Credit Spreads and the Severity of Financial Crises *

Arvind Krishnamurthy — Stanford GSB

Tyler Muir — Yale SOM

March 25, 2015

Abstract

We study the behavior of credit spreads and their link to economic growth during financial crises. We have three main findings. First, credit spreads accurately forecast the depth of the financial crisis in terms of GDP losses and thus the aftermath of a financial crisis is predictable based on the initial increase in spreads. Moreover much of this forecasting power comes from forecasting the worst quartile of GDP outcomes. That is, high spreads indicate a much higher probability of a bad tail event. Second, credit spreads the year before a crisis are largely uninformative about future GDP, suggesting that crises are largely a surprise. Third, while financial crises feature "slow recoveries" in terms of GDP and economic growth, credit spreads revert to pre-crisis levels more quickly, suggesting a separate role for financial and real factors in explaining the evolution of the macroeconomy in the aftermath of a financial crisis.

*Stanford Graduate School of Business and Yale School of Management.

1 Introduction

We examine the behavior of credit spreads and output around financial crises in an international panel spanning from the 1800's 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 (2009) 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 sorting out several of the issues that arise with existing studies. First, 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 2). The oft-cited number from (Reinhart and Rogoff (2009)) is that crisis contraction is 9%, with this mean measured from a mix of many different financial crises. We show that 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.

Second, almost all of the existing research dates crises ex-post. The Great Depression is a financial crisis because we observe ex-post that it involved bank failures and a significant collapse in output. The 1998 hedge fund financial crisis is not labeled a crisis in (Reinhart and Rogoff (2009)) and Jorda *et al.* (2010) because there little apparent contagion to the banking sector and the real economy. In both of these events, credit spreads rose sharply on the incidence of the crisis. But apparently in the 1998 event, either through luck or adequate government intervention, the financial crisis dissipated and did not have knock-on effects. By dating crises ex-post, research can be biased toward selecting the worst financial crises episodes. By focusing on ex-ante measure, we overcome this bias. Using an ex-ante measure of crises based on spreads, we find that the median peak-to-trough GDP decline in a crisis is -0.2% while the standard deviation of this measure is 11.7%. The mean decline is -6.9%. As we discuss, high spreads changed the conditional distribution of crises outcomes, so that it is the worst quartile of the distribution that evidences a severe contraction.

Last, there is a growing body of research examining the forecastability of crises. Jorda *et al.* (2010) have documented prominently that credit growth helps forecast financial crises. We show that credit spreads once a crisis begins helps forecast the severity of the crisis. However, credit spreads the year prior to the crisis have little forecasting power. Put differently, all the forecasting power of credit spreads comes from the surprise "jump" on impact when the crisis begins, and little of it comes from credit spreads starting at high levels. Moreover, we show that in regressions including both the Jorda *et al.* (2010) measure and credit spreads, both have independent explanatory power for the severity of the crisis.

Using credit spreads, we are able to paint a better picture of crises than existing studies. There is a lot of variation within crises. Indeed a central finding of our paper is that the standard deviation of outcomes is as interesting a variable as mean outcomes. There are financial crises such as the 1998 hedge fund crisis which resolve quickly, while there are others such as the Great Depression that are persistent and severe. In thinking about the US experience, a relevant benchmark is to ask how slow a recovery should have been expected given the spreads in 2008. We find that the current crisis behaved remarkably close to what was implied by 2008 spreads.

Our research does not allow us to address important causality questions such as, do high spreads cause crises? But it does document facts about financial crises that macroeconomic models of crises need address. One important finding is that spreads in a financial crisis recover more quickly than output. The average time for spreads to revert back to normal levels is between one and two years, while the average number of years for GDP to reach its trough after a crisis is closer to four years. This fact suggests that there at least two state variables driving financial and real variables in a crisis episode.

We conduct a number of robustness exercises related to our main three results. A major worry is that the sample of financial crises is relatively small and consists of large outliers that influence our main result. We show that, consistent with previous findings, the relationship between credit spreads and GDP is stable and holds during normal times, during recessions, and during financial crises. The coefficient in our forecasting regression is similar when we expand to using all recessions and when we forecast GDP growth unconditionally. This gives us much more confidence that the coefficient found during crises is reasonable.

We also consider a number of alternative dating schemes including Reinhart and Rogoff (2009), Jorda *et al.* (2010), and dates based on credit spread spikes. Specifically we date a crisis not based on existing dates, but based on a large spike in spreads. This has advantages as well as drawbacks. A main advantage is that this might capture the timing of crises more accurately and also allows us to add more observations by varying the threshold of spread increases. The downside is the dates are less comparable to those in the literature.

The most closely related papers to ours are Reinhart and Rogoff (2009), Jorda *et al.* (2010), and Romer and Romer (2014) who study the behavior of macro variables during financial crises (see also Bordo and Haubrich (2012)). Relative to these papers, we bring data on credit spreads and show that this can add substantially to our understanding of crises. Our paper is also closely related to work on credit spreads and economic growth, most notably Gilchrist and Zakrajsek (2012). Relative to this work we study the behavior of spreads specifically in financial crises and study an international panel of bond price 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 the correlation we find.

