed at a monthly frequency. Even after adjusting for these facts when comparing the AUROC values, the NFCI remains just as predictive, if not more predictive, than any of the three competing indexes at a contemporaneous horizon.[11] While not apparent in the table, this is also true at leading horizons when judging the NFCI relative to the others solely on their shared history. All of these indexes were very low in advance of the most recent crisis, which is reflected in an AUROC value well below 0.5 at a three-year-ahead horizon.

4.2 Credit-to-GDP-Based Measures

At the bottom of table 4 are the AUROC values for the HP-filtered private credit-to-GDP measure put forth as a leading indicator of financial stress by Drehmann et al. (2010).[12] While the NFCI matches its predictive ability at short horizons, its appeal can be seen in its very high and statistically significant AUROC values at leads from two to three years using either of the alternative dating

[10]We thank Troy Matheson for kindly making available a time series of his index.

[11]The same can be said of the quarterly index of Hatzius et al. (2010).

[12]We measure private credit as the sum of household mortgage and consumer credit market debt combined with the credit market debt of non-financial businesses. For the smoothing parameter of the HP filter, we use the preferred value of lambda in Drehmann et al. (2010).

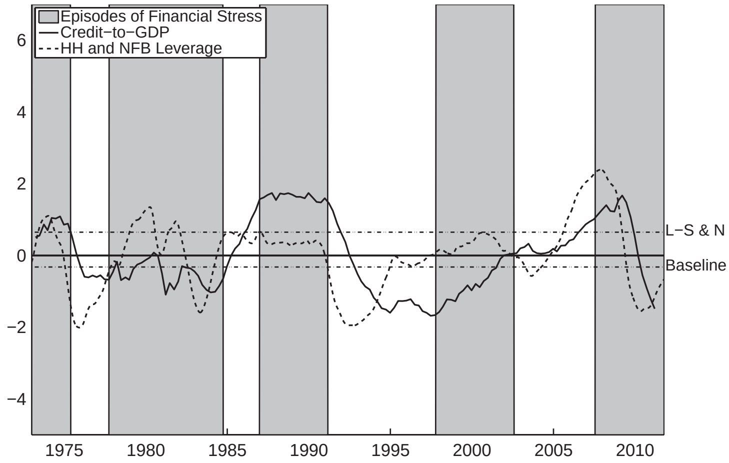
conventions we consider. Similarly, while the post-1984 NFCI performs just as well up to a lead of two years with the López-Salido and Nelson (2010) crisis dates, credit-to-GDP dwarfs it when considering the Laeven and Valencia (2010) dates.

These results are, however, very sample dependent. For instance, both indexes are superior to private credit-to-GDP at all horizons under our baseline crisis dates. To see why, consider figure 8, which plots the credit-to-GDP measure, shading the periods corresponding with our baseline crisis dates. The signal sent by this indicator is very strong for three of the five crisis episodes, i.e., the 1973–75, 1987–91, and most recent crises. It has very little to say about the 1977–84 and 1997–2002 periods we consider, hence the lower AUROC value. In contrast, the alternative crisis dates omit much of these periods, resulting in a higher AUROC value.

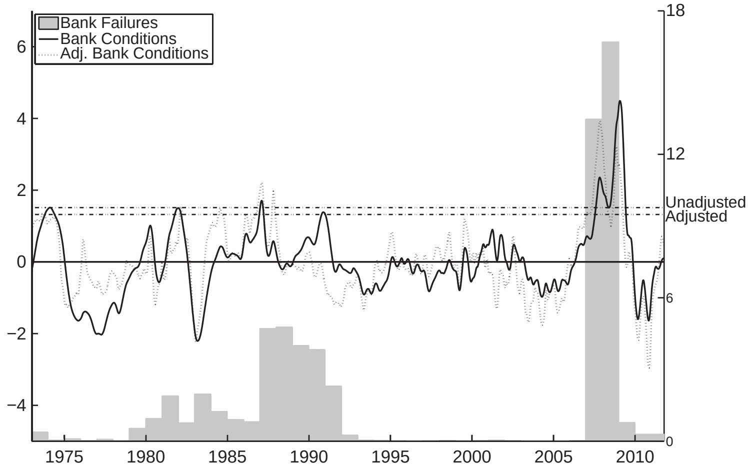
For comparison, consider the sub-index constructed from the two leverage indicators in the NFCI that most closely resemble the inputs to the private credit-to-GDP ratio of Drehmann et al. (2010), referenced as HH and NFB Leverage in the top panel of figure 8 and Non-Financial Leverage in table 4. The signal that emerges offers a picture over time that is more consistent than private credit-to-GDP. It captures all five crisis periods while still producing AUROC values at leading horizons that are on par with private credit-to-GDP using our alternative crisis dates. In addition, it leads all three of the most severe López-Salido and Nelson (2010) crises, with by far its strongest signal occurring in the run-up to the recent crisis.

