Credit Spreads and the Severity of Financial Crises *

Arvind Krishnamurthy
Stanford GSB and NBER

Tyler Muir
Yale SOM

June 3, 2015

Abstract

We study the behavior of credit spreads and their link to economic growth during financial crises. Our main finding is that the recessions that surround financial crises are longer and deeper than the recessions surrounding non-financial crises. The slow recovery from the 2008 crisis is in keeping with historical experience of recoveries from financial crises. We reach this conclusion by examining the relation between credit spreads and economic growth in the cross-section, across many countries and crisis-events. This cross-sectional approach differs from much of the existing literature, which instead studies the average GDP performance in countries for a set of specified crisis dates. We argue that the cross-sectional approach avoids some important shortcomings in existing studies.

*Stanford Graduate School of Business and Yale School of Management. We thank Alan Taylor, Francis Longstaff, and seminar/conference participants at the AFA, Stanford University, and University of California Davis.

1 Introduction

We examine the behavior of credit spreads and output around financial crises in an international panel spanning from the 1800s to the present, uncovering new facts on the depth and duration of financial crises. Our research contributes to a growing literature on financial crises, with prominent recent contributions by Reinhart and Rogoff (2009b) on the slow recoveries from financial crises and Jorda *et al.* (2010) on the role of credit growth in forecasting crises.

Credit spreads are known to be a leading indicator for economic activity (see, for example, Gilchrist and Zakrajsek (2012)). The forecasting power of spreads is because: (1) default rates are higher in downturns, and since spreads embed expectations of future default, they offer a leading indicator of downturns; (2) Economic risk premia attached to default states are higher than to boom states, and since spreads reflect economic risk premia, they are a leading indicator of downturns; (3) In many models of financial frictions, credit spreads measure the external finance premium (the tightness of borrowing constraints), so that high spreads cause a reduction in borrowing, investment, and economic activity.

Because credit spreads offer a forward looking market-based assessment of financial crises, they can be helpful in better understanding the behavior of the economy in the aftermath of a financial crisis. Existing research on financial crises groups crises together, without distinguishing among the severity of different crisis events. In our sample of crises, the mean peak-to-trough contraction is -6.2%, but the standard deviation of this measure is 7.3% (see Table 8). The oft-cited number from Reinhart and Rogoff (2009a) is that the peak-to-trough contraction in crises is 9.3%, with this mean measured from a select sample of financial crises. Many academics and policy-makers are interested in answer the question, "How slow a recovery should we expect following the 2008-2009 financial crisis in the US?" We provide a more precise answer than the mean decline across a sample of crises.

Credit spreads in the first year of a crisis correlate with the subsequent severity of the crisis, so that credit spreads allow one to distinguish meaningfully among different crises-events. We find that there is large variation in what to expect when a crisis occurs based on the behavior of spreads and that there is substantial heterogeneity in the severity of crises. We use the rise in the spread in the initial year of a crisis to benchmark the severity of the financial disruption. Using the 2008 spread to benchmark the recent crisis, we plot the predicted path of GDP. This path is remarkably close to the actual path of GDP, suggesting that the realized growth of GDP in the US is in line with what should have been expected based on past financial crises.

A crisis involves a large reduction in the mean growth of GDP, but a much smaller reduction in the median growth of GDP. We date a crisis as an episode where spreads spike above a threshold. Using this method (which differs from the approach taken in other studies), we find that the conditional distribution of output growth is substantially left-skewed. The mean peak-to-trough GDP decline is -5.9% while the median decline is -2.5%. We examine the forecasting power of spreads for output growth via quantile regressions. Most of the forecasting power comes in the lower growth quantiles. The picture that emerges from this data is that when there is a financial crisis, as measured by a spike in credit spreads, the economy follows one of two trajectories: (1) either the crisis dissipates quickly with spreads coming back down and GDP unaffected; or, (2) the economy enters a deep and protracted recession – as it has in the slow recovery following the 2008-2009 financial crisis.

These effects we document on the conditional distribution of output growth point to a bias in existing studies. Existing studies date crises ex-post. The Great Depression is dated a financial crisis because ex-post there were bank failures and a significant collapse in output. However, the 1970 Penn Central crisis or the 1998 hedge fund crisis are not labeled crises in Reinhart and Rogoff (2009b) and Jorda *et al.* (2010) because there is little apparent contagion to the banking sector and the real economy. However, in both of these events, credit spreads rose sharply on the incidence of the crisis. But apparently in these events, either through luck or adequate government intervention, the financial crisis dissipates without leading to knock-on effects to the real economy. By dating crises ex-post, research may be biased toward selecting the worst financial crises episodes, which then biases inferences about the depth and duration of a crisis. Quantitatively this bias is apparent for the crisis dates of Jorda *et al.* (2010) and less so for the dates of Reinhart and Rogoff (2009b).

We investigate whether there are ex-ante variables that can help inform us on when a crisis will be long and protracted. We find that the leverage measure in Jorda *et al.* (2010) is one variable. A sharp rise in spreads that occurs when Jorda *et al.* (2010)'s leverage measure is also high leads to worse GDP outcomes. One way of understanding this result is that sometimes a financial disruption triggers losses within the financial sector which then leads to a rise in risk premia on intermediated assets (e.g., credit), as in the "intermediary asset pricing" model of He and Krishnamurthy (2012). This trigger has a large impact on the real sector only if it interacts with a real amplification mechanism. Jorda *et al.* (2010)'s leverage variable may measure the strength of the real amplification mechanism.

Last, we address the question of whether recoveries following financial crises are especially slow. We compare the growth path of output following a financial crisis with the growth path of output in a deep recession, conditional on the same increase in credit spreads. In particular, we estimate the relation between spreads and subsequent output in recessions and the relation between spreads and subsequent output in financial crises. We find that the coefficient on spreads in the latter relation is much higher than in the former. Then we consider a hypothetical experiment where spreads increase by 100 basis points in both a recession and a financial crisis, and use the historical relation to compute the difference in GDP growth rates across these episodes. By conditioning on the same credit spread shock, we are attempting to "match" on similar real economic circumstances, although the match is likely imperfect. Under a variety of approaches, we consistently find that GDP growth is lower in recessions accompanying financial crises than in recessions without financial crises. We also find that credit spreads recover more quickly than output following a financial crises. Spreads fall back to near normal with one-year, while output typically is still low 4-years after the crisis. This suggests that while a financial crisis is short-lived, the real disturbance associated with the crisis is long-lived. This historical pattern, which emerges from many episodes across many countries, is remarkably consistent with the experience of the US economy in the 2008 financial crisis and subsequent slow recovery. These results suggest that the recovery from financial crises is indeed slower than the recovery from recessions.

Our research does not address important causality questions such as, do high spreads cause crises? Rather it documents facts on the behavior of GDP and credit spreads, and helps to better answer the important question of, is growth especially low following financial crises. The facts we uncover are also important for future research. For example, our finding that credit spreads revert more quickly to normal than GDP is a fact that macroeconomic models must address, as it suggests a complex dynamic relation between financial markets and the real sector.

Our paper contributes to a growing recent literature on the aftermath of financial crises. The most closely related papers to ours are Reinhart and Rogoff (2009b), Jorda *et al.* (2010), Bordo and Haubrich (2012), Cerra and Saxena (2008), Claessens, Kose and Terrones (2010) and Romer and Romer (2014). This literature generally finds that the recoveries after financial crises are particularly slow compared to deep recessions, although Bordo and Haubrich (2012) examine the US experience and dispute this finding, showing that the slow-recovery pattern is true only in the 1930s, the early 1990s and the 2008-2009 financial crisis. Relative to these papers, we consider data on credit spreads. In much of the literature, crisis dating is binary, and variation within events that are dated as crises is left unstudied. An important contribution of our paper is to use credit spreads to understand the variation within crises. Romer and Romer (2014) take a narrative approach based on a reading of OECD accounts of financial crises to examine variation within crises. They also find that more intense crises are associated with slower recoveries. Our paper is also closely related to work on credit spreads and economic growth, most notably Gilchrist and Zakrajsek (2012) and Michael D. Bordo (2010). Relative to this work we study the behavior of spreads specifically in financial crises and study an international panel of bond price data as opposed to only US data. Our paper is also related to Giesecke *et al.* (2012) who study the knock-on effects of US corporate defaults and US banking crises, in a sample going back to 1860, and find that banking crises have significant spillover effects to the macroeconomy, while corporate defaults do not. We find that corporate bond spreads offering an indicator of the severity of crises, and taken with the evidence that the incidence of defaults do not correlate with the severity of downturns or with credit spreads (see Giesecke *et al.* (2011)), the data suggest that it is variation in default risk premia that may be driving our findings.

2 Data and Definitions

Crisis dates come from two sources: Reinhart and Rogoff (2009b) and Jorda *et al.* (2010) (henceforth RR and ST). The dates provided by Reinhart and Rogoff (2009b) are based on a major bank run or bank failure and contain the year the crisis itself began. The data from Jorda *et al.* (2010) instead dates the business cycle peak associated with the crisis. This may have occurred before or after the actual bank run or bank failure. However, there are appealing features of this dating convention. First, it is directly comparable to the exercise of comparing crisis outcomes to regular recessions or business cycle peaks. Second, it gives a clear guideline for choosing dates, rather than trying to understand when a systemic bank run occurred. Third, it guarantees we will measure a contraction in GDP going forward, which may not be the case when using the RR dates. In contrast the RR dates will tend to measure the more acute phase of the crisis. We show robustness to both sets of dates.

Our data on credit spreads come from a variety of sources. Table 1 details the data coverage. The bulk of our data covers a period from 1869 to 1929. We collect bond price, and other bond specific information (maturity, coupon, etc.), from the Investors Monthly Manual, a publication from the Economist, which contains detailed monthly data on individual corporate and sovereign bonds traded on the London Stock Exchange from 1860-1930. The foreign bonds in our sample include banks, sovereigns, and railroad bonds, among other corporations. The appendix describes this data source in more detail. We use this data to construct credit spreads, formed within country as high yield minus lower yield bonds. Lower yield bonds are meant to be safe bonds analogous to Aaa rated bonds. We select the cutoff for these bonds as the 10th percentile in yields in a given country and month. An alternative way to construct spreads is to use safe government debt as the benchmark. We find that our results are largely robust to using UK government debt as this alternative benchmark.[1] We form this spread for each country in each month and then average the spread over the last quarter of each year to obtain an annual spread measure.[2] This process helps to eliminate noise in our spread construction.

From 1930 onward, our data coverage is more sparse. We collect data, typically from central banks, on the US, Japan, South Korea, Hong Kong, and Sweden. These data include a number of crises, such as the Asian crisis, and the Nordic banking crisis. We also collect data on Ireland, Portugal, Spain and Greece over the period from 2000 to 2014 using bond data from Datastream, which covers the recent European crisis. We intend to expand our coverage as this research progresses. Our data appendix discusses the details and construction of this data extensively.

Finally, data on real per capita GDP are from Barro and Ursua (see Barro *et al.* (2011)). We examine the information content of spreads for the evolution of per capita GDP.

Figure 1 plots the incidence of crises, as dated by both RR and ST over our sample (i.e. the intersection of their sample and ours that contain data on bond spreads). The figure reveals that our coverage is dense in the late 19th century, and more sparse since.