2 Data and Definitions

Crisis dates come from two sources: Reinhart and Rogoff (2009) and Jorda *et al.* (2010) (henceforth RR and ST). The dates provided by Reinhart and Rogoff (2009) 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. The first is it is directly comparable to the exercise of comparing crisis outcomes to regular recessions or business cycle peaks. The second is it gives a clear guideline for choosing dates, rather than trying to understand when a systemic bank run occurred. The third is 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.

Data on credit spreads come 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. When constructing spreads, we form within country 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.

For data from 1930 onward, we use various sources where the source will depend on the country. For the United States, we use the BaaAaa default spread from 1930-present. Our data appendix discusses the details and construction of this data extensively.

Data on real per capita GDP are from Barro and Ursua (see Barro *et al.* (2011)). We construct three measures of severity of a crisis based on the behavior of GDP in the years after the crisis occurs. 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. 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. This is one reason why we prefer the ST dates in our main results. The last measure of severity is simply the 5 year cumulative growth in GDP after a crisis has occured. We choose 5 years again to account for persistent negative effects to GDP after crises. The 5 year growth rate will aso capture experiences were growth is low relative to trend but not necessarily persistently negaitve (i.e., Japan in 1990). Our other measures will not pick up these effects. Our primary measure is the peak to trough decline in GDP as our measure of severity, but our aim is to use these other benchmarks for robustness as

[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.

[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.

there is no absolute clear measure of crisis severity.

We define the "duration" of a crisis in terms of real economic growth as the number of years it takes GDP to reach trough. We consider an alternative measure of duration that counts the number of years it takes GDP to return to its value at the beginning of a crisis (this measure will always be longer than our standard duration measure). We label the first "time to trough" and the second time back to previous peak. Finally, when using the "flexible depth" measure of severity described above, we define an analagous "flexible duration" measure that counts the number of years until GDP reaches this local minimum value.

We define "spread duration" as the number of years, after the onset of the crisis, it takes credit spreads to return to their pre-crisis level, namely the level of spreads the year before the onset of the crisis. If spreads have reverted back to pre-crisis levels, it seems unlikely that high spreads would be the continual cause of drops in GDP, so this definition helps us assess how long it takes spreads to normalize and compare this with GDP. This is a strict definition of recovery because spreads before crises may be very low as there is often a credit boom. Spreads may recover to very near pre-crisis levels but stay slightly above these levels. Therefore we have also considered using the half life of spreads, or number of years it takes for spreads to reach half of their level observed in the crisis.

In all cases our forecasting variables are measured in the year of the crisis, which we label $t = 0$. Real outcomes are forecast looking forward from this date.

3 Model of credit spreads and output

Denote the growth of output of country i between s and $s+1$ as g^i_{s+1}. We write,

$$g^i_{s+1} = A^i + x^i_s + \epsilon^i_{s+1} \tag{1}$$

where A^i is a mean growth rate for the country and x^i_s is a variable that reflects the expected growth rate. We assume that,

$$x^i_{s+1} = \rho x^i_s + v^i_{s+1}, \quad \rho < 1 \tag{2}$$

A shock of $v^i_s < 0$ that reduces growth for the next period, also reduces growth for the following period, dying out over many periods at rate ρ. One should think of the shock v^i_{s+1} that corresponds to a crisis as large, perhaps an extreme realization from the shock distribution.

At date t, investors know x_t^i and thus form expectations over growth rates for $s > t$. We price a corporate bond based on this information. The price of the bond, P_t^C is,

$$P_t^C = e^{-r_t}\left((1-\lambda_t)E_t[P_{t+1}^C] + \lambda_t E_t[D_{t+1}]\right)$$

Here λ_t is the risk-neutral probability of the bond defaulting over the next period, while D_{t+1} is the recovery of the bond in default, r_t is the riskless one period interest rate, and all expectations are under the risk-neutral measure. It is important to underline that λ_t is the risk-neutral default probability; one can always express this default probability as $\lambda_t^P\gamma_t$ where λ_t^P is the true probability of default and γ_t is a risk-adjustment which captures the risk aversion of the marginal investor. Thus, λ_t reflects both the probability of default and a possibly time-varying economic risk premium on default.

We follow Duffie and Singleton (1999) and assume that we can write,

$$E_t[D_{t+1}] = E_t[P_{t+1}^C](1-l_t)$$

This is their "recovery of market value" assumption. With this assumption, we find that,

$$P_t^C \approx E_t[P_{t+1}^C]e^{-r_t - l_t\lambda_t}$$

so that the bond price at t is the discounted expected value of the bond price at $t+1$, given no-default, where the discount rate is $r_t + l_t\lambda_t$, and reflects the probability of default. Iterating forward to maturity at time $s > t$, we can rewrite this expression as $E_t[e^{-\sum_{\tau=t}^{s} r_\tau}]\ E_t[e^{-\sum_{\tau=t}^{s} l_\tau\lambda_\tau}]$. We note that $E_t[e^{-\sum_{\tau=t}^{s} r_\tau}] = e^{-R_{t,s}}$ where $R_{t,s}$ is the interest rate on a riskless bond with maturity s. Thus the yield on this corporate bond is,

$$Yield_{t,s} = -ln\ P_t^C = R_{t,s} - ln\ E_t[e^{-\sum_{\tau=t}^{s} l_\tau\lambda_\tau}]$$

The yield on the corporate bond in excess of the riskless rate reflects three objects: the probability of default in each period between t and s (i.e. $\lambda_t^P = \lambda_t/\gamma_t$), the loss given default in each period between t and s (i.e., l_t), and the market price of default risk (or economic risk premium) that investors charge for bearing default risk for each period between t and s (i.e. γ_t).