There are several reasons why our measure differs so markedly from that of Drehmann et al. (2010) despite the use of similar data. The first is that it is unnecessary for us to take a stance on the nature of the trends in household and non-financial business leverage, as we focus on their growth rates. Little is lost in doing so, as it is already apparent from table 2 that the growth rate of non-financial business debt relative to GDP is a highly predictive leading indicator of financial stress. Our measure is also more flexible in that it weights household and non-financial business leverage unequally. Household leverage receives roughly 1.3 times the weight that non-financial business leverage does, as opposed to its equal treatment in the private credit-to-GDP ratio. The weighting we use for our measure based on the systemic decomposition of the NFCI turns

Figure 8. Leading Indicators of Financial Stress



Crisis Dating Convention	Baseline	L-S & N
Two-Year-Ahead Threshold	–0.32	0.65



Crisis Dating Convention	Unadj.	Adj.
Two-Year-Ahead Threshold	1.51	1.32

Note: "Bank Failures" refers to the annual percentage of failed/assisted institutional assets according to FDIC statistics relative to the total assets of the commercial banking sector taken from the Federal Reserve's H.8 statistical release.

out to be non-trivial, as there are several instances in time where the movements in the two measures are conflicting.

In addition, we measure household leverage as household mortgage and consumer debt relative to residential investment and personal consumption of durable goods rather than to overall GDP. This also turns out to be non-trivial because, expressed in this way, household leverage receives a positive weight in the NFCI. It thereby serves to amplify the signal from non-financial business leverage, whereas expressed relative to GDP it has the opposite effect. The flexibility thus provided by our measure makes it possible to produce a combined signal that is superior to that of each of the individual measures, particulary at leading horizons.

The positive weight assigned to each of these measures in the NFCI reflects the fact that rising household and non-financial business leverage are typically associated with increasingly tighter financial conditions. This is in contrast to many of the measures of financial leverage in the NFCI which instead show the opposite correlation with financial conditions. It also makes them characteristic of the financial accelerator mechanism of Bernanke, Gertler, and Gilchrist (1999). Increasingly tighter financial conditions are associated with rising risk premia and declining asset values. Hence, the net worth of households and firms is reduced at the same time that credit tightens. This leads to a period of deleveraging and ultimately to lower economic activity, making these measures relevant for both credit and business cycles.

It is also possible to produce meaningful thresholds at a leading horizon for our measure based on our baseline and the López-Salido and Nelson (2010) crisis dates. We plot these thresholds in the top panel of figure 8 for a leading horizon of two years. Instances of where this measure falls above both thresholds are characteristic of the most severe crises. It is instructive that these coincide with the three deepest and longest U.S. recessions in the post-War era. Instances in between are consistent with more moderate crises and less pronounced recessions. This suggests that policies which aim to minimize the deviation of measures of non-financial leverage from their long-run averages are likely to fall within the realm of both financial stability and the welfare costs of business cycles.

The longer leading horizon considered here is important given that implementation of any such policy would require sufficient lead time. As an example, consider again the most recent crisis. Our measure was at its historical average as recently as the end of 2004. It first crossed the upper threshold in late 2005, nearly two years before the onset of the crisis in August 2007. That said, revision error is likely to be problematic for inference, because the underlying data are subject to large revisions and lags in availability. However, our method of constructing this measure using the architecture for the NFCI should help to minimize these concerns. This is because we can project our measure forward in time based on the higher frequency availability of the NFCI, thereby disciplining its near-term movements by its correlations with the other indicators in the index. In this way, the NFCI still plays an important role in monitoring financial stability in real time.