3 Spreads, Output, and Normalization

We examine the relation between credit spreads and the evolution of GDP using variations on the following specification:

$$\ln(y_{i,t+k}/y_{i,t}) = a_i + b \times s_{i,t} + c' x_{i,t} + \varepsilon_{i,t+k}. \tag{1}$$

This is a panel data regression, where i indexes countries, t indexes time, and k indexes the forecasting horizon. We regress output growth at horizon-$k(\ln(y_{i,t+k}/y_{i,t}))$ on a country dummy, a spread measure for country-i at date-t $(s_{i,t})$ which we discuss further below, and a vector of country-time controls $(x_{i,t})$.

There is a large literature examining the forecasting power of credit spreads for economic activity (see Friedman and Kuttner (1992), Gertler and Lown (1999), Philippon (2009),

[1] One issue with UK government debt is that it does not appear to serve as an appropriate riskless benchmark during the period surrounding World War I as government yields rose substantially in this period.

[2] We use the average over the last quarter rather than simply the December value to have more observations for each country and year. Our results are robust to averaging over all months in a given year but we prefer the 4th quarter measure as our goal is to get a current signal of spreads at the end of each year.

and Gilchrist and Zakrajsek (2012)). Credit spreads reflect the probability of a borrower defaulting ($P_{default}$), the loss given default (LGD), and the economic risk premium that investors attach to suffering losses at times when marginal utility may be high ($\frac{Q_{default}}{P_{default}}$, where Q refers to Arrow-Debreu prices):

$$spread_{i,t} = \overbrace{P_{default} \;\; LGD}^{\text{Expected default}} \;\times\; \overbrace{\frac{Q_{default}}{P_{default}}}^{\text{Risk premium}} \tag{2}$$

Both the expected default component and the risk premium component will rise when output is expected to be low, i.e. $\ln(y_{i,t+k}/y_{i,t})$ is expected to be low, which underlies the predictive power of spreads for economic activity, and motivates the regression (1).

Algebraically, suppose growth is,

$$\ln(y_{i,t+k}/y_{i,t}) = A_i + B \times z_{i,t} + C'x_{i,t} + \epsilon_{i,t+k}. \tag{3}$$

and, spreads are,

$$Spread_{i,t} = \alpha_i + \beta_i z_{i,t} + u_{i,t}. \tag{4}$$

Then, spreads offer a measure of expected growth ($z_{i,t}$) which will be useful in forecasting actual growth.

In much of our analysis, we use spreads simply as a metric for the severity of a financial crisis. Thus it does not matter which component of spreads forecasts economic activity. On the other hand, some of our results are consistent with risk premia being an important component in forecasting crises. These results are consistent with Gilchrist and Zakrajsek (2012), who provide evidence that the informative component of spreads for future output is the default risk premium component rather than the expected default component. There is also a theoretical literature based on financial frictions in the intermediation sector, which draws a causal relation between increases in credit spreads and future economic activity (see He and Krishnamurthy (2012)). For most of our analysis, we do not take a stand on whether or not the relation between spreads and activity reflects causation or correlation.

Table 2 examines the forecasting power of spreads for 1-year output growth in our sample. We run,

$$\ln(y_{i,t+1}/y_{i,t}) = a_i + a_t + b_0 \times spread_{i,t} + b_1 \times spread_{i,t-1} + \varepsilon_{i,t+k}. \tag{5}$$

That is, we examine the forecasting power of spreads and lagged spreads for 1-year output growth. The first and second columns contain these results for the raw spread. The first column only includes country fixed effects, while the second column includes country and time fixed effects. We report coefficients and standard errors, clustered by year, in parentheses.

Columns (1) and (2) show that spreads do not forecast well in our sample. But there is a simple reason for this failing. Across countries, our spreads measure differing amounts of credit risk, so that the coefficients β_i in equation (4) differs across country. For example, in US data, we would not expect that Baa-Aaa spread and Ccc-Aaa spread contain the same information for output growth, which is what is required in running (1) and holding b constant across countries. In the 2008-2009 Great Recession in the US, high yield spreads rose much more than investment grade spreads. It is necessary to normalize the spreads in some way so that the spreads from each country contain similar information. We try a variety of approaches.

In, columns (3) and (4), we normalize spreads by dividing by the average spread for that country. That is, for each country we construct:

$$\hat{s}_{i,t} \equiv Spread_{i,t}/\overline{Spread^i} \tag{6}$$

A junk spread is on average higher than an investment grade spread, and its sensitivity to the business cycle is also higher. By normalizing by the mean country spread we assume that the sensitivity of the spread to the cycle is proportional to the average spread. Formally, if we assume that,

$$\textbf{A1:}\ \ \beta^i = \phi\alpha^i \tag{7}$$

then,

$$\hat{s}_{i,t} = 1 + \phi z_{i,t} + \frac{1}{\alpha^i} u_{i,t} \tag{8}$$

so that the normalized spread has a loading of ϕ on $z_{i,t}$ that is constant across all countries. In this case, the regression (1) recovers $b = \frac{1}{\phi}\beta$.

The results in columns (3) and (4) show that this normalization considerably improves the forecasting power of spreads. Both the R^2 of the regression and the t-statistic of the estimates rise.

The rest of the columns report other normalizations. In columns (5) and (6) we report results from normalizing by the median spread. In columnns (7) and (8), we report results from converting the spread into a Z-score for a given country, while in columns (9) and (10) we convert the spread into its percentile in the distribution of spreads for that country. All of these approaches do better than the non-normalized spread, both in terms of the R^2 and the t-statistics in the regressions. But none of them does measurably better than the mean normalization. We will focus on the mean normalization in the rest of the paper – a variable we refer to as $\hat{s}_{i,t}$. Our results are broadly similar when using other normalizations.

4 Spreads and Crises

4.1 Variation within crises

Reinhart and Rogoff in their research emphasize that recessions accompanied by financial crises are particularly severe. Across a select sample of banking crises, Reinhart and Rogoff (2009a) report a mean peak-to-trough decline in output of 9.3%.

However, there is enormous variation in outcomes across the literature's defined financial crises. Figure 2 illustrates this point. We focus on crisis dates identified by ST and plot histograms of different output measures across the crisis dates. We use use three measures of severity of a crisis. The first is to use the standard peak to trough decline in GDP locally as the last consecutive year of negative GDP growth after the crisis has started. The second is to look for the minimum value of GDP in a 10 year window after the crisis begins. The second definition allows for the possibility of a double dip, which may be important as crises tend to have slow recoveries. We call this measure the "flexible depth" measure.[3] The last measure of severity is simply the 5 year cumulative growth in GDP after a crisis has occurred. We choose 5 years to again account for persistent negative effects to GDP after crises. The 5 year growth rate will also capture experiences were growth is low relative to trend but not necessarily persistently negative (i.e., Japan in 1990). Our other measures will not pick up these effects.

Focusing on the peak-to-trough decline, in the upper left panel of the figure, we see that there is considerable variation within crises. Moreover, we see that the distribution is left-skewed. The top panel of Table 8 provides statistics on the variation for the ST dates. The mean peak-to-trough decline is -6.2%, but the standard deviation is 7.3%. The median is -3.9%, which is smaller in magnitude than the mean, indicating that the distribution is left-skewed. The table also reports statistics for the RR dates. The declines are smaller under RR's dating convention because the declines are measured from the date of the crisis to the trough rather than from the previous peak. But we see the same general pattern of enormous variation and a left-skewed distribution.

[3] Note that when using the ST dates both of these values will never be zero as by definition GDP contracts the following year. In contrast, using the RR dates this value may be zero if the crisis occurs after the trough.

4.2 Spreads as a measure of the severity of crises

The extent of variation within crises is in large part due to the convention of dating an episode a "crisis" or "non-crisis." With this binary approach, different crises with varying severity are grouped together. We can do better in understanding crises with a more continuous measure of the severity of crises. Romer and Romer (2014) pursue such an approach based on narrative assessments of the health of countries' financial systems. They describe financial stress using an index that takes on integer values from zero to 15, and show that this index offers guidance in forecasting the evolution of GDP over a crisis. We follow the Romer-Romer approach, but use credit spreads in the first year of a crisis to index the severity of the crisis. Relative to the Romer-Romer approach, credit spreads have the advantage that they are market-based. In addition, since they are based on asset prices they are automatically forward-looking indicators of economic outcomes.

Table 3 presents regressions of credit spreads on the peak-to-trough decline in GDP, as a measure of the severity of crises. Each data point in these regressions is a crisis in a given country-year (i, t), where crises are defined either using the ST or the RR chronology.

$$decline_{i,t} = a + b \times \hat{s}_{i,t} + \varepsilon_{i,t} \tag{9}$$

It is important to emphasize that the regression relates cross-sectional variation in spreads and the measure of severity. The average severity of crises is absorbed into the constant. Other papers, such as Reinhart and Rogoff, focus on the average severity in crises (i.e. the -9.3% decline cited above from Reinhart and Rogoff (2009a)). In this regard, our research adds new information relative to the existing literature.

The spread has statistically and economically significant explanatory power for crisis severity. A one-sigma change in $\hat{s}_{i,t}$ of 1.2% translates to a 2.64% decrease in peak-to-trough GDP using the ST dates and a 1.7% decrease using the RR dates. The spreads also meaningfully capture variation in crisis severity, meaning it is a salient approach to index crises. The standard deviation of the peak-to-trough decline in GDP for the ST dates is 8.0%. The variation that the spread variable captures is 4.3% for the ST crises.

The bottom panel of the table presents results where we include lagged spreads, $\hat{s}_{i,t-1}$ and leverage growth (Δlev_t, change in the ratio of total loans to money supply) from Jorda *et al.* (2010) which is known to be a predictor of financial crises. The sample shrinks in these regressions because the ST variable is not available for all of our main sample.

We note that the explanatory power increases measurably when including these other variables. Comparing columns (3) and (4) corresponding to the ST crises, the variation that is picked up by the independent variables rises from 4.8% of GDP to 6.0% of GDP. The increase in explanatory power is only present for the crises and not for the recession dates. Second, we see that the lagged spread has a positive and significant sign for the crisis dates (not for the recession date), indicating that the change in the spread from the prior year is more indicative of the severity of the recession. In fact, the autocorrelation of spreads is about 0.70 in our sample, which is also roughly the ratio of the coefficients on $\hat{s}_{i,t-1}$ and $\hat{s}_{i,t}$, indicating a special role for the innovation in spreads.

Why does the innovation in spreads matter? One may expect that since the level of spreads embody expectations of future default probabilities and loss-given-default, that it should only be the level of spreads that should forecast output, as in equation (4). We think that this effect arises through the rise in the risk premium component of spreads (see equation (2)). The finance literature documents a relation between jumps, volatility and risk premia (see Andersen *et al.* (2014)). It is thus likely that a large surprise innovation in spreads is correlated with an increase in the risk premium component of spreads. Since the forecasting power of spreads, as in (2), comes from the separate relations between expected default and output growth as well as risk premium and output growth, the importance of the lagged spreads in the crisis dates indicates a role for the risk premium.[4] Indeed, in focusing on the recession regression in column (6), we see that the lag has little explanatory power, and it is likely that in these episodes there is little rise in volatility and risk premia.