We focus on the spread between two corporate bonds. For example, for the case of the US we construct the spread between Baa rated bonds and Aaa rated bonds. This spread picks up the difference in $ln\ E_t[e^{-\sum_{\tau=t}^{s} l_\tau\lambda_\tau}]$ across these two bonds. We assume that we can write this spread as a linear function of the underlying growth state variable:

$$Sp_t^i = a_0^i + b_x^i x_t^i + e_t^i \tag{3}$$

The spread for this bond has mean value parameterized by a_0^i. In the time series, changes in expected growth raise or lower this spread proportionately to b_x^i.

3.1 Normalized spread

We construct bond spreads across many different countries for which the underlying bond default risks, as captured by a_0^i, and b_x^i are likely to be different. For example, while for the US we construct the Baa-Aaa spread, for Sweden we construct the spread between bank loans and government bond yields. We need to make further assumptions to normalize and allow for comparability of spreads across countries.

We follow two approaches. Our main approach is to assume that we can write,

$$\textbf{A1:} \quad a_0^i = c^i a_0, \qquad b_x^i = c^i b_x \tag{4}$$

for all countries. This assumption says that the average level of any given spread is proportional to the sensitivity of the spread to the growth state. While we cannot verify the assumption across countries, it is possible to check this assumption *within* countries. That is, we can check how the average sensitivity of spreads to the business cycle is related to the average level of spreads in the cross-section of US corporate bonds. We have investigated this relation in the US, and find results consistent with our assumption.

With this assumption, we note that,

$$E[Sp_t^i] = c^i a_0,$$

so that,

$$Spreadnorm_t^i \equiv \frac{Sp_t^i}{E[Sp_t^i]} = 1 + \frac{b_x}{a_0} x_t^i + \frac{e_t^i}{c^i a_0}. \tag{5}$$

where now $\frac{b_x}{a_0}$ is constant across all countries. Thus, we can express,

$$\begin{aligned} E[g_{t+k}^i | x_t^i] &= A^i + E[x_{t+k}^i | x_t^i] \\ &= A^i + \rho^k x_t^i \\ &= A^i + \rho^{s-t} \frac{a_0}{b_x} \left(Spreadnorm_t^i - 1 - \frac{e_t^i}{c^i a_0} \right), \end{aligned}$$

where we note that the coefficient in front of $Spreadnorm_t^i$ does not vary across countries. We can then run OLS regressions:

$$g_{t+k}^i = \alpha^i + b \times Spreadnorm_t^i + \epsilon_{t+k}^i \tag{6}$$

where we measure $E[Sp_t^i]$ using the sample average spread for a given country.

3.2 Weighted spread

We take a second approach that does not require assumption **A1**, but does require more data. In this second approach, we first run OLS regressions of,

$$g^i_{t+k} = \alpha^i + \beta^i Sp^i_t + \epsilon^i_{t+k}$$

for each country using only non-crisis observations. That is, we estimate β^i as the sensitivity of growth to spreads in booms and normal recessions. We then construct a weighted spread,

$$Spreadweighted^i_t = \widehat{\beta^i} Sp^i_t \tag{7}$$

and run regressions in the cross-section of countries where we forecast growth rates in crisis observations using *Spreadweighted*.

This second approach essentially orders the sensitivity of spreads in a crisis according to the sensitivity of spreads to normal economic fluctuation. Note it does not impose that the sensitivities across crises and normal fluctuations are the same, and rather it just requires that the crisis sensitivities be proportional to the normal sensitivities in the cross-section of countries.

4 Results

4.1 Forecasting crises using spreads

Table 1 presents regressions of credit spreads on various measures of the severity of crises described earlier. Each data point in these regressions is a crisis in a given country-year (i, t), where crises are defined either use the ST or the RR chronology. The regression includes a constant, but this constant is fixed across country-i in these regressions. That is the regressions estimate the cross-sectional relation between measures of severity and spreads.

The first panel of the table uses the raw credit spread as explanatory variable. The estimated relation is weak, indicating that the credit risk underlying the spreads is heterogenous across countries. The second panel shows our preferred measure (*spreadnorm*) where we normalize the spreads following (5). For most measures, the spread has statistically and economically significant explanatory power for crisis severity. A one-sigma change in *spreadnorm* of 1.2% translates to a 2.6% decrease in peak-to-trough GDP using the ST dates and a 2.2% decrease using the RR dates. The effect on 5-year growth is -0.36% (ST) and -0.6% (RR). The last panel of the table uses the weighted spread measure, following

(7), and this measure also works, although the goodness-of-fit R^2 is somewhat smaller than when using *spreadnorm*.