4.3 Bank Conditions Measures

It is interesting that our measure of non-financial leverage also remains a superior leading indicator to private credit-to-GDP even when using the Laeven and Valencia (2010) crisis dates. These dates were designed to capture periods of systemic bank failure in the United States, i.e., primarily the savings and loan crisis of the late 1980s and the recent crisis. Based on this fact, it is easy to see why in figure 8 private credit-to-GDP was proposed by Drehmann et al. (2010) as a useful leading indicator of systemic risk, as its two highest peaks precede both crises. One might, however, question the magnitude of the signals with the movements in the 1980s more persistent and just as pronounced as the recent crisis. Our measure of non-financial leverage does not feature these concerns.

To assess the robustness of our measure, we also constructed sub-indexes from a broader set of NFCI indicators considered by Laeven and Valencia (2008) as indicators of systemic risk. The indicators included those in our non-financial leverage measure above as well as public debt-to-GDP, bank liabilities and conditions data, and asset prices. Laeven and Valencia (2008) also suggest using measures of economic conditions to assess systemic risk. To achieve this, we constructed sub-indexes of these measures using both the NFCI and ANFCI methodologies, the latter of which adjusts the indicators for growth in economic activity and inflation, while the former does not.

The bottom panel of figure 8 depicts both sub-indexes as well as the relative size of the assets of failed/assisted banking institutions. We include the latter to highlight the magnitudes of the two systemic banking crises. Both sub-indexes quite clearly capture the recent crisis as the worst period for bank conditions in our sample. Other spikes in bank failures are also captured in the peaks of the indexes, although few other than the peaks preceding the savings and loan crisis breach the unbiased two-year-ahead thresholds shown in the figure for both sub-indexes.

Looking again at table 4, it is clear that the difference in predictive ability at shorter forecast horizons between the two sub-indexes (referred to as Bank Conditions and Adj. Bank Conditions) is rather small, as the AUROC values are very similar across all three crisis dating conventions. At longer forecast horizons, the adjusted version is superior only in the case of our baseline crisis dates, and this is also the only instance where either of the two indexes produces a leading signal comparable to that of our measure of non-financial leverage. In this sense, there seems to be little additional value to be had from combining non-financial leverage with other indicators to assess the risk of systemic bank failure.

5. Conclusion

Our analysis suggests that the National Financial Conditions Index (NFCI) produced by the Federal Reserve Bank of Chicago is a highly robust and accurate indicator of financial stress at leading horizons of up to one year. Furthermore, breaking the index into its subcomponents of risk, credit, and leverage can enhance the nature of the signal provided by the NFCI as to the severity of the crisis, with leverage playing a crucial role in signaling financial imbalances as in Adrian and Shin (2010). Moving to a multi-factor representation of the NFCI, which makes full use of the different signals in its three types of indicators and their dynamic properties, is a non-trivial extension which we leave to future research.

At forecast horizons beyond one year, we show that a particular combination of household and non-financial business leverage measures proves to be a consistent leading indicator of financial stress and its impact on economic activity. This result is by virtue of its resemblance to the financial accelerator mechanism of Bernanke, Gertler, and Gilchrist (1999). Increases over time in these measures are associated with increasingly tighter financial conditions, which lead to periods of deleveraging and ultimately lower economic activity. By combining their signal into a single measure based on their relative weighting in the NFCI, we arrive at a combined signal that is superior to several alternative measures of systemic risk suggested in the literature.

The flexibility of the ROC methodology is its greatest asset. However, it leaves unaddressed the types of policy actions that can and should be used to address any financial imbalances it identifies, as well as the magnitude of their likely impact. At the very least, though, our analysis establishes properties that could be explored in the formulation of optimal policy within a general equilibrium framework.[13] In particular, policies which aim to minimize the deviations of financial and non-financial leverage from their long-run averages are consistent with our analysis of financial stability and may also be beneficial for minimizing the welfare costs of business cycles.