Finally, we see that the leverage growth variable from ST has independent explanatory power for the severity of the crisis. If we repeat the regression in columns (2) and (4) of the bottom panel, dropping $\hat{s}_t$ and only including Δlev_t we find that the coefficients are quite close to the regression coefficients in the regression with spreads. That is, spreads and credit growth have independent forecasting power for crises. This result is similar to Greenwood and Hanson (2013) who find that a quantity variable that measures the credit quality of corporate debt issuers deteriorates during credit booms, and that this deterioration forecasts low excess returns on corporate bonds even after controlling for credit spreads. Our finding confirms the Greenwood and Hanson result in a much larger cross-country sample.

[4] Note the current spread also reflects the risk premium. So the statement is that the risk premium has different forecasting power for economic growth than expected default, which is why including the risk premium (proxied for as a large rise in spreads) has forecasting power for growth above that embodied in the level of the spread. For example, if a 1% rise in the credit spread due to a rise in the risk premium and a 1% rise in the credit spread due to a rise in the expected default are both associated with an equal fall in output growth, then it is not possible to separately identify a role for risk premia and expected default.

4.3 Spreads and the evolution of output

Table 4 and 5 presents regressions where we use all of the data in panel regressions. We regress future GDP growth at the 5-year (in Table 4, and 3-year in Table 5) horizon on current $\hat{s}_{i,t}$, including a country dummy that absorbs mean differences in country growth rates. We also include the lagged value of the spread, two lags of GDP growth, and leverage growth from ST. Note that the independent variable can now reflect the innovation in spreads. The regression specification in column (1) of the tables is,

$$\ln\left(y_{i,t+k}/y_{i,t}\right) = a_i + b_0 \times \hat{s}_{i,t} + b_{-1} \times \hat{s}_{i,t-1} + \sum_{j=0}^{1} c_j \times \Delta \ln(y_{i,t-j}) + d \times \Delta lev_t + \varepsilon_{i,t+k} \quad (10)$$

This regression pools both crises and non-crises together and indicates that there is a relation between spreads and subsequent GDP growth, consistent with results from the existing literature (see, for example, Gilchrist and Zakrajsek (2012)).

Columns (2)-(5) allow the coefficient on spreads to vary across crises and non-crises (or recessions and non-recessions). That is, we run,

$$\begin{aligned} \ln\left(y_{i,t+k}/y_{i,t}\right) &= a_i + 1_{crisis}\left(a^{crisis} + b_0^{crisis} \times \hat{s}_{i,t} + b_{-1}^{crisis} \times \hat{s}_{i,t-1}\right) \quad (11) \\ &\quad + 1_{no-crisis}\left(a^{no-crisis} + b_0^{no-crisis} \times \hat{s}_{i,t} + b_{-1}^{no-crisis} \times \hat{s}_{i,t-1}\right) + c' x_t + \varepsilon_{i,t+k} \end{aligned}$$

with controls of lagged GDP growth and leverage growth, as in (10).

The results are in line with our findings in Table 3. High current spreads forecast more severe crises. The lagged spread comes in with a positive coefficient that is significantly different than zero in many of the specifications. Thus, again the innovation in spreads plays an important role in explaining the severity of crises. The effects are statistically stronger at the 3-year horizon as reported in Table 5.

Figure 3 plots the evolution of GDP to a 100 basis points shock to spreads, conditional on ST crises (top panel) and RR crises (lower panel). Focusing on the ST dates, we see that output falls, reaching a low at the 4-year horizon of -6% before recovering. The results for the RR dates are similar, although smaller in magnitude, which is likely due to the fact that RR date the crises typically later than ST.

The impulse responses in Figure 3 are computed by forecasting GDP individually at all horizons from 1 to 5 years using the local projection methods in Jordà (2005) (see also Romer and Romer, 2014). That is, we estimate (11) for $k = 1, ..., 5$ and use the individual coefficients on spreads to trace out the effect on output given a 100 basis point shock to our normalized spreads. Thus the plot in Figure 3 is the difference in output paths for two financial crises, one of which has a 100 basis point higher spread. We use the Jorda methodology rather than imposing more structure as in a VAR as it is more flexible and does not require us to specify the dynamics of all variables.

Table 4 and 5 also reveal that the coefficient on spreads in crises is larger in magnitude than the coefficient outside crises (which is near -0.9 as in the full sample regression, and which we omit to save space). We discuss this result further in the next section in order to compare recoveries from financial crises to recessions.[5]

4.4 2008 crisis and recovery

Reinhart and Rogoff (2009a)'s mean estimate of -9.3% peak-to-trough decline in GDP in financial crises has been taken as the benchmark to compare the experience of the US after the 2008 financial crisis. We can provide a different benchmark based on our approach of examining the cross-sectional variation in crisis severity.

Figure 4 top-panel plots the actual and predicted path of output for the 2008-2013 period based on the spread in the last quarter of 2008. The lower panel plots the actual and predicted path of spreads for the 2008-2013 using the (11) with spread as dependent variable. The actual and predicted output paths are remarkably similar, indicating that at least for this crisis, what transpired is exactly what should have been expected. The result supports Reinhart and Rogoff (2009a)'s conclusion that the recoveries from financial crises are protracted. Our forecast path is not purely from the historical average decline across crises as in Reinhart and Rogoff, but is also informed by the historical cross-section of crises severity and the spread in 2008.

We also note that the actual reduction in spreads is faster than the reduction that would have been predicted by our regressions, while GDP growth is faster than predicted. That is, the residuals from the forecasting regressions are negatively correlated. This result could be interpreted to mean that the aggressive policy response in the recent crisis allowed for a better outcome than historical crises. Many of the historical crises in our sample come from a period with limited policy response.

[5] Note that it is tempting to read the higher coefficients associated with crisis observations as evidence of non-linearity, as suggested by theoretical models such as He and Krishnamurthy (2014). However this is not correct. In He and Krishnamurthy, *both* the spread and the path of output are a non-linear function of an underlying financial stress state variable. It is not the case that output is a non-linear function of spreads, but rather that both are non-linear functions of a third variable. Since we regress output on spreads, rather than either stress or output on an underlying financial shock, the regressions need not be evidence of non-linearity.

4.5 Forecasting future spreads

Table 6 presents regressions where we forecast $\hat{s}_{i,t+1}$ using current and lagged values of variables, as in (11). Since a rise in spreads indicates a decline in future output, this regression asks whether ex-ante variables can forecast a rise in spreads. The most interesting result from this table is that leverage growth forecasts a rise in spreads, confirming ST results that leverage growth is an ex-ante measure of the likelihood of crises. Lopez-Salido, Stein, and Zakrajsek (2014) report a related finding in US data, suggesting that high leverage is part of a pattern of an overheating boom/bust cycle.

5 Spread-based crises

The methods of dating crises of ST, RR and Romer and Romer (2014) are subject to a potential bias. Since financial crises are dated with knowledge of eventual GDP outcomes, it is possible that only financial events that result in large declines in output are labeled crises, so that we infer that all financial crises lead to large declines in output. This bias will affect the predicted path of output in Figure 3.

5.1 Skewness in output

Figure 6 plots the distribution of GDP growth at the 1-year and 5-year horizons based on a kernel density estimation. The blue line plots the distribution of GDP growth when spreads are in the lower 30% of their realizations, while the red-dashed line plots the distribution when spreads are in the highest 30% of their realizations. A comparison of the blue to red lines indicates that high spreads shifts the conditional distribution of output growth to the left, with a fattening of the left tail.

Table 7 presents quantile regressions of output growth on $\hat{s}_{i,t}$ and $\hat{s}_{i,t-1}$. We see that the forecasting power of spreads for output increases as we move to the lower quantiles of the output distribution. At the median, the coefficient on $\hat{s}_t$ is -0.64 (and is $+0.62$ on the lag), while it is -1.07 (and 0.83 on the lag) at the 25th quantile.

Figure 5 plots the impulse response of different moments of GDP to a 100 basis point shock to spreads. We see that the median response is smaller than the mean response, indicating that high spreads are associated with skewness. The 10th percentile shows a dramatic reduction in output, roughly twice the size of the mean of the response, with the trough of output occurring about 4 years after the shock. These results suggest that a spike in spreads increases the likelihood that the economy will suffer a deep and protracted slump.

		Macro (GDP)	
		No Drop	Large Drop
Risk Premium (Spread)	No Rise	Normal periods	Wars, recessions, natural disasters, etc.
	Large Rise	Crises which did not realize: Penn Central, LTCM etc.	“Systemic crises": Laevan-Valencia, Reinhart-Rogoff Great Depression, 2008, etc.

The above diagram helps to explain these effects. A spike in spreads, associated with a financial disturbance, results in one of two outcomes: either there is pass-through to the real economy, or the financial shock dissipates quickly. The Penn Central crisis of 1970 is a financial disturbance, but either through the timely intervention of the Federal Reserve or through luck, the disturbance fades and there is little impact on the real sector. In contrast, in 2008-2009, there is a financial disturbance that despite intervention by the government (or bad luck) results in a protracted recession.

The diagram also helps to shed light on RR and ST's crisis dating. Neither RR nor ST date events in the lower left box of the table as crises, while if one takes the spread perspective, then these events were crises that were somehow avoided. The 1970 Penn Central crisis is an example of a crisis event that RR/ST do not date as a crisis. Thus, one way of looking at RR/ST is that they tell us about outcomes in the lower quantile of the output distribution. That is, they only focus on events in the lower right box of the above diagram.

5.2 Dating crises based on spreads

We can avoid the ex-post bias by dating crises based on spreads. We formulate a new definition of crises as follows. We define a crisis if,

$$\hat{s}_{i,t} - \hat{s}_{i,t-1} \geq \bar{\Delta}. \tag{12}$$

We have seen that information in forecasting GDP evolution during a crisis is contained in the innovation in spreads. Our definition follows this logic and looks for an innovation above a threshold to define a "crisis." We choose $\bar{\Delta}$ so that the proportion of crises in our sample is similar to that of the ST and RR crises. There are 51 ST crises, 50 RR crises, and 56 spread crises in our sample.

Figure 6 provides a visual representation of how spread-dated crises overlap with the ST crises. There is considerable overlap in the dates, although there are many events that are labeled "spread crises" that are not ST crises. Table 6 provides statistics on the overlap between spread-based crises and the ST and RR crises.

The bottom panel of Table 8 presents summary statistics for the spread-based crises, allowing for a comparison to ST and RR crises. Crises are associated with large mean declines in output as with the RR and ST dating. The mean output decline is larger under ST than under spread-crises, indicating that there is likely a dating bias in RR and ST. Additionally, note that under the spread dating, the median is smaller while the standard deviation of output declines is larger. That is the conditional distribution of output growth is skewed towards low realizations, which we can also see by comparing the 10th percentile of outcomes across these dates. This further indicates that the RR and ST dating methods are trimming off a set of financial stress events that do not have significant ex-post effects on output.

5.3 Leverage and skewness

Why do some financial disturbances end up in the lower right box of the diagram, while others end up in the lower left box? Table 9 offers one answer. We forecast output growth based on our classification of spread-crises as well as the interaction between spread-crisis and ST's leverage variable. A shock to spreads that occurs after a ST leverage buildup leads to larger declines in output. This effect is present out to the 4-year horizon. The strongest effect is at the 3 year horizon, where we see that the coefficient on SpreadCrisis is statistically not different from zero, while the coefficient on the interaction term is large and statistically significant.

One way of understanding these results is in terms of the “triggers" and "vulnerabilities" dichotomy outlined in Bernanke (April 13 2012). Sometimes a financial disruption triggers losses within the financial sector which then leads to a rise in risk premia on intermediated assets (e.g., credit), as in the “intermediary asset pricing" model of He and Krishnamurthy (2012). This trigger has a large impact on the real sector only if it interacts with a real amplification mechanism. Jorda *et al.* (2010)'s leverage variable may measure the strength of the real amplification mechanism.