Table 3 presents regressions where we use all of the data in panel regressions. The top panel presents results where we regress future GDP growth at 1 and 5 year horizons on current *spreadnorm*, including a country dummy that absorbs mean differences in country growth rates. This regression includes both crisis and non-crisis periods, and indicates that there is relation between spreads and subsequent GDP growth, consistent with results from the existing literature (see, for example, Gilchrist and Zakrajsek (2012)). The second panel allows for a direct comparison of crisis and non-crisis periods. We run a panel regression where the dependent variable is cumulative GDP growth over the next k years:

$$ln\ (y^i_{t+k}/y^i_t) = \alpha^i + 1_{nocrisis,t} b_1 spreadnorm^i_t + 1_{crisis,t} \left(b_2 spreadnorm^i_t + b_3\right) + \epsilon^i_t. \quad (8)$$

The regression includes a country fixed effect (Table 5 considers a specification where we include a time fixed effect). The coefficients b_1 and b_2 reflect the normal and crisis-time sensitivity of output growth to spreads. We see that both coefficients are negative across all specifications. We also see that that b_2 is larger in magnitude than b_1. This difference is likely because spreads are more informative of aggregate outcomes prior to a crisis than during normal periods. That is, the spreads are a noisy measure of x_t and the measurement error is highest in non-crisis periods when spreads are low. The measurement error leads to an attenuation bias in the coefficients. The bottom panel of Table 3 allows for a comparison across recessions and non-recessions, and further broken down across non-financial recessions and the complement. For the recession comparison, the b_2 and b_1 coefficients are similar to those for crises/non-crises. If we look further into the non-financial recession, where spreads are not rising as significantly, b_1 and b_2 are similar, indicating that it is indeed something about the signal strength of the spread in financial events that drives the forecasting power of spreads for output.

Table 5 presents regressions where we include the credit growth measure (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. Including the ST measure has no appreciable effect on the *spreadnorm* coefficient, although the ST measure does independently help to predict 5 year GDP growth. If we repeat the regression in column (2) of both panels, dropping *spreadnorm* and only including $\Delta STlev$ we find that the coefficients are -0.10 (top panel) and -0.50 (bottom panel), which are quite close to the regression coefficients in the bivariate regression. 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.

Table 5 also includes columns where we include time fixed effects. We do this as a robustness check to deal with concerns regarding common international crises (e.g., the Great Depression occurs in many countries at the same time). Not surprisingly, including time fixed effects increases the R^2 in the regressions substantially. Moreover, we note that including the time effect largely strengthens the statistical and economic significance of $spreadnorm$.

4.2 Pre-crisis spreads

Table 6 presents regressions where we use spreads lagged by one year to forecast crises. The first panel is based on all dates and shows that lagged spreads have little forecasting power once one controls for the current spread. The table shows moreover that most of the information comes from the innovation in the spread from $t-1$ to t. The bottom panel repeats the exercise for various samples including the crisis dates. The most interesting finding in this table is that the lagged spread comes in with a positive coefficient, albeit statistically significant only for the ST dates. This suggests that a jump from low spreads to high spreads is most associated with subsequent crises. In other words, if spreads are persistently high prior to a crisis this isn't necessarily informative, but a sharp increase in spreads before a crisis is particularly informative. Lopez-Salido, Stein, and Zakrajsek (2014) report a related finding in US data, suggesting that this is evidence of an overheating boom/bust cycle.

4.3 Spread-based crises and skewness

In Table 3, the coefficient b_3 measures the expected growth in crises after conditioning on the higher levels of spreads. For the ST dates, this coefficient is negative indicating that the ST dates are likely affected by an ex-post bias whereby the worst GDP outcomes are labeled crises. To be precise, here is an example of the bias that is likely present. If one conditioned only on the time-t spread, there are two ex-post outcomes, a very low growth crisis outcome and an outcome where no crisis comes to pass. ST label the former as a crisis and drop the latter. Our regression indicates that dating the crisis ex-post leads to a 1% worse GDP outcome for the ST crises. Interestingly, the issue does not arise for the RR dates, where the b_3 coefficients are statistically not different than zero.

To overcome the ex-post dating bias, we formulate a new definition of crises based on ex-ante measures. We define a crisis if,

$$Spreadnorm_{i,t} - Spreadnorm_{i,t-1} \geq \bar{\Delta}. \tag{9}$$

We have seen that lagged spreads have little explanatory power for crises, and that information forecasting 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. In the bottom panel of Table 2 we report summary statistics for the spread-based crises. Crises are associated with large mean declines in output as with the RR and ST dating. But note that under the spread dating, the median is near zero and the standard deviation of output declines is large. That is the conditional distribution of output growth is skewed towards low realizations.

The bottom panel of Table 7 provides information on how well the spread-dated crises overlap with the ST and RR crises. There is some overlap, but not as much as one would like. In investigating the data, it appears that RR and ST include a number of slow-moving crises such as the S&L crisis or the Japan crisis of the 1990s, which don't involve a significant spike in spreads. On the other hand, there are crises in our sample, such as the 1998 LTCM crisis that involves a spike in spreads, but that does not lead to a subsequent output contraction. Neither RR nor ST date this event as a crisis. These differences account for the partial overlap in crisis dates.