Appendix. A Dynamic Factor Model of Financial Conditions

The model for the National Financial Conditions Index is fundamentally similar to many of the dynamic factor models summarized in Stock and Watson (2011). Its state-space representation is shown below, where y_t is a vector of stationary variables that have been demeaned and standardized; ε_t and η_t are idiosyncratic error vectors with $\varepsilon_t \sim N(0, H)$ and $\eta_t \sim N(0, Q)$; α_t is a vector made up of a latent coincident factor f_t and its $L^\star - 1$ lags; and $t = 1, \ldots, \hat{T}$, where $\hat{T}$ is the longest time-series length of the collection of $\hat{N}$ financial indicators in the NFCI.

$$y_t = Z\alpha_t + \varepsilon_t \tag{6}$$

$$\begin{aligned} \alpha_{t+1} &= T\alpha_t + R\eta_t, \\ & a_1 = \mathbb{E}[\alpha_1] \text{ and } P_1 = Var[\alpha_1] \text{ given.} \end{aligned} \tag{7}$$

[13]For instance, Brave and Genay (2011) provide some evidence that during the recent crisis, traditional and non-traditional monetary policy actions contributed to the recovery in financial conditions as measured by the NFCI and ANFCI in a joint model of Federal Reserve policy and financial conditions.

Written in this form, the latent factor is identified based on both the historical cross-correlations of the vector of variables y_t and its own historical autocorrelations embedded in the system matrix T in the state equation, (7). In contrast, the variable cross-correlations enter via the model's observation equation, (6), relating each of the $\hat{N}$ variables to the coincident latent factor via the factor loadings λ included in the system matrix Z. Identification is achieved only up to scale, as initial conditions for the mean and variance of the latent factor—a_1 and P_1, respectively—are necessary to complete the model.

This model can be estimated using the procedure outlined in Doz, Giannone, and Reichlin (2006) based on the EM algorithms of Shumway and Stoffer (1982) and Watson and Engle (1983) in order to obtain quasi-maximum-likelihood estimates of the system matrices and subsequently the latent factor. In general, it requires one pass through the Kalman filter and smoother, and then reestimation of the system matrices—Z, T, H, and Q—using ordinary least squares (OLS) estimation at each iteration.[14] The resulting sequence of log-likelihood function valuations is non-decreasing, and convergence of the algorithm is governed by its stability according to the suggested procedure in Doz, Giannone, and Reichlin (2006).

A unique characteristic of y_t in the context of the NFCI is that it contains series of varying reported frequencies and series that start and end at different times within the sample. The EM algorithm proves advantageous in this setting because it allows for a complete characterization of the data-generating process using incomplete data. However, applying this procedure to our particular data set requires some changes to the standard Kalman filter and smoother equations as described below.

Missing-Value Kalman Filter

Due to the irregular observation of the data in our framework, two extensions to the standard Kalman recursion equations need to first be made before running the EM algorithm. The first alteration involves setting up the Kalman filter to deal with missing values as

[14] A small alteration in the least-squares step is required to account for the fact that the unobserved components of the model must first be estimated.

discussed by Durbin and Koopman (2001). As one moves through time, the vector of observables for the NFCI, y_t, changes size from period to period. Consequently, it is necessary to accommodate the partially observed vector $y_t^\star$ following Durbin and Koopman (2001) by using the known matrix W_t whose rows are a subset of the rows of $I(\hat{N})$ such that $y_t^\star = W_t y_t$ to alter the system matrices at each point in time. Taking this W_t as given, the system matrices Z and H are replaced with $Z^\star = W_t Z$ and $H^\star = W_t H W_t'$, respectively. Substituting these matrices into the standard Kalman filter and smoother equations allows one to proceed as usual through the recursive equations.

Temporal Aggregation and the Harvey Accumulator

The second necessary modification involves including additional state variables that evolve deterministically to properly adjust for the varying temporal aggregation properties of the mixed-frequency data used to create the NFCI. By applying the accumulator of Harvey (1989), one can manage this data irregularity with relative ease. The goal of the accumulator is to augment the state, α_t, with a deterministically evolving indicator that is a summary of all past values of the unobserved factor aggregated in such a way as to correspond with the nature of the observed data.