5.4 2008 crisis and recovery, V2

Figure 8 revisits the exercise of forecasting GDP growth and spreads for the 2008-2013 period based on the spread in the last quarter of 2008, but now using our new crisis dating based on spreads which is free of an ex-post dating bias. The upper panel plots the actual and predicted path of output for the 2008-2013 using specification (11). The actual is in black while the blue dashed lines are the forecast based on the ST dates, where we have seen earlier that output grows faster than forecast. The red line is based on spread-crisis dates. Now we see that predicted output is higher than actual output. This result is consistent with the existence of a bias in the ST dates. The average spread crisis forecast is based on an average over a crisis that dissipates and one that turns into a protracted recession as in the 2008-2013 period. The green-dot line presents results based on the specification of Table 9 where we condition on both spread-crises and leverage. Leverage was high prior to the 2008 crisis. The forecast exercise now results in predicted GDP that lies between ST and spread-crisis and similar to actual output. Thus, we again find that the recovery is slow and in keeping with patterns from past crises.

The lower panel presents results for the actual and predicted path of spreads. We consistently find that spreads in the recent crisis recovered faster than output.

6 Slow recoveries from financial crises

We now turn to comparing financial crises to non-financial crisis episodes. Cerra and Saxena (2008) and Claessens, Kose and Terrones (2010) document that recessions that accompany financial crises are deeper and more protracted than recessions that do not involve financial crises. They reach this conclusion by examining the average non-financial crisis recession to the average financial recession. Using spreads, we can offer a new estimate for recovery patterns.

6.1 A spread-based comparison

Suppose we are able to observe two episodes, one where a negative shock (z_t) leads to a deep recession but no financial disruption, and one where the same negative z_t shock lead to a financial disruption/crises and a deep recession. Then, the measured difference in long-term growth rates in these two episodes is the slow recovery that can be attributed to the financial crisis.

We try to measure this difference as follows. We have noted that crises are associated with high expected default and high risk premia, while recessions are only associated with high expected default. If we can compare the dynamics of GDP in two episodes with the same expected default, but in one of which there are also high risk premia, then the difference between in GDP dynamics across these two events is the pure effect of a financial crisis.

We use the coefficients in the spread regressions in Table 4/5 across crises and recessions to compute a long-run cost to growth. We consider a 100 basis point shock to the spread in different events, and trace out the impulse response of this shock for GDP using our different crisis and non-crisis events.

It is likely that this approach leads to an underestimate of the crisis effect. This is because the shock in a recession, $z_t^{recession}$ that leads a 100 basis point change in spreads is likely larger than the shock, z_t^{crisis}, that leads to a 100 basis point change in spreads. In the crisis, the shock z_t^{crisis} increase expected default and risk premia, while the same shock in recession likely largely increases expected default.

Figure 9 presents the results. The top panel presents results based on unconditional regressions, i.e., a regression pooling crises and non-crises dates. We see that output declines by about 1% 5 years after a shock to spreads. The lowest panel presents results for non-financial crisis recessions. Here also we see a decline of about 1% 5 years out, but the estimates are very imprecise. The middle panels presents results for our three dating conventions. We uniformly see larger and significant declines in GDP as far as 5 years out. The largest effects are with the ST dates (around 5% decline), while the RR dates and our spread-crisis dates provide similar results of around a 3% decline.

Our results affirm the findings of others that financial crises do result in deeper and more protracted recessions. We emphasize that we have reached this conclusion by examining the cross-section of countries rather than the mean decline across crises. Indeed the mean decline across crises plays no role in the impulse responses in Figure 9 because the plot is of the forecast GDP path in a crisis for a 100 basis point worse crisis (or recession). The mean decline across crises is differenced out, rendering the impulse response a "diff-diff" estimate.

Thus our results are new to the literature.

6.2 Long-lasting effects of financial disruptions

The dynamics of spreads and output in response to the spread shock are quite different, with spreads recovering more quickly than output. We can see this in the 2008 financial crisis, where the spread reverts to near normal levels within one year, while output remains depressed for considerably longer. This spread behavior is likely due to the dynamics of the risk premium in equation (2). We have noted that with the onset of a crisis, coincident with a jump in spreads, the risk premium component of spreads is likely high (as it was in 2008). This risk premium component reverts back more quickly than output. Muir (2014) presents evidence consistent with this observation using both spreads and dividend yields to measure risk premia. Muir notes that the spike-fall behavior of risk premia is special to financial crises and not to macroeconomic events such as wars where consumption drops dramatically. The sharp rise and fall an risk premia can be thought of as reflecting shocks to intermediary capital in the model of He and Krishnamurthy (2012).

The financial disruption implied by high spreads dissipates relatively quickly, yet output continues to be depressed. Table 10 and 11 presents regression results to underscore this observation. Consider the first three columns in the top panel. We run regressions where the dependent variable is output growth between years $t+1$ and $t+4$ (i.e. for 3 years), where t denotes the beginning of an ST crisis. The independent variable in the first column is the innovation in spreads from the year before the crisis to the year $t+1$ of the crisis. In the second column, the independent variable is the innovation in the spread from the year before the crisis tot the year t of the crisis. Comparing these columns, we see that the initial spike in spreads has as much statistical significance and explanatory power as the spread in year $t+1$ for forecasting output growth beyond year $t+1$. The third column presents a regression where the independent variable is the integral of the spread innovation over the first year of the crisis. This variable thus measures the duration and intensity of the spike in spreads at the start of a crisis. This variable also helps explains subsequent growth. The dependent variable in columns (4), (5), and (6) is output growth between years $t+2$ and $t+4$, while independent variables are now the initial innovation in spreads, the innovation to year $t+2$, and the integral of the innovation over the first 2 years of the crisis. Here we see that the initial shock and the integral of the innovation have significant explanatory power, but the spread innovation to year 2 of the crisis has little explanatory power. The last three columns of the table repeat to the exercise but using only one year of output growth, between years $t+3$ and $t+4$. The results are broadly similar but the t-statistics on spreads falls considerably because one-year of output growth, three years after a crisis is hard to predict. Expanding the window, say to forecast output from $t+3$ to $t+6$ does not help in this regard. The bottom panel of Table 7 presents results for the RR crises. The results are broadly in line with those for the ST crises. The initial spike in spreads and integral measure are better forecasters of future growth than even spreads in the years after the crisis. Table 11 presents results for the spread-crises dates. The statistical significance for these results are weaker, but in line with those presented in Table 10.

These results suggest that financial crises have long-lasting effects on the real sector, even after the financial sector is normalized. The experience of the 2008-2009 crisis and slow recovery is thus the rule and not the exception. Mian and Sufi (2014) have argued that the slow recovery pattern is due to the slow recovery of household balance sheets, and our evidence is consistent with their arguments, but with the nuance that the credit disruption in the initial phase of the crisis exacerbated the damage to household balance sheets.

7 Conclusion

This paper studies the behavior of credit spreads and their link to economic growth during financial crises. Our main finding is that the recessions that surround financial crises are longer and deeper than the recessions surrounding non-financial crises. The slow recovery from the 2008 crisis is in keeping with historical patterns surrounding financial crises. We have reached this conclusion by examining the cross-sectional variation between credit spreads and crisis outcomes rather computing the average GDP performance for a set of specified crisis dates. This is the main contribution of the paper.

These results are in keeping with the conclusions of Reinhart and Rogoff (2009a) but different than Romer and Romer (2014) who approach crises using a cross-sectional approach, as we have. The differences between our paper and the Romer and Romer paper is that our sample largely consists of crises over the 1870-1930 period, while they focus on crises in the last 50 years in OECD countries. Additionally, we use credit spreads to index the severity of crises rather than their narrative approach. We hope that future versions of this paper will include more spreads and crises from the last 50 years so that we can better understand these differences.

Researchers have begun to construct macroeconomic models that generate slow recoveries from financial crises. Our study offers more facts for constructed models to match. First, we provide quantitative guidance on the speed of recoveries after financial crises. Second, we find that risk premia spike in crises and revert back to normal more quickly than GDP reverts back to normal. Third, we find that many of these spikes, especially if they occur when leverage has been low, lead to little effects on GDP.

8 Data Appendix

Credit spreads from 1860-1930. Source: Investor's Monthly Manual (IMM) which publishes a consistent widely covered set of bonds from the London Stock Exchange covering over 60 countries. We take published bond prices, face values, and coupons and convert to yields. Maturity or redemption date is typically included in the bond's name and we use this as the primary way to back out maturity. If we can not define maturity in this way, we instead look for the last date at which the bond was listed in our dataset. Since bonds almost always appear every month this gives an alternative way to roughly capture maturity. We check that the average maturity we get using this calculation almost exactly matches the year of maturity in the cases where we have both pieces of information. In the case where the last available date is the last year of our dataset, we set the maturity of the bond so that its inverse maturity (1/n) is equal to the average inverse maturity of the bonds in the rest of the sample. We equalize average inverse maturity, rather than average maturity, because this results in less bias when computing yields. To see why note that a zero coupon yield for a bond with face value \$1 and price p is $-\frac{1}{n}\ln p$. Many of our bonds are callable and this will have an effect on the implied maturity we estimate. Our empirical design is to use the full cross-section of bonds and average across these for each country which helps reduce noise in our procedure, especially because we have a large number of bonds. For this reason, we also require a minimum of 20 bonds for a given country in a given year for an observation to be included in our sample.

US spread from 1930-2014. Source: Moody's Baa-Aaa spread.

Japan spread from 1989-2001. Source: Bank of Japan.

South Korea spread from 1995-2013. Source: Bank of Korea. AA- rated corporate bonds, 3 year maturity.

Sweden spread from 1987-2013. Source: Bank of Sweden. Bank loan spread to non-financial Swedish firms, maturities are 6 month on average.

Hong Kong 1996-2012. Source: .

European spreads (Ireland, Portugal, Spain, Greece) from 2000-2014. Source: Datastream. We take individual yields and create a spread in a similar manner to our historical IMM dataset.

GDP data. Source: Barro and Ursua (see Robert Barro's website). Real, annual per capital GDP at the country level. GDP data for Hong Kong follows the construction of Barro Ursua using data from the WDI.

Crisis dates. Source: Jorda, Schularick, and Taylor / Schularick and Taylor ("ST" dates), Reinhart and Rogoff (“RR” dates, see Kenneth Rogoff’s website).

Leverage, Credit to GDP data. Source: Schularick and Taylor.

References

Torben G. Andersen, Victor Todorov, and Nicola Fusari. The risk premia embedded in index options. *working paper*, 2014.

Robert Barro, Emi Nakamura, Jon Steinsson, and Jose Ursua. Crises and recoveries in an empirical model of consumption disasters. *working paper*, 2011.

Ben S. Bernanke. Some reflections on the crisis and the policy response. *At the Russell Sage Foundation and The Century Foundation Conference on "Rethinking Finance," New York, New York*, April 13, 2012.

Michael D. Bordo and Joseph G. Haubrich. Deep recessions, fast recoveries, and financial crises: Evidence from the american record. *NBER working paper*, 2012.

Valerie Cerra and Sweta Chaman Saxena. Growth dynamics: The myth of economic recovery. *The American Economic Review*, 98(1):439–457, 2008.