The top panel of Table 7 considers our panel specification (8) for the spread-based crisis dates. We see results consistent with the earlier ones in Table 3. There is a larger relation between spreads and output declines in crises than outside crises. Now however the differences between b_2 and b_1 are not as stark, likely because there is no bias in our dating. We also see that b_3 is statistically not different from zero, again confirming that the spread indicator does not suffer from the selection bias.

For the spread measure, we again investigate the relation between *spreadnorm* and severity measures. This is reported in the middle panel of Table 7. These coefficients should be compared to those in the middle panel of Table 1. We see that the coefficients are statistically significant, but the magnitudes are reduced. For example, the peak-to-trough measure has coefficients of -1.1 for spread-crises, while it is -2.0 and -1.4 for ST and RR, respectively. This result is consistent with our observation that there is likely some ex-post bias in the

RR/ST measures.

The differences between the ST/RR results and our spread-based results suggest that ST/RR essential trim off an important set of dates, such as the 1998 crisis, which resolve quickly. If one included these dates, then another essential fact of crises appears: crises lead to a skewed distribution of output growth. We investigate this further in Table 8 where we present quantile regressions of *spreadnorm* on output growth. We see that higher spreads primarily forecast the lower quantiles of GDP growth, with little explanatory power for the median. The -0.40 coefficient for 5-year GDP growth in the 10th percentile is close to the estimates from the RR and ST crisis dates.

4.4 Impulse responses

Figure 1 plots the forecast of the evolution of GDP and *spreadnorm* to a shock to *spreadnorm* normalized as 1.5 (a one-sigma shock to *spreadnorm* is 1.2). The impulse response is computed by forecasting GDP individually at all horizons from 1 to 5 years using the local projection methods in Jorda (see also Romer and Romer 2014).

$$\ln\left(y_{i,t+k}/y_{i,t}\right) = \alpha_i + b \times spreadnorm_{i,t} + \Sigma_{j=1}^{3} c_j spreadnorm_{i,t-j} + \Sigma_{j=0}^{3} d_j \Delta \ln(y_{i,t-j}) + \varepsilon_{i,t} \quad (10)$$

We run this specification for $k = 1, ..., 5$ and use the individual coefficients on spreads to trace out the effect on output given a shock to our normalized spreads. We then re-run the above equation allowing for the coefficient of spreads during crises to be different from that in normal times. That is, we replace the *spreadnorm* term with $b_1 \times 1_{nocrisis,t} \times spreadnorm_{i,t} + b_2 \times 1_{crisis,t} \times spreadnorm_{i,t} + b_3 \times 1_{crisis,t}$. This forms our crisis impulse response. We use this 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.

The top panels of the figure trace the evolution of GDP and spreads annually using our estimates for non-crisis periods. We see that output falls, reaching a low of -1% 4 years out, before recovering. Spreads fall uninformly over the 5 year period.

The bottom panel plots the same objects, but using our crisis estimates. The output response has roughly the same shape, reaching a low at 4 years, but the fall in output is much larger at -4%. It is interesting to note that the crisis is just a larger version of the normal, in the sense that the duration of the fall in output is 4 years in both crises and normal periods. Spreads fall through the 5 years, albeit somewhat faster in a crisis, as we can also see in Table 4.

The dynamics of spreads and output in response to the spread shock are quite different. Spreads recover more quickly than output. This suggests, interestingly, that to forecast output growth say between years 3 and 4 after a crisis, we need information not only the spread at year 3, but also the spread at year 0. Figure 2 helps to further clarify this point. We plot the expected path of GDP based on the initial impulse to a shock to spread in solid lines. We also plot the expected path of GDP for each consecutive year, t to $t+1$, but each time conditioning only on the spread at t and ignoring the spread at the start of the crisis. This path is plotted in dashed lines. The figure shows that the path of GDP implied by the dynamics of spreads is less negative than that implied by the initial spread spike. This also highlights we why do not specify a standard VAR for our main impulse response as it would rely on these spread dynamics.

One way of understanding this result is to think about a model with two-state variables. Suppose that there is a variable that especially affects the pricing of financial risk, f_t^i, where,

$$f_{t+1}^i = \rho^f f_t^i + v_{t+1}^i, \quad \rho^f < 1 \tag{11}$$

Shocks to x and f are common, but suppose that f can have different dynamics than x. In particular, consider models where $\rho^f << 1$ so that shocks to the financial sector's risk bearing capacity die out more quickly than shocks to the real economy. Muir (2014) presents evidence consistent with this observation. The variable f_t can be thought of as reflecting a "risk aversion" shock, such as shocks to intermediary capital in the model of He and Krishnamurthy (2012). Gilchrist and Zakrajsek (2012) show that these risk premium shocks play an important role in driving the correlation between corporate bond spreads and real activity for the US, using data from 1973 to 2010.

Next suppose that,

$$Sp_t^i = a_0^i + b_x^i x_t^i + b_f^i f_t^i + e_t^i \tag{12}$$

where $b_f^i > b_x^i$, while,

$$g_{t+1}^i = A^i + x_t^i + f_t^i + \epsilon_{t+1}^i \tag{13}$$

This model can represent the dynamics we see in the data. A shock at t of v_t^i increases both f_t and x_t. The shock has effects on both Sp_t^i and g_{t+1}^i. But as f_t^i reverts to its mean of zero more quickly, the spread falls faster relative to output; this is because $b_f^i > b_x^i$. Consider the

At present we have not estimated this two-state variable model in our data, but intend to do so in further versions of this research paper.