More specifically, variables viewed as a “stock,” or a *snapshot* in time, will not need such aggregation of past realizations of the factor. Variables that correspond to sums or averages over the higher base frequency of the factor will need to accumulate the factor realizations over a defined period in order to properly account for the contemporaneous factor’s contribution to what is being observed. Any “stocks” that are differenced can be interpreted as sum variables and treated as such.

The NFCI includes variables that resemble both “sums” and “averages,” in addition to variables that are first-differenced at lower frequencies than the weekly (base) frequency. Combining this with monthly and quarterly frequencies of observation, the model for the NFCI requires three Harvey accumulators in the state, one each for (i) monthly averages, (ii) monthly sums, and (iii) quarterly sums.

Sum-Variables Accumulator

For both monthly and quarterly sums, we follow Aruoba, Diebold, and Scotti (2009)'s implementation of the Harvey (1989) accumulator. The accumulators for sum variables are denoted by S_t. By construction, any sum accumulator should represent the sum of all of the factor realizations that have occurred within the current period of the lower frequency of observation. Additionally, the accumulator is defined recursively so as to be included in the state-space equations (6)–(7). Analytically, the sum accumulator evolves each period according to the following equation:

$$S_{t+1} = s_t S_t + f_{t+1},$$

where s_t is a calendar-determined indicator that evolves according to

$$s_t = \begin{cases} 0 & \text{if } t \text{ is the last period (base frequency)} \\ & \quad \text{within the lower frequency} \\ 1 & \text{otherwise.} \end{cases}$$

For notational purposes, it is assumed in what follows that f_{t+1} is an AR(1) process defined by $f_{t+1} = \rho f_t + R\eta_t$. Incorporating this representation of the accumulator into the state-space model follows from a simple substitution of the contemporaneous factor as outlined by Aruoba, Diebold, and Scotti (2009).

Average-Variables Accumulator

We denote the desired accumulator for the average variables with M_t and derive it as though we are aggregating from a weekly base frequency to monthly observations.[15] By construction, this accumulator should represent the *current* average of all of the factor realizations (occurring every week) that have occurred within the current month (frequency that is being observed) and be defined

[15] All methods outlined in this section generalize fully to any particular combination of base and observation frequencies that one might encounter, with the only necessary modifications occurring in the evolution of the calendar indicator m_t or s_t.

recursively for seamless addition to the state-space equations (6)–(7). Analytically, the average accumulator evolves each period by the following equation:

$$M_{t+1} = \frac{(m_t - 1)}{m_t} M_t + \frac{1}{m_t} f_{t+1},$$

where m_t is a calendar-determined indicator that evolves:

$$m_t = \begin{cases} 1 & \text{if } t \text{ is the } \textit{last} \text{ week of the month} \\ 2 & \text{if } t \text{ is the first week of the month} \\ 3 & \text{if } t \text{ is the second week of the month} \\ \textit{etc.} & \end{cases}$$

Explicitly including the accumulator in the state requires augmenting the state and some substitution. The resulting formulation is given by

$$\begin{bmatrix} f_{t+1} \\ M_{t+1} \end{bmatrix} = \begin{bmatrix} \rho & 0 \\ \frac{\rho}{m_t} & \frac{m_t - 1}{m_t} \end{bmatrix} \begin{bmatrix} f_t \\ M_t \end{bmatrix} + \begin{bmatrix} R \\ \frac{R}{m_t} \end{bmatrix} \eta_t.$$

Building the State

This section gives a more detailed explanation of how to build the system matrices of the state space given the two necessary extensions to accommodate the irregular observation of the NFCI's observables described above. The latent factor is assumed to have some finite-order dynamics $L^\star$, given by the vector of coefficients, ρ. Augmenting the state to include $L^\star - 1$ lags of f_t yields the following state equation, with $\alpha_t = [f_{t-i}]$ for $i = 0, \ldots, L^\star - 1$:

$$\alpha_{t+1} = \begin{bmatrix} \rho & 0 \\ I(L^\star - 1) & 0_{L^\star - 1 \times 1} \end{bmatrix} \alpha_t + \begin{bmatrix} 1 \\ 0_{L^\star - 1 \times 1} \end{bmatrix} \eta_t.$$