Benjamin M. Friedman and Kenneth N. Kuttner. Money, income, prices, and interest rates. *The American Economic Review*, 82(3):pp. 472–492, 1992.

M Gertler and CS Lown. The information in the high-yield bond spread for the business cycle: evidence and some implications. *Oxford Review of Economic Policy*, 15(3):132–150, 1999.

Kay Giesecke, Francis A. Longstaff, Stephen Schaefer, and Ilya Strebulaev. Corporate bond default risk: A 150-year perspective. *Journal of Financial Economics*, 102(2):233 – 250, 2011.

Kay Giesecke, Francis A. Longstaff, Stephen Schaefer, and Ilya Strebulaev. Macroeconomic effects of corporate bond default crises: A 150-year perspective. *working paper*, 2012.

Simon Gilchrist and Egon Zakrajsek. Credit spreads and business cycle fluctuations. *The American Economic Review*, forthcoming, 2012.

Zhiguo He and Arvind Krishnamurthy. Intermediary asset pricing. *The American Economic Review*, forthcoming, 2012.

Oscar Jorda, Moritz Schularick, and Alan M. Taylor. Financial crises, credit booms, and external imbalances: 140 years of lessons. *NBER working paper*, 2010.

Òscar Jordà. Estimation and inference of impulse responses by local projections. *The American Economic Review*, 95(1):pp. 161–182, 2005.

Atif Mian and Amir Sufi. House of debt: How they (and you) caused the great recession, and how we can prevent it from happening again. *University of Chicago Press*, 2014.

Joseph G. Haubrich Michael D. Bordo. Credit crises, money and contractions: An historical view. *Journal of Monetary Economics*, 57(1):1 – 18, 2010.

Thomas Philippon. The bond market's q. *The Quarterly Journal of Economics*, 124(3):1011–1056, 2009.

Carmen M. Reinhart and Kenneth S. Rogoff. The aftermath of financial crises. *American Economic Review*, 99(2):466–72, 2009.

Carmen M. Reinhart and Kenneth S. Rogoff. *This time is different: Eight centuries of financial folly.* Princeton University Press, Princeton, NJ, 2009.

Christina D. Romer and David H. Romer. New evidence on the impact of financial crises in advanced countries. *working paper*, 2014.

9 Tables and Figures

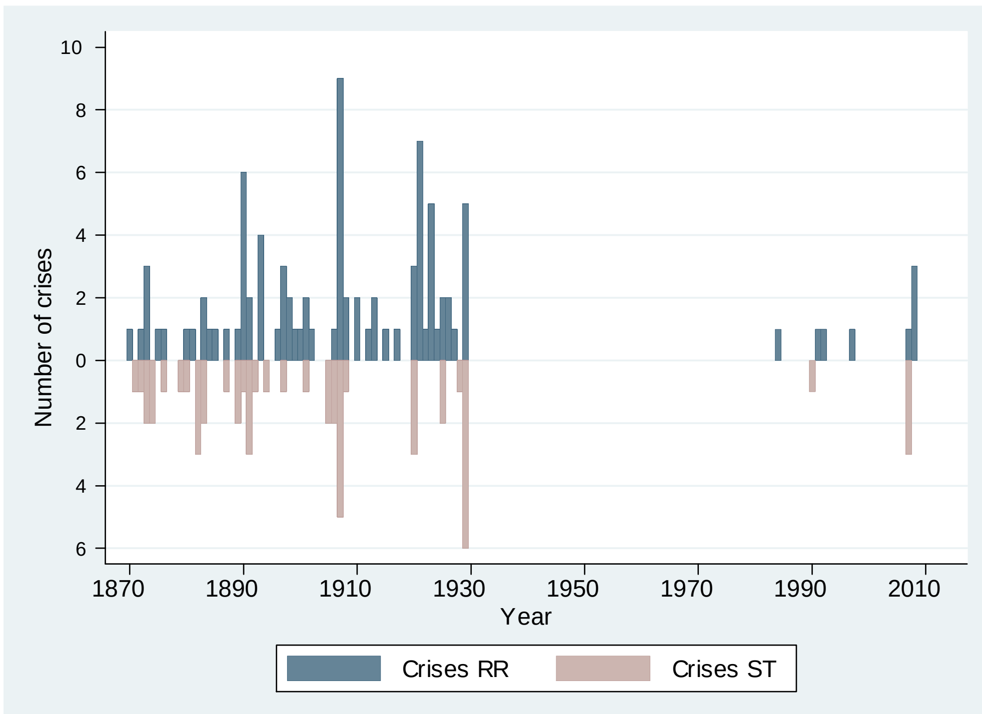

Figure 1: This figure plots the incidence of crises over time across various countries from 1870-2008. RR denotes those measured by Reinhart and Rogoff and ST denotes those measured by Schularick and Taylor. We only plot these variables for countries and dates for which we have credit spread data to give a sense of the crises covered by our data.

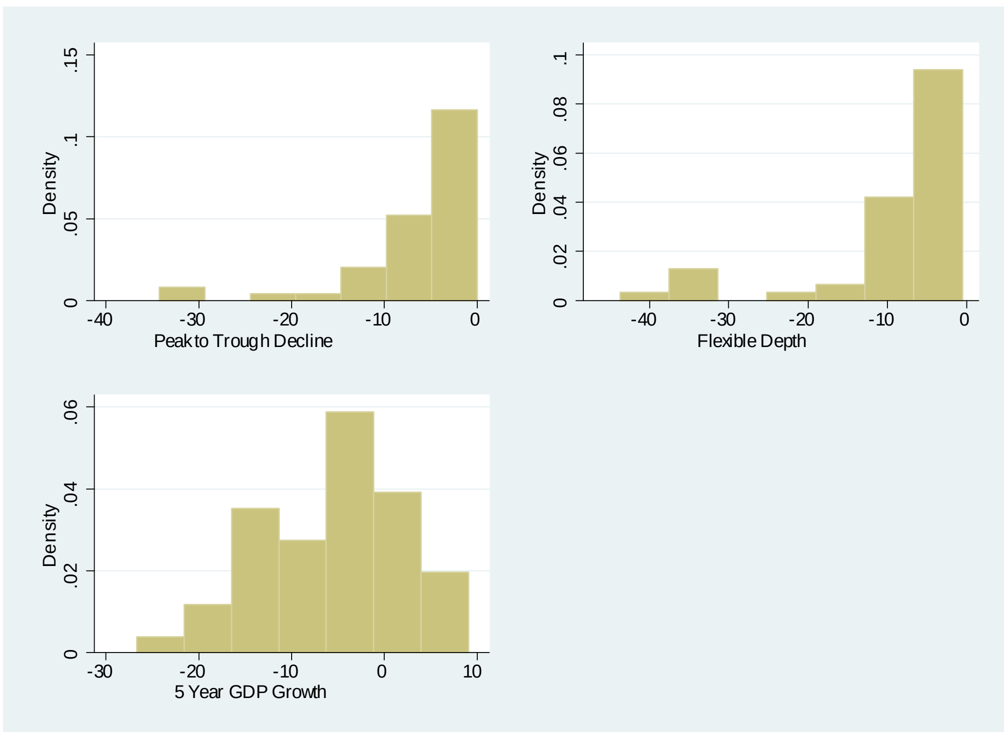

Figure 2: Distribution of outcomes in GDP across financial crises. See text for descriptions.

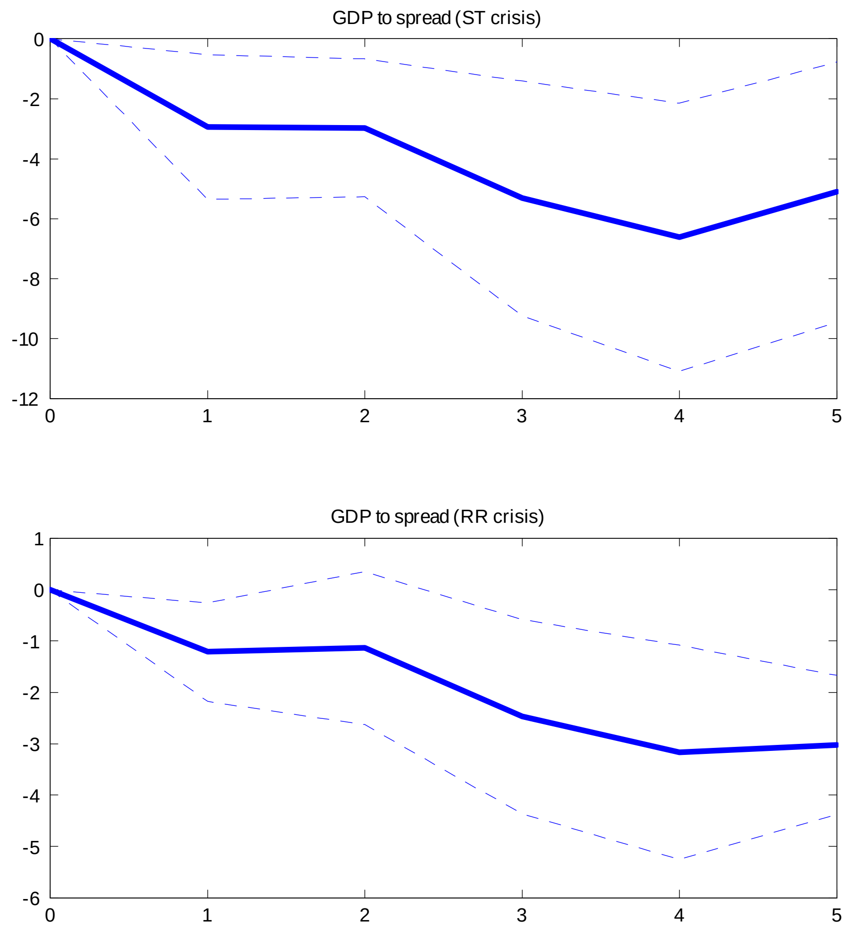

Figure 3: This figure plots the impulse responses of GDP and normalized spreads to an innovation in our spread measure. We show this unconditionally in the top panel (labeled normal times) as well as during crises in the lower panel. Impulse responses are computed using local projection measures where we forecast GDP independently at each horizon.