4.5 2007- crisis and slow recovery

An accepted wisdom following the work of Reinhart and Rogoff is that the recoveries from financial crises are especially slow. The slow recovery from the recent crisis especially confirms this hypothesis. We can revisit this issue using our historical data where we can benchmark the recent crisis relative historical crises based on the increase in spreads in 2008. Figure 3 presents these results.

In the top panel, we plot the actual path of GDP (normalized in 2008 at 100) compared to the predicted path of GDP based on the 2008 spread. We see that this crisis in terms of output matches historical crises well based on spreads in the initial years. However, we find that this crisis is worse than historical crises in the sense that in the period beyond 2012, recovery in the recent crisis is below historical recoveries. The bottom panel plots the same time path for spreads. Here we see that the actual path of spreads has come down much more quickly than would have been expected based on historical experience. The fact that spreads and output seem to have evolved differently in the recent episode than in historical crises is interesting and surprising, and that needs to be studied further.

5 Conclusion

This paper studies the behavior of credit spreads and their link to economic growth during financial crises. We have three main findings. First, credit spreads accurately forecast the depth of the financial crisis in terms of GDP losses and thus the aftermath of a financial crisis is predictable based on the initial increase in spreads. Moreover much of this forecasting power comes from forecasting the worst quartile of GDP outcomes. That is, high spreads indicate a much higher probability of a bad tail event. Second, credit spreads the year before a crisis are largely uninformative about future GDP, suggesting that crises are largely a surprise. Third, while financial crises feature "slow recoveries" in terms of GDP and economic growth, credit spreads revert to pre-crisis levels more quickly, suggesting a separate role for financial and real factors in explaining the evolution of the macroeconomy in the aftermath of a financial crisis.

6 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. Many of our bonds are callable and this will have an effect on the implied maturity we estimate. The key feature necessary for these spreads to provide informative estimates about growth is that these effects do not induce strong time-series variation in spreads. For example, if these issues affect only the level of yields this will not affect our results. Moreover, part of 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.

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.

European spreads (Ireland, Portugal, Spain, Greece) from 200-2014. 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.

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

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

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

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.

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

7 Tables and Figures

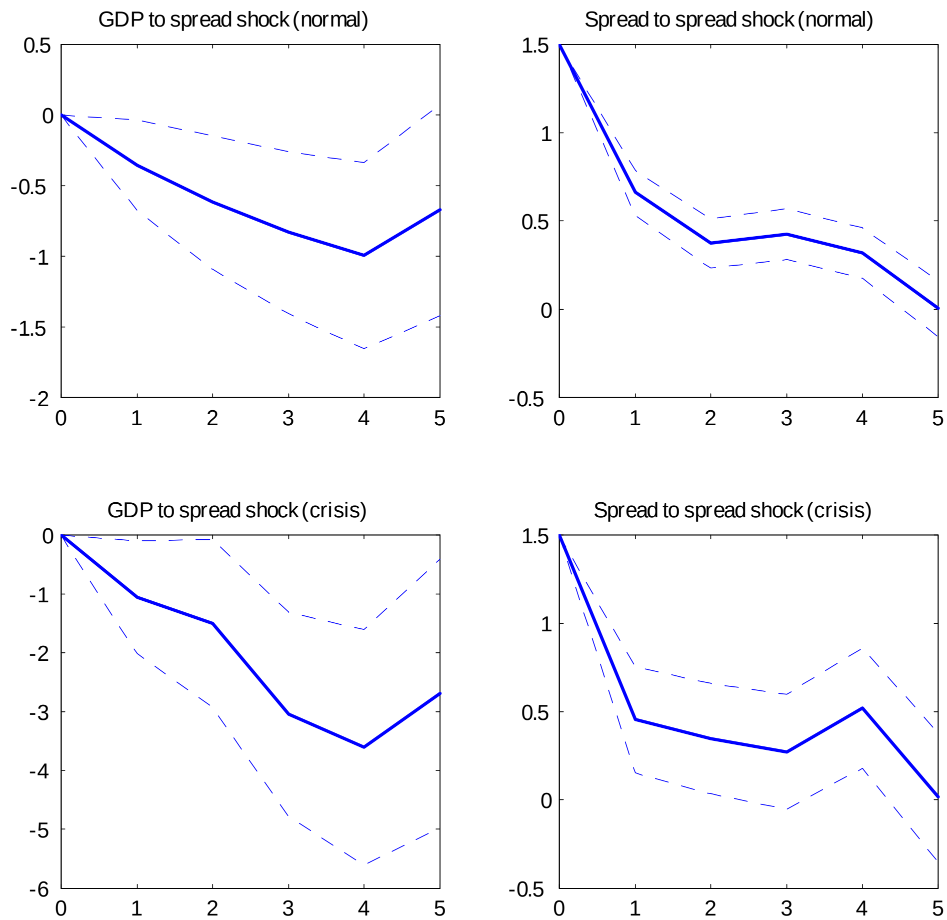

Figure 1: This figure plots the impulse responses of GDP and normalized spreads to an innovation in the spreadnorm measure. We show this unconditionally in the top panel (labeled normal times) as well as during crises in the lower panel. Impulse responses are computed using local projection measures where we forecast GDP independently at each horizon.