Now, in this particular representation of the state, the system matrix $\tilde{T}$ is defined by

$$\tilde{T} = \begin{bmatrix} \rho & 0 \\ I(L^\star - 1) & 0_{L^\star - 1 \times 1} \end{bmatrix}.$$

One must also augment the state (currently a $L^\star$ long vector) by the additional states needed for each of the accumulators derived in the previous section to yield the state equation (taking the ρ dynamics again as given):

$$\begin{bmatrix} \alpha_{t+1} \\ M_{t+1} \\ S_{t+1} \end{bmatrix} = \begin{bmatrix} \tilde{T} & 0 & 0 & 0 \\ \frac{\rho}{m_t} & \frac{m_t - 1}{m_t} & 0 & 0 \\ \rho & 0 & s_t & 0 \end{bmatrix} \begin{bmatrix} \alpha_t \\ M_t \\ S_t \end{bmatrix} + \begin{bmatrix} 1 \\ 0_{L^\star - 1 \times 1} \\ \frac{1}{m_t} \\ 1 \end{bmatrix} \eta_t. \qquad (8)$$

It should be noted that, as written above, the $\tilde{T}$ *within* the general transition system matrix, T, here is time invariant, and subsequently the dynamics being estimated (essentially a reestimation of ρ) at each iteration of the EM algorithm are from a *time-invariant* system. However, our (accumulator-augmented) state transition system matrix (as well as the coefficient matrix on the η_t) *does vary over time* due to the different number of weeks in a given month, or quarter, and the particular evolution of the average accumulators.

Moving to the measurement equation, assume that a priori the vector of factor loadings λ is known. Then, taking the state equation (8) as given, the Z measurement system matrix is simply an $\hat{N}$ by $L^\star + 3$ matrix, where each row has the particular loading (the particular element of λ) in either the first column (if it corresponds with a weekly or stock variable) or one of the last three columns (corresponding to one of the three accumulators, i.e., monthly average, monthly sum, or quarterly sum) and zeros everywhere else. Finally, the system matrices H and Q are the standard variance-covariance matrices of ε_t and η_t, respectively, as described in Doz, Giannone, and Reichlin (2006), where H is assumed to be a diagonal matrix and Q normalizes the scale for the latent factor.

Kalman Filter and Smoother Recursive Equations

Adopting the notation of Durbin and Koopman (2001), the Kalman filter and smoother equations taking into account the state-space structure above are as follows.[16] With a_1 and P_1 given, the filter equations are

[16] For more details on the derivation of these equations, see Durbin and Koopman (2001, pp. 64–73).

$$
\begin{aligned}
v_t &= y_t^\star - Z^\star a_t & F_t &= Z^\star P_t Z^{\star\prime} + H^\star \\
K_t &= T_t P_t Z^{\star\prime} F_t^{-1} & L_t &= T_t - K_t Z^\star \\
a_{t+1} &= T_t a_t + K_t v_t & P_{t+1} &= T_t P_t L_t' + R_t Q R_t',
\end{aligned}
$$

where $y_t^\star$, $Z^\star$, and $H^\star$ are the truncated versions of the more general vector and system matrices due to any potential missing observations in the vector y_t as described above.

Likewise, the equations for the backwards smoother are given by[17]

$$
\begin{aligned}
r_{t-1} &= Z^{\star\prime} F_t^{-1} v_t + L_t' r_t & N_{t-1} &= Z^{\star\prime} F_t^{-1} Z^\star + L_t' N_t L_t \\
\tilde{\alpha}_t &= a_t + P_t r_{t-1} & V_t &= P_t - P_t N_{t-1} P_t \\
J_t &= P_{t-1} L_{t-1}' (I - N_{t-1} P_t), & &
\end{aligned}
$$

with $r_{\hat{T}} = 0$, and $N_{\hat{T}} = 0$.