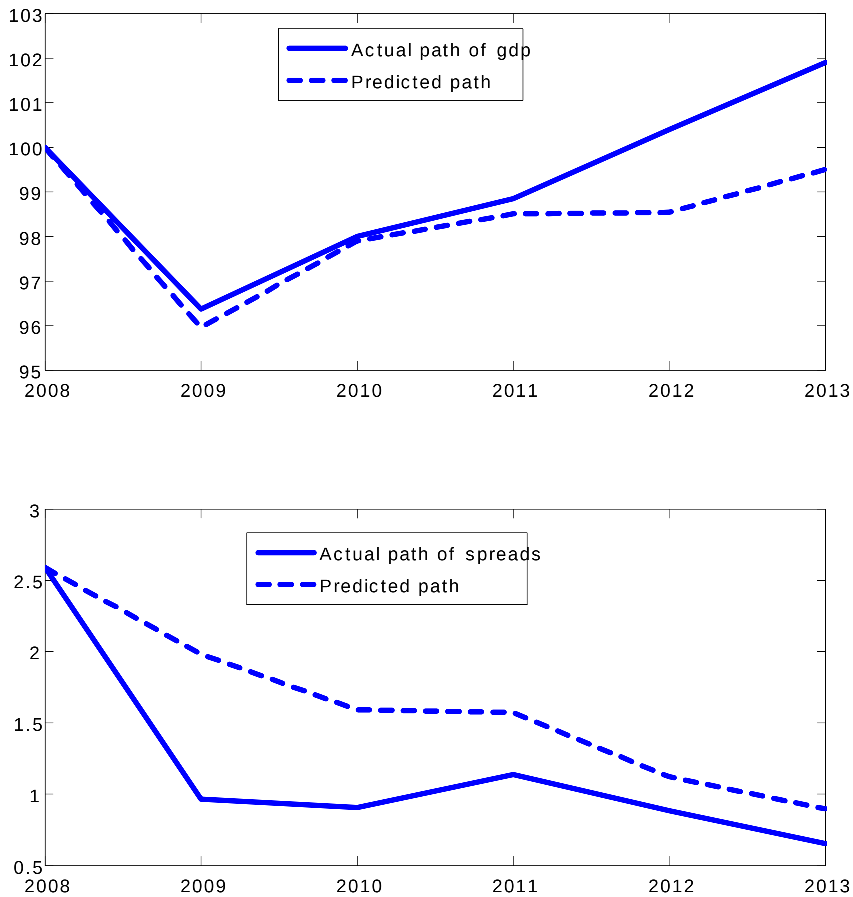

Figure 4: We predict outcomes of output and spreads during the 2008 US financial crisis using predicted values from our regressions and data up to 2008. The top panel, GDP, is cumulative from a base of 100 in 2008. The lower panel, spreads, uses the last quarter value of the BaaAaa spread in 2008.

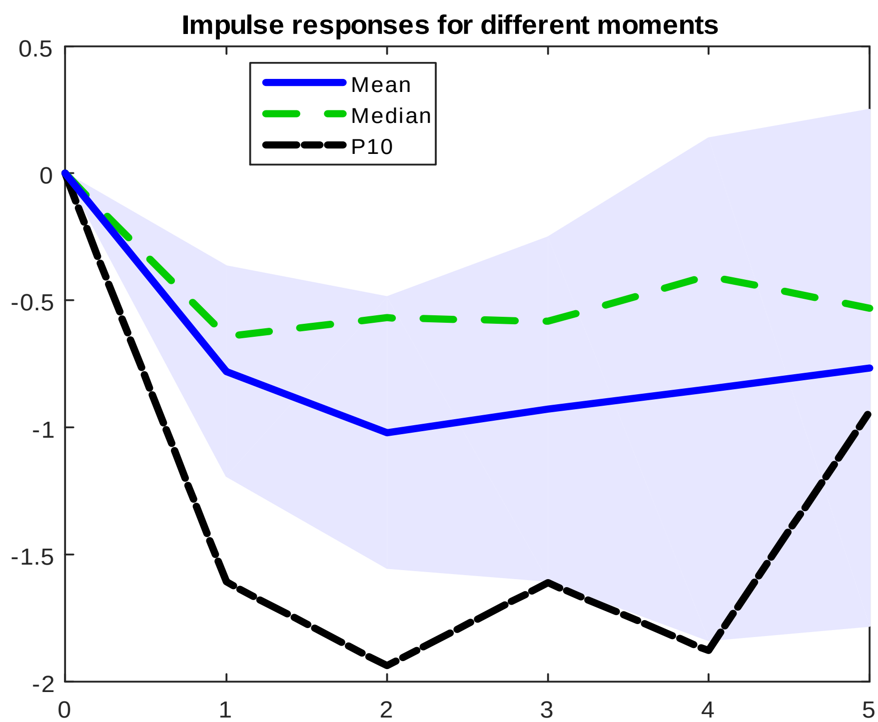

Figure 5: We plot impulse responses of GDP to spreads for various moments: the mean, median, and 10th percentile. All impulse responses use the Jorda local projection method where we use quantile regression or OLS depending on the moment plotted. 95% confidence intervals are given in colored shaded regions.

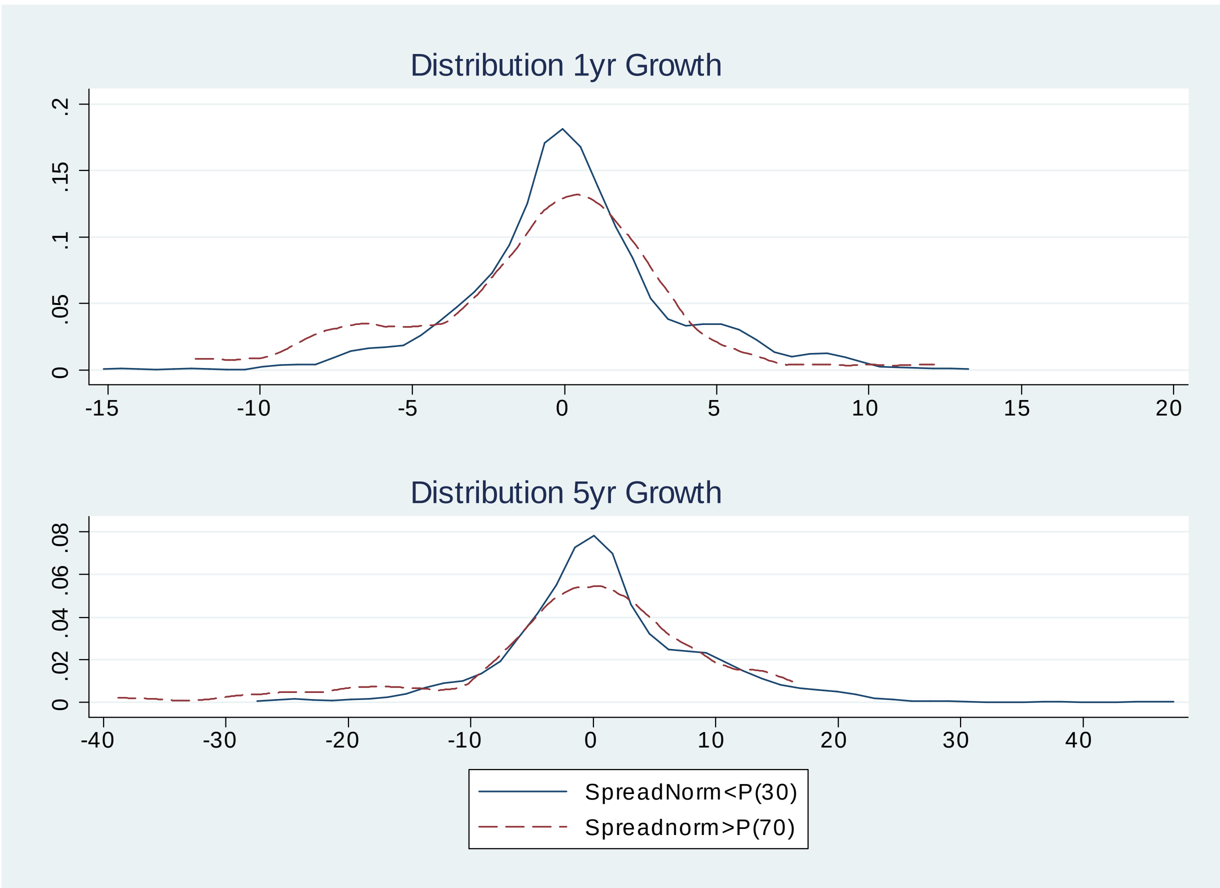

Figure 6: This figure plots the distribution of GDP growth at various horizons conditional on spreads based on a kernel density estimation. The blue solid line plots the distribution of GDP growth when spreads are in the lower 30% of their realizations, the red dashed line plots the distribution when spreads are in the highest 30% of their realizations.

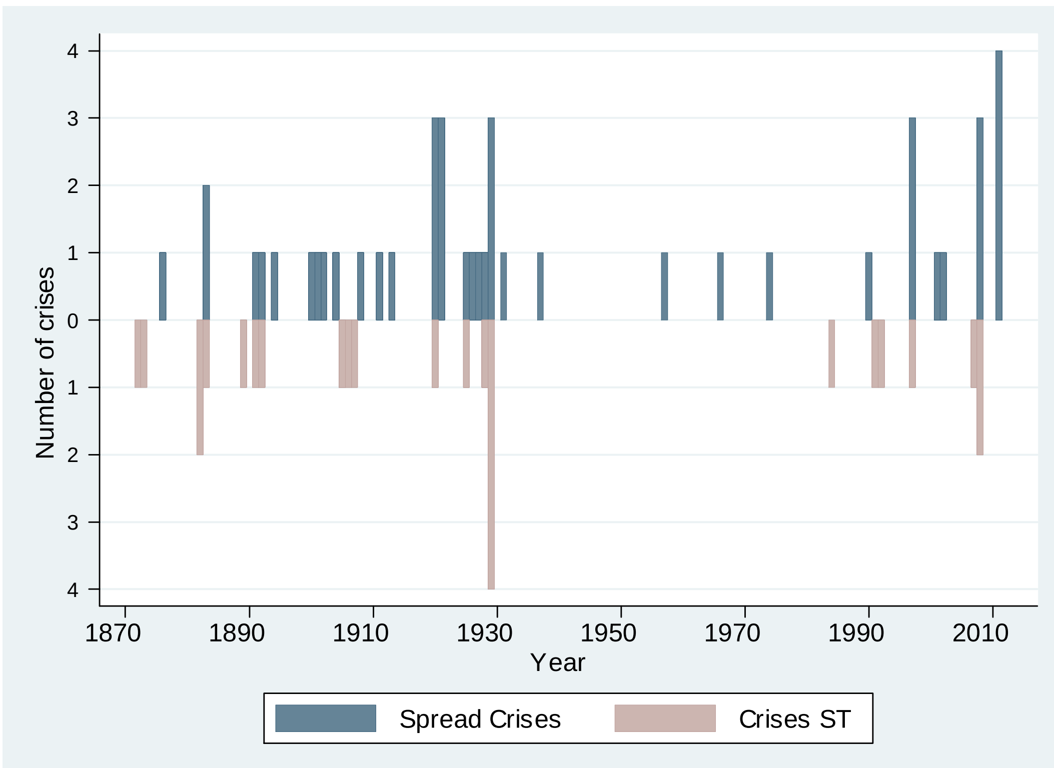

Figure 7: We plot our spread crises counts by year along with counts from Schularick and Taylor for crisis dates.