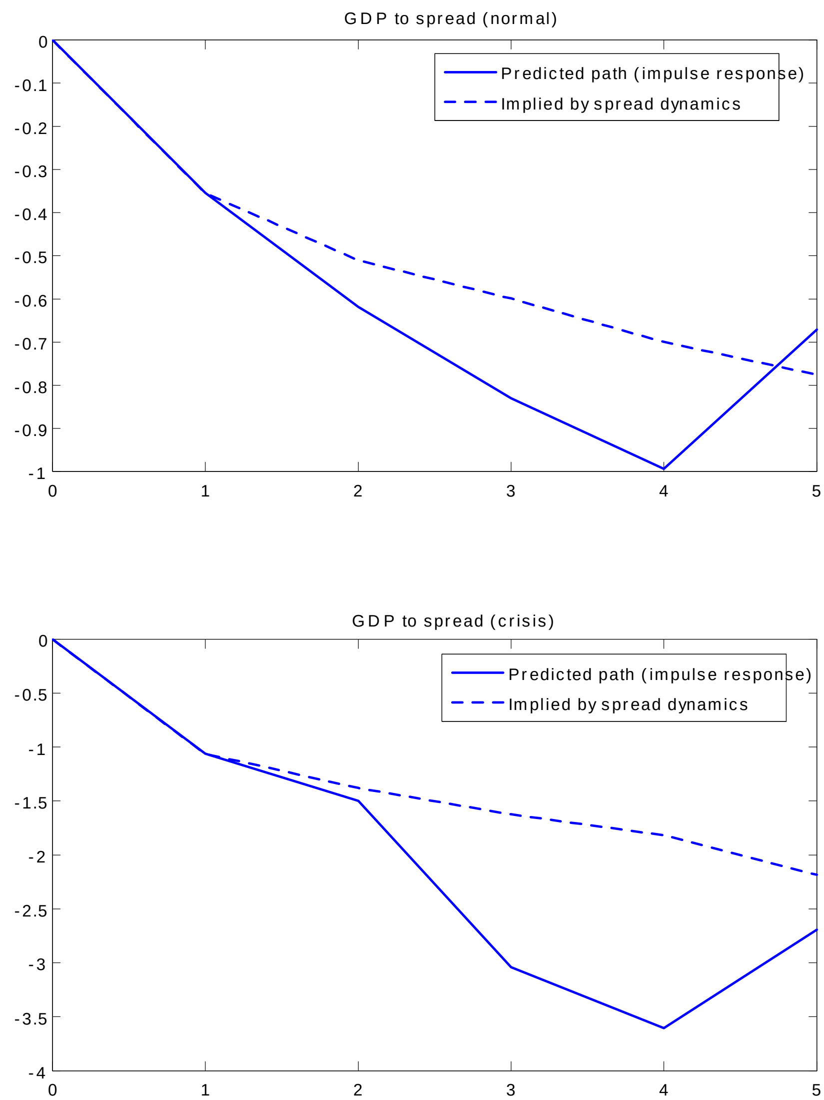

Figure 2: This figure shows the impulse response of GDP to normalized spreads computed using local projection methods along with the path of GDP implied by the persistence of spreads. For the latter, we forecast spreads at future dates and use the relationship between spreads and GDP at the one year horizon.