The EM Algorithm

Given the missing observation and accumulator extensions to the Kalman filter, the system parameters can be estimated via the EM algorithm of Shumway and Stoffer (1982) and Watson and Engle (1983). Starting at some initial values for every system matrix—Z^0, H^0, T^0, Q^0, P_1^0, and a_1^0—each iteration of the algorithm consists of one pass through the Kalman filter and smoother using the system matrices estimated at the previous iteration r—Z^r, H^r, T^r, Q^r, P_1^r, and vector a_1^r. The typical critique of the algorithm, its slow convergence rate, is not problematic in this setting due to the size of the time-series and cross-section dimensions for the NFCI, which allow for consistent initial values using the iterative PCA techniques described in Stock and Watson (2002).

The initial values of the Z system matrix Z^0 include the estimated factor loadings λ, while the initial guess of the H system matrix H^0 is simply the estimated variances of the idiosyncratic errors from the same analysis. The initial guess of the T system

[17]It should be noted that the additional matrix, J_t, is being calculated so that the maximization step in the EM algorithm can take into account the uncertainty in the estimation of the state.

matrix T^0 is given by standard OLS techniques on the initial estimate of the factor F_t^0 with a chosen lag length of fifteen weeks, or roughly one quarter's worth of observations, determined based on the BIC. The initial guess of a_1^0 is set to zero and the initial guess of P_1^0 is set at some reasonable baseline value. Furthermore, we discard two years worth of estimated data to avoid any issues with this initialization.

The lack of identification that is common to these models requires that we restrict the scale of either the factor loadings or the factor. We use the normalization of Doz, Giannone, and Reichlin (2006) and restrict the variance of the state disturbances to be 1 to set the scale of the factor. By utilizing both the smoothed estimates and their covariance matrices, one can update the expectation of the conditional log-likelihood function, the (E) step. A concise version of the log-likelihood, and the one that can be computed at each iteration, is as follows:

$$\log L = -\frac{1}{2}\sum_{t=1}^{\hat{T}} \big(\log|F_t| + v_t' F_t^{-1} v_t\big), \tag{9}$$

where both v_t and F_t are given by the forward recursion equations of the Kalman filter. Then, using OLS techniques, the system matrices are reestimated as before, the (M) step. This reestimation yields new values for the system matrices—Z^{r+1}, H^{r+1}, T^{r+1}, Q^{r+1}, P_1^{r+1}, and vector a_1^{r+1}. Repeating this process yields a non-decreasing sequence of log-likelihood values. With the (E) and (M) steps completely defined, one can iterate between the two until (9) becomes stable.[18]

Constructing Sub-Indexes

To construct sub-indexes of the NFCI, post-estimation we replace the elements of λ corresponding to variables not in the sub-index with zeros. We then make an additional pass through the equations

[18] As a convergence criterion, we used $|\log L(r) - \log L(r-1)/((\log L(r) + \log L(r-1))/2)| < 10^{-6}$. With this criterion, the EM algorithm converges rather quickly, i.e., generally within 150 iterations.

of the Kalman filter and smoother with the system matrices estimated from the last (M) step of the EM algorithm for the NFCI. In this fashion, we hold fixed both the relevant factor loadings and dynamics of the latent factor from the fully estimated model. As with the NFCI, the scale of each sub-index is normalized according to the variance-covariance matrix Q. This method is equivalent to the two-step consistent estimation procedure as described in Doz, Giannone, and Reichlin (2006).

References

Adrian, T., and H. S. Shin. 2010. "Liquidity and Leverage." *Journal of Financial Intermediation* 19 (3): 418–37.

Aruoba, S. B., F. X. Diebold, and C. Scotti. 2009. "Real-Time Measurement of Business Conditions." *Journal of Business and Economic Statistics* 27 (4): 417–27.

Baker, S. G., and B. S. Kramer. 2007. "Pierce, Youden, and Receiver Operating Characteristic Curves." *The American Statistician* 61 (4): 343–46.

Berge, T. J., and O. Jordà. 2011. "Evaluating the Classification of Economic Activity into Recessions and Expansions." *American Economic Journal: Macroeconomics* 3 (2): 246–77.

Bernanke, B. S., M. Gertler, and S. Gilchrist. 1999. "The Financial Accelerator in a Quantitative Business Cycle Framework." In *Handbook of Macroeconomics*, ed. J. B. Taylor. Elsevier.