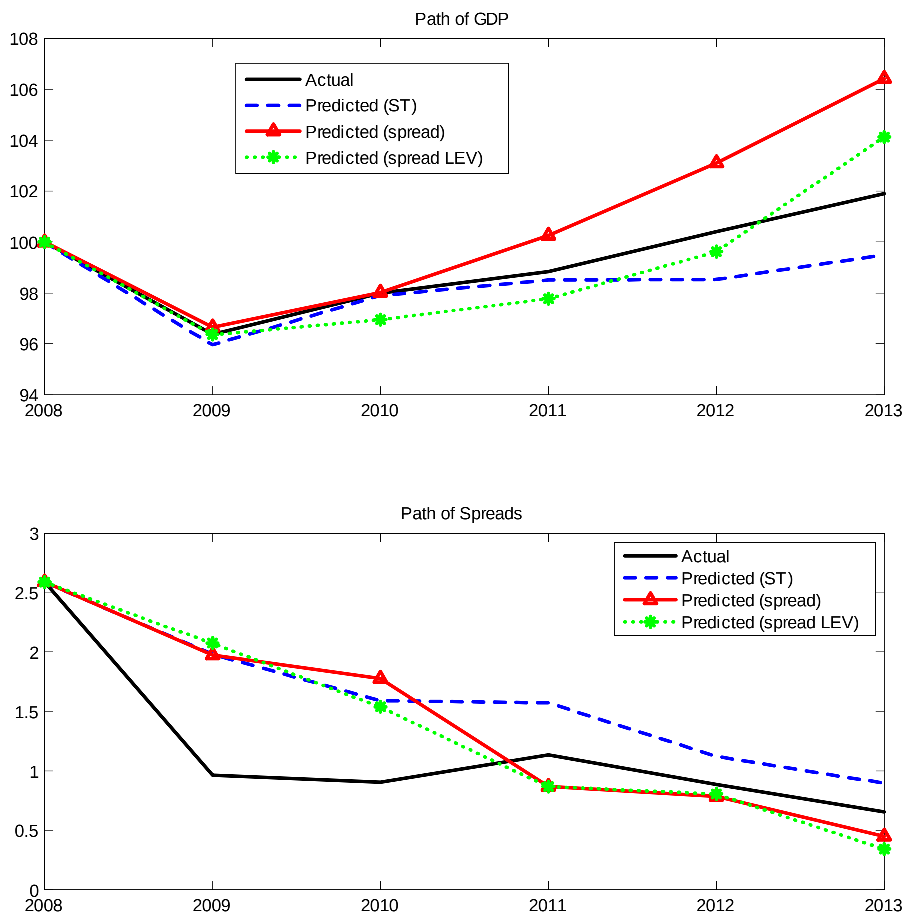

Figure 8: We predict outcomes of output and spreads during the 2008 US financial crisis using predicted values from our regressions and data up to 2008. The top panel, GDP, is cumulative from a base of 100 in 2008. The lower panel, spreads, uses the last quarter value of the BaaAaa spread in 2008.

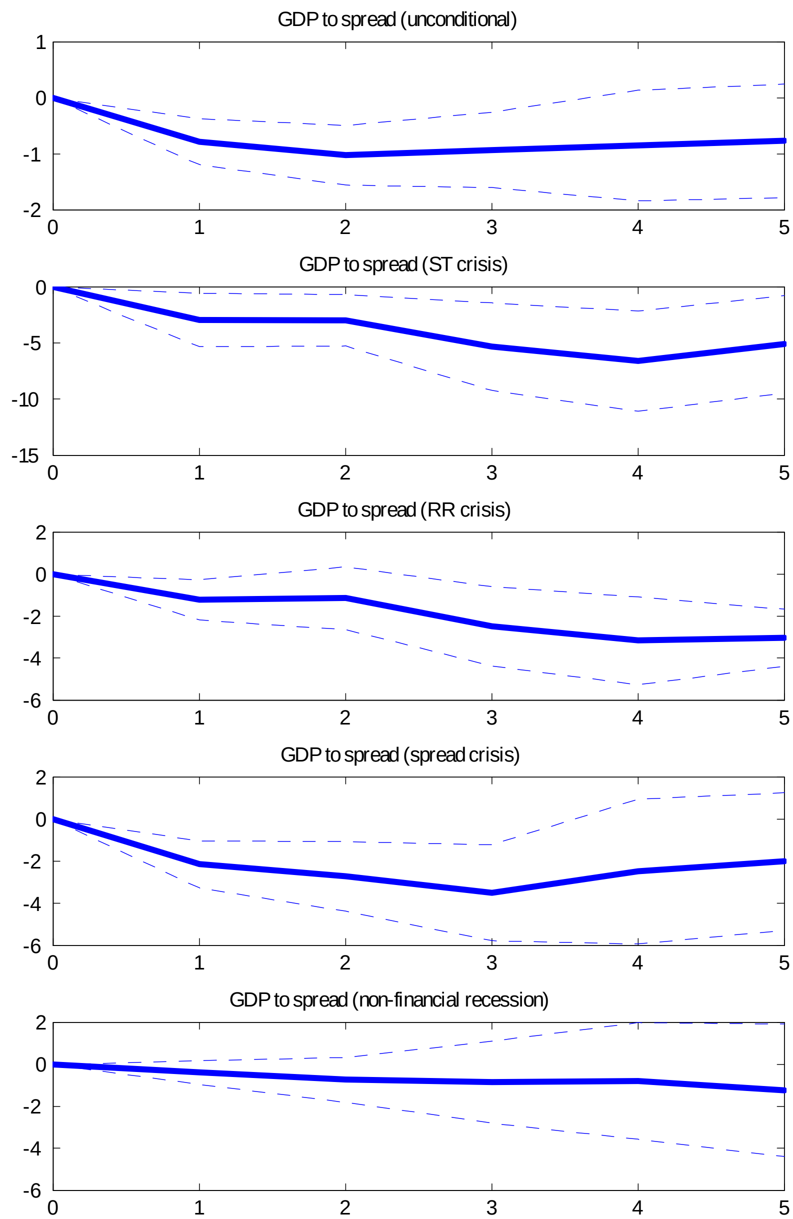

Figure 9: This figure plots the impulse responses of GDP and normalized spreads to an innovation in our spread measure conditional on various episodes. We show this unconditionally in the top panel (labeled normal times) as well as during crises in the lower panels, where crises are defined using Schularick and Taylor (ST), Reinhart and Rogoff (RR), or spread dates. The last panel gives these results using non-financial recessions. Impulse responses are computed using local projection measures where we forecast GDP independently at each horizon.

Table 1: This table provides basic summary statistics on the bonds in our sample.

Panel A: Bond Statistics for 1869-1929

Observations	Unique bonds	% Gov't	% Railroad	% Other
194,854	4,464	23%	27%	50%
Median Yield	Median Coupon	Median Discount	Avg Maturity	Median Spread
5.5%	4.2%	6%	17 years	1.9%

Panel B: Full Sample Coverage by Country

Country	First Year	Last Year	Total Years	ST Sample
Argentina	1869	1929	60	N
Australia	1869	1929	60	Y
Austria	1869	1929	57	N
Belgium	1869	1929	57	N
Brazil	1869	1929	60	N
Canada	1869	1929	60	Y
Chile	1869	1929	60	N
China	1877	1929	52	N
Colombia	1869	1929	58	N
Denmark	1869	1929	51	Y
Egypt	1869	1929	60	N
Finland	1909	1929	20	N
France	1869	1929	60	Y
Germany	1871	1929	43	Y
Greece	1869	2012	67	N
Hong Kong	1890	2012	58	N
Iceland	1922	1929	8	N
India	1869	1929	60	N
Indonesia	1869	1929	22	N
Italy	1869	1929	60	Y
Japan	1870	2001	70	Y
Korea	1995	2013	19	N
Malaysia	1908	1929	21	N
Mexico	1869	1929	60	N
Netherlands	1869	1929	60	Y
Norway	1876	1929	53	Y
Peru	1869	1929	60	N
Philippines	1890	1928	38	N
Portugal	1869	2012	66	N
Russia	1869	1929	60	N
S Africa	1869	1929	60	N
Singapore	1878	1929	34	N
Spain	1869	2012	72	Y
Sweden	1869	2013	87	Y
Switzerland	1899	1929	29	Y
Turkey	1869	1929	60	N
United Kingdom	1869	1929	60	Y
United States	1869	2014	145	Y

Table 2: This table provides regressions of future 1 year GDP growth on credit spreads where we consider different normalizations of spreads. The text discusses these normalizations in more detail. Standard errors clustered by year.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
VARIABLES	Raw	Raw	MeanNorm	MeanNorm	MedNorm	MedNorm	Zscore	Zscore	Percentile	Percentile
spread	-0.08	-0.07								
	(0.08)	(0.07)								
lspread	0.08	0.07								
	(0.05)	(0.05)								
spreadnorm			-0.81	-0.71						
			(0.35)	(0.30)						
lspreadnorm			0.76	0.54						
			(0.28)	(0.31)						
spreadnormmed					-0.15	-0.14				
					(0.10)	(0.09)				
lspreadnormmed					0.15	0.08				
					(0.08)	(0.08)				
spreadzscore							-1.16	-0.91		
							(0.42)	(0.39)		
lspreadzscore							1.04	0.69		
							(0.31)	(0.30)		
spreadpcrank									-3.13	-1.12
									(0.82)	(1.07)
lspreadpcrank									2.20	0.16
									(0.69)	(0.98)
Observations	639	639	639	639	639	639	639	639	639	639
R-squared	0.04	0.35	0.06	0.36	0.04	0.35	0.08	0.36	0.05	0.35
Country FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Year FE	N	Y	N	Y	N	Y	N	Y	N	Y

Table 3: This table shows the forecasting power of credit spreads for the severity of financial crises in terms of the peak to trough declines in GDP. We provide results for both the Schularick and Taylor as well as Reinhart and Rogoff dates as well as dates that mark regular non-financial recessions. The top panel uses only spreads, the bottom panel adds control variables including the change in leverage from Schularick and Taylor which is a measure of credit growth. See text for further description.

$$decline_{i,t} = a + b\widehat{s}_{i,t} + \varepsilon_{i,t}$$

VARIABLES	(1) Crisis RR	(2) Crisis RR	(3) Crisis ST	(4) Crisis ST	(5) Recess	(6) Recess
$\widehat{s}_{i,t}$	-1.08	-1.47	-2.40	-2.51	-0.97	-1.62
	(0.45)	(0.68)	(0.56)	(0.64)	(0.49)	(0.51)
Observations	91	36	51	39	123	79
R-squared	0.06	0.12	0.27	0.29	0.03	0.11
Restricted Sample	N	Y	N	Y	N	Y
Variation in Realized Severity $\sigma(decline)$	5.9	6.4	7.3	8.0	7.1	5.7
Variation in Expected Severity $\sigma(E_t[decline])$	1.5	2.3	3.8	4.3	1.3	1.9

Adding Controls

VARIABLES	(1) Crisis RR	(2) Crisis RR	(3) Crisis ST	(4) Crisis ST	(5) Recess	(6) Recess
$\widehat{s}_{i,t}$	-1.66	-4.09	-2.48	-5.57	-1.92	-2.27
	(0.79)	(2.72)	(0.73)	(1.69)	(0.71)	(1.30)
$\widehat{s}_{i,t-1}$		2.34		4.08		0.51
		(3.32)		(1.90)		(1.46)
$\Delta lev_{i,t}$		-68.59		-32.54		20.44
		(30.54)		(19.35)		(23.52)
Observations	29	29	29	29	70	70
R-squared	0.11	0.21	0.28	0.42	0.08	0.07
Controls	N	Y	N	Y	N	Y
Variation in Realized Severity $\sigma(decline)$	7.0	7.0	8.7	8.7	8.0	8.0
Variation in Expected Severity $\sigma(E_t[decline])$	2.6	3.8	4.8	6.0	2.5	2.6

Table 4: This table provides regressions of future GDP growth on credit spreads at the 1 and 5 year horizon (top and bottom panels, respectively). We include interactions with crisis or recession dummies to assess whether spreads become more informative during crisis periods.

	(1)	(2)	(3)	(4)	(5)
VARIABLES	5yr	5yr	5yr	5yr	5yr
$\widehat{s}_{i,t}$	-0.90				
	(0.52)				
$\widehat{s}_{i,t-1}$	1.38				
	(0.63)				
$\widehat{s}_{i,t} \times 1_{crisis,i,t}$		-4.96			
		(2.14)			
$\widehat{s}_{i,t-1} \times 1_{crisis,i,t}$		5.95			
		(2.29)			
$\widehat{s}_{i,t} \times 1_{crisisRR,i,t}$			-3.13		
			(0.68)		
$\widehat{s}_{i,t-1} \times 1_{crisisRR,i,t}$			1.44		
			(0.65)		
$\widehat{s}_{i,t} \times 1_{recess,i,t}$				-1.22	
				(1.61)	
$\widehat{s}_{i,t-1} \times 1_{recess,i,t}$				0.01	
				(1.30)	
$\widehat{s}_{i,t} \times 1_{spread,i,t}$					-2.00
					(1.67)
$\widehat{s}_{i,t-1} \times 1_{spread,i,t}$					3.54
					(1.57)
Observations					
R-squared	0.55	0.57	0.56	0.56	0.77
Controls	Y	Y	Y	Y	Y

Table 5: This table provides regressions of future GDP growth on credit spreads at the 3 year horizon. We include interactions with crisis or recession dummies to assess whether spreads become more informative during crisis periods.

VARIABLES	(1) 3yr	(2) 3yr	(3) 3yr	(4) 3yr	(5) 3yr
$\widehat{s}_{i,t}$	-0.95				
	(0.36)				
$\widehat{s}_{i,t-1}$	0.69				
	(0.50)				
$\widehat{s}_{i,t} \times 1_{crisis,i,t}$		-5.21			
		(1.96)			
$\widehat{s}_{i,t-1} \times 1_{crisis,i,t}$		5.48			
		(1.78)			
$\widehat{s}_{i,t} \times 1_{crisisRR,i,t}$			-2.57		
			(0.94)		
$\widehat{s}_{i,t-1} \times 1_{crisisRR,i,t}$			0.73		
			(0.49)		
$\widehat{s}_{i,t} \times 1_{recess,i,t}$				-0.84	
				(0.99)	
$\widehat{s}_{i,t-1} \times 1_{recess,i,t}$				-0.03	
				(0.88)	
$\widehat{s}_{i,t} \times 1_{spread,i,t}$					-3.50
					(1.16)
$\widehat{s}_{i,t-1} \times 1_{spread,i,t}$					0.46
					(1.41)
Observations	467	467	467	467	229
R-squared	0.56	0.59	0.58	0.57	0.79
Controls	Y	Y	Y	Y	Y

Table 6: This table provides regressions of our normalized spread measure on lagged spreads conditional on crises and normal times. Standard errors clustered by year.

$\widehat{s}_{i,t+1} = a_i + b\widehat{s}_{i,t} + cx_{i,t} + e_{i,t+1}$				
	(1)	(2)	(3)	(4)
VARIABLES				
Δlev	4.04	3.87	4.09	4.12
	(1.61)	(1.44)	(1.61)	(1.58)
$\widehat{s}_{i,t}$	0.67			
	(0.15)			
$\widehat{s}_{i,t} \times 1_{crisis,i,t}$		0.50		
		(0.33)		
$\widehat{s}_{i,t} \times (1 - 1_{crisis,i,t})$		0.68		
		(0.16)		
$\widehat{s}_{i,t} \times 1_{crisisRR,i,t}$			0.44	
			(0.36)	
$\widehat{s}_{i,t} \times (1 - 1_{crisisRR,i,t})$			0.69	
			(0.16)	
$\widehat{s}_{i,t} \times 1_{recess,i,t}$				1.16
				(0.26)
$\widehat{s}_{i,t} \times (1 - 1_{recess,i,t})$				0.62
				(0.15)
$\widehat{s}_{i,t-1}$	0.12	0.12	0.13	0.09
	(0.14)	(0.14)	(0.14)	(0.13)
$1_{crisis,i,t}$		0.51		
		(0.25)		
$1_{crisisRR,i,t}$			0.58	
			(0.27)	
$1_{recess,i,t}$				-0.32
				(0.23)
Observations	449	449	449	449
R-squared	0.51	0.51	0.52	0.53
Country FE	Y	Y	Y	Y
Controls	Y	Y	Y	Y

Table 7: Quantile Regressions. We run quantile regressions of output growth on spreads.

Quantile Regressions					
	(1)	(2)	(3)	(4)	(5)
VARIABLES	Q 90th	Q 75th	Q Median	Q 25th	Q 10th
$\widehat{s}_{i,t}$	-0.30	-0.36	-0.64	-1.07	-1.61
	(0.32)	(0.18)	(0.19)	(0.21)	(0.40)
$\widehat{s}_{i,t-1}$	0.69	0.73	0.62	0.83	1.11
	(0.35)	(0.20)	(0.21)	(0.24)	(0.45)
Observations	637	637	637	637	637
Country FE	Y	Y	Y	Y	Y
Controls	Y	Y	Y	Y	Y
Pseudo R2	0.11	0.08	0.05	0.07	0.12

Table 8: This table provides summary statistics for declines in GDP around crisis episodes.

Distribution of declines in GDP across episodes

Financial Crises (ST dates)						
	Mean	Median	Std Dev	P 10th	P 90th	N
Full Sample	-6.2	-3.9	7.3	-12.2	-0.1	51
Restricted	-7.2	-4.9	8.0	-17.4	-0.7	39
Financial Crises (RR dates)						
	Mean	Median	Std Dev	P 10th	P 90th	N
Full Sample	-3.6	-1.1	5.9	-8.8	0	92
Restricted	-4.4	-3.0	6.4	-10.9	0	37
Spread crises						
	Mean	Median	Std Dev	P 10th	P 90th	N
Full Sample	-5.9	-2.5	9.3	-23.4	0	44
Restricted	-6.2	-2.4	9.6	-24.0	0	40

Table 9: Which spread crises turn out badly? We run regressions where the left hand side is GDP growth at various horizons. The right hand side contains a dummy for whether there was a crisis according to our spread variable. It also interacts this dummy with leverage at the beginning of the crisis. The lower panel instead intereacts spreads with a dummy for when leverage growth is high, defined based on the 90th percentile of leverage growth. The table shows that high spreads are bad news for output on average, but are particularly bad and long lasting when leverage is high. Standard errors in parenthesis.

When is an increase in spreads particularly bad for GDP?

VARIABLES	(1) 1yr	(2) 1yr	(3) 2yr	(4) 2yr	(5) 3yr	(6) 3yr	(7) 4yr	(8) 4yr	(9) 5yr	(10) 5yr
SpreadCrisis	-4.38		-3.47		-2.54		-0.41		1.76	
	(0.86)		(1.29)		(1.69)		(1.95)		(2.22)	
(SpreadCrisis) x		-6.75		-7.27		-7.30		-4.56		-0.63
(Leverage)		(1.23)		(1.93)		(2.54)		(2.96)		(3.39)
Observations	242	242	241	241	240	240	239	239	238	238
R-squared	0.17	0.19	0.16	0.19	0.19	0.21	0.21	0.21	0.20	0.20
Controls	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y

VARIABLES	(1) 1yr	(2) 2yr	(3) 3yr	(4) 4yr	(5) 5yr
(HighLeverage) x	-1.27	-2.74	-4.30	-5.84	-6.20
(Spread)	(0.41)	(1.06)	(1.48)	(2.16)	(2.35)
Observations	458	457	457	456	455
R-squared	0.09	0.16	0.24	0.27	0.23
Controls	Y	Y	Y	Y	Y

Table 10: Initial spread vs integral vs future spread.

Panel A: ST crisis dates

VARIABLES	(1) 2yr	(2) 2yr	(3) 2yr	(4) 3yr	(5) 3yr	(6) 3yr	(7) 4yr	(8) 4yr	(9) 4yr
Shock $\widehat{s}_{i,t} \times 1_{crisis,i,t}$		-3.13			-2.93			-1.58	
		(1.72)			(1.12)			(1.06)	
Shock $\widehat{s}_{i,t+1} \times 1_{crisis,i,t}$	-1.55								
	(0.81)								
$\int_0^1$Shock $\widehat{s}_{i,t+k} \times 1_{crisis,i,t}$			-1.34						
			(0.57)						
Shock $\widehat{s}_{i,t+2} \times 1_{crisis,i,t}$				-0.61					
				(0.82)					
$\int_0^2$Shock $\widehat{s}_{i,t+k} \times 1_{crisis,i,t}$						-0.62			
						(0.19)			
Shock $\widehat{s}_{i,t+3} \times 1_{crisis,i,t}$							-0.64		
							(0.51)		
$\int_0^3$Shock $\widehat{s}_{i,t+k} \times 1_{crisis,i,t}$									-0.16
									(0.14)
N	558	558	558	539	539	539	520	520	520
R-squared	0.51	0.51	0.51	0.50	0.51	0.51	0.39	0.39	0.39

Panel B: RR crisis dates

VARIABLES	(1) 2yr	(2) 2yr	(3) 2yr	(4) 3yr	(5) 3yr	(6) 3yr	(7) 4yr	(8) 4yr	(9) 4yr
Shock $\widehat{s}_{i,t} \times 1_{crisis,i,t}$		-0.64			-0.98			-0.27	
		(1.51)			(1.07)			(1.09)	
Shock $\widehat{s}_{i,t+1} \times 1_{crisis,i,t}$	-2.21								
	(1.09)								
$\int_0^1$Shock $\widehat{s}_{i,t+k} \times 1_{crisis,i,t}$			-0.82						
			(0.40)						
Shock $\widehat{s}_{i,t+2} \times 1_{crisis,i,t}$				-0.16					
				(1.23)					
$\int_0^2$Shock $\widehat{s}_{i,t+k} \times 1_{crisis,i,t}$						-0.46			
						(0.22)			
Shock $\widehat{s}_{i,t+3} \times 1_{crisis,i,t}$							0.26		
							(0.48)		
$\int_0^3$Shock $\widehat{s}_{i,t+k} \times 1_{crisis,i,t}$									-0.07
									(0.17)
N	558	558	558	539	539	539	520	520	520
R-squared	0.51	0.51	0.51	0.50	0.50	0.50	0.38	0.38	0.38

Table 11: Initial spread vs integral vs future spread.

	Spread Crises								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
VARIABLES	2yr	2yr	2yr	3yr	3yr	3yr	4yr	4yr	4yr
Shock $\widehat{s}_{i,t} \times 1_{crisis,t}$		0.59			1.57			0.34	
		(2.74)			(2.10)			(1.30)	
Shock $\widehat{s}_{i,t+1} \times 1_{crisis,t}$	-4.39								
	(1.60)								
$\int_0^1$Shock $\widehat{s}_{i,t+k} \times 1_{crisis,t}$			-1.43						
			(1.10)						
Shock $\widehat{s}_{i,t+2} \times 1_{crisis,t}$				-0.54					
				(1.16)					
$\int_0^2$Shock $\widehat{s}_{i,t+k} \times 1_{crisis,t}$						-0.23			
						(0.36)			
Shock $\widehat{s}_{i,t+3} \times 1_{crisis,t}$							-5.74		
							(3.54)		
$\int_0^3$Shock $\widehat{s}_{i,t+k} \times 1_{crisis,t}$									0.11
									(0.38)
N	310	310	310	300	300	300	290	290	290
R-squared	0.69	0.67	0.68	0.69	0.69	0.69	0.61	0.60	0.60