Brave, S., and R. A. Butters. 2011. "Monitoring Financial Stability: A Financial Conditions Index Approach." *Economic Perspectives* (Federal Reserve Bank of Chicago) 35 (Q1): 22–43.

Brave, S., and H. Genay. 2011. "Federal Reserve Policies and Financial Market Conditions during the Crisis." Federal Reserve Bank of Chicago Working Paper No. 2011-04.

DeLong, E., D. DeLong, and D. Clarke-Pearson. 1988. "Comparing the Areas under Two or More Correlated Receiver Operating Characteristic Curves: A Nonparametric Approach." *Biometrics* 44: 837–45.

Doz, C., D. Giannone, and L. Reichlin. 2006. "A Quasi Maximum Likelihood Approach for Large Approximate Dynamic Factor Models." ECB Working Paper No. 674.

Drehmann, M., C. Borio, L. Gambacorta, G. Jimenez, and C. Trucharte. 2010. "Countercyclical Capital Buffers: Exploring Options." BIS Working Paper No. 317.

Durbin, J., and S. J. Koopman. 2001. *Time Series Analysis by State Space Methods.* Oxford University Press.

FDIC (Federal Deposit Insurance Corporation). 1984. *The First Fifty Years: A History of the FDIC 1933–1983.* Federal Deposit Insurance Corporation.

———. 1997. *History of the Eighties — Lessons for the Future.* Federal Deposit Insurance Corporation.

Hakkio, C. S., and W. R. Keeton. 2009. "Financial Stress: What Is It, How Can It Be Measured, and Why Does It Matter?" *Economic Review* (Federal Reserve Bank of Kansas City) (Second Quarter): 5–50.

Harvey, A. 1989. *Forecasting, Structural Time Series Models and the Kalman Filter.* Cambridge University Press.

Hatzius, J., P. Hooper, F. Mishkin, K. L. Schoenholtz, and M. W. Watston. 2010. "Financial Condition Indexes: A Fresh Look After the Financial Crisis." U.S. Monetary Policy Forum Paper Series.

Illing, M., and Y. Liu. 2006. "Measuring Financial Stress in a Developed Country: An Application to Canada." *Journal of Financial Stability* 2 (3): 243–65.

Laeven, L., and F. Valencia. 2008. "Systemic Banking Crises: A New Database." IMF Working Paper No. 08/224.

———. 2010. "Resolution of Banking Crises: The Good, the Bad, and the Ugly." IMF Working Paper No. 10/146.

López-Salido, D., and E. Nelson. 2010. "Postwar Financial Crises and Economic Recoveries in the United States." Mimeo, Board of Governors of the Federal Reserve System.

Matheson, T. 2011. "Financial Conditions Indexes for the United States and Euro Area." IMF Working Paper No. 11/93.

Nelson, W. R., and R. Perli. 2007. "Selected Indicators of Financial Stability." *IFC Bulletin* (Irving Fisher Committee on Central-Bank Statistics) 23 (October): 92–105.

Reinhart, C. M., and K. S. Rogoff. 2009. *This Time Is Different: Eight Centuries of Financial Crises.* Princeton, NJ: Princeton University Press.

Schreft, S. L. 1990. "Credit Controls: 1980." *Economic Review* (Federal Reserve Bank of Richmond) (Nov/Dec): 25–55.

Shumway, R. H., and D. S. Stoffer. 1982. "An Approach to Time Series Smoothing and Forecasting Using the EM Algorithm." *Journal of Time Series Analysis* 3 (4): 253–64.

Spero, J. E. 1999. *The Failure of the Franklin National Bank: Challenge to the International Banking System.* Beard Books.

Stock, J. H., and M. W. Watson. 2002. "Forecasting Using Principal Components from a Large Number of Predictors." *Journal of the American Statistical Association* 97 (460): 1167–79.

———. 2011. "Dynamic Factor Models." In *Oxford Handbook of Economic Forecasting*, ed. M. P. Clements and D. E. Hardy, 35–60. Oxford University Press.

Watson, M. W., and R. F. Engle. 1983. "Alternative Algorithms for the Estimation of Dynamic Factor, Mimic and Varying Coefficient Regression Models." *Journal of Econometrics* 23 (3): 385–400.