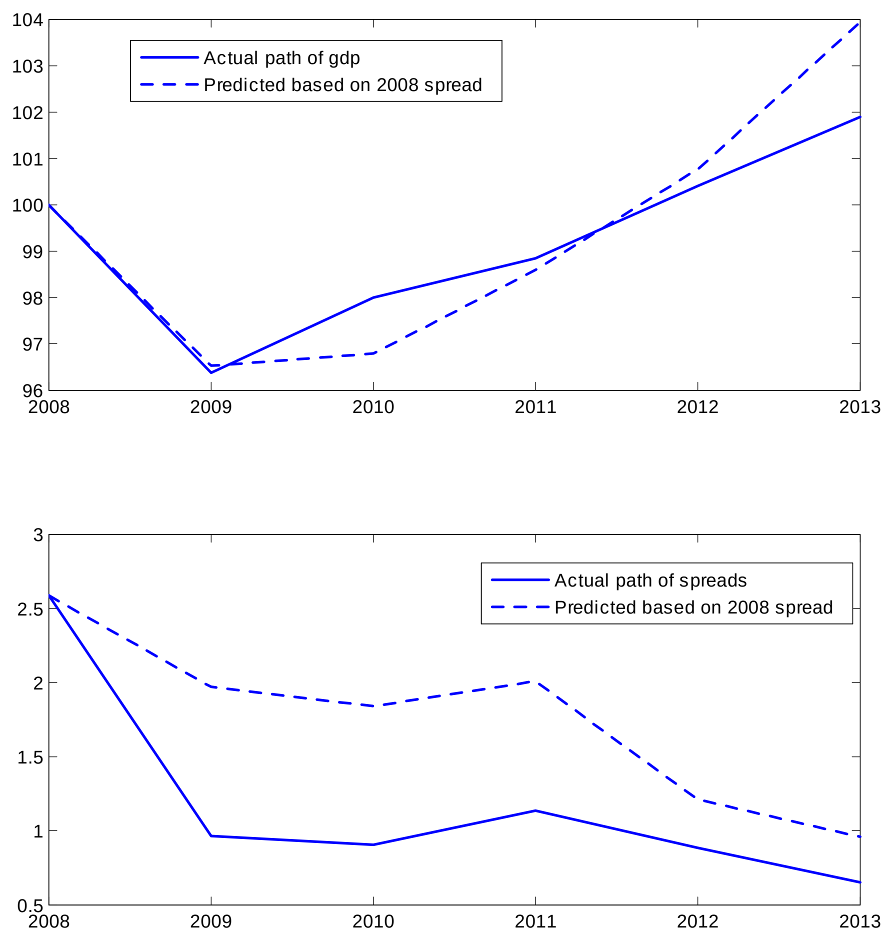

Figure 3: We predict outcomes of output and spreads during the 2008 US financial crisis using predicted values from our regressions and data up to 2008. The top panel, GDP, is cumulative from a base of 100 in 2008. The lower panel, spreads, uses the last quarter value of the BaaAaa spread in 2008.

Table 1: This table shows the forecasting power of credit spreads for the severity of financial crises in terms of the peak to trough decline in GDP. We provide results for both the Schularick and Taylor as well as Reinhart and Rogoff dates. We consider several alternative notions of severity of crises beyond peak to trough declines: flexible depth looks for the lowest value of output in a 10 year window allowing for a double dip, 5 year growth measures the 5 year growth in GDP following the crisis. We forecast using: raw spreads (top panel), normalized spreads which are spreads divided by the average spread level in each country (middle panel), and weighted spreads where the weights are determined by an unconditional regression of output growth on spreads for each country (lower panel). See text for further description.

$$severity_{i,t} = \alpha + b \times spread_{i,t} + \varepsilon_{i,t}$$

Dating	Outcome	b	se(b)	AdjR^2	N
ST	Peak to trough	-0.36	.25	9%	51
ST	Flexible depth	-0.31	.21	4%	50
ST	5 yr growth	-0.06	.04	4%	50
RR	Peak to trough	-0.07	.05	2%	91
RR	Flexible depth	-0.20	.09	7%	87
RR	5 yr growth	-0.04	0.18	3%	88

$$severity_{i,t} = \alpha + b \times spreadnorm_{i,t} + \varepsilon_{i,t}$$

Dating	Outcome	b	se(b)	AdjR^2	N
ST	Peak to trough	-2.2	.43	35%	51
ST	Flexible depth	-1.8	.44	14%	50
ST	5 yr growth	-0.3	.12	13%	50
RR	Peak to trough	-1.4	.70	14%	91
RR	Flexible depth	-1.3	.90	7%	87
RR	5 yr growth	-0.5	.26	1%	88

$$severity_{i,t} = \alpha + b \times weightedspread_{i,t} + \varepsilon_{i,t}$$

Dating	Outcome	b	se(b)	AdjR^2	N
ST	Peak to trough	5.3	1.02	27%	51
ST	Flexible depth	4.5	0.96	11%	50
ST	5 yr growth	1.00	0.25	16%	50
RR	Peak to trough	1.20	0.71	8%	91
RR	Flexible depth	2.04	1.12	12%	87
RR	5 yr growth	0.50	0.14	11%	88

Table 2: This table provides summary statistics on crisis episodes and compares those to typical non-financial recessions. We exclude recessions due to major world wars. The bottom panel provides summary statistics to episodes where our normalized spread measure increases by 1.5, providing an alternative crisis measure.

Financial Crises

Outcome variable	Mean	Median	Std Dev	P 10th	P 90th
Peak to trough	-6.2	-3.9	7.3	-0.5	-11.6
Time to trough	1.8	2	1.0	1	3
Time back to previous peak	4.6	4	2.7	2	10
Flexible depth	-8.9	-5.7	11.8	-1.0	-20.7
Flexible duration	3.4	2	2.8	1	7.5
Spread duration	1.6	1	1.8	1	3

Normal Recession (excluding world wars)

Outcome variable	Mean	Median	Std Dev	P 10th	P 90th
Peak to trough	-4.0	-2.4	5.8	-0.3	-7.4
Time to trough	1.5	1	0.9	1	3
Time back to previous peak	3.3	2	2.2	1	7
Flexible depth	-6.2	-3.0	8.6	-0.4	-13.8
Flexible duration	2.6	1	2.4	1	7
Spread duration	1.7	2	2.1	0	5

Spread crises

Outcome variable	Mean	Median	Std Dev	P 10th	P 90th
Peak to trough	-6.9	-0.2	11.7	0	-28.9
Time to trough	2.2	2	3	0	6
Time back to previous peak	2.6	1	2.7	1	8
Flexible depth	-7.9	-1.0	12.1	0	-30.1
Flexible duration	2.4	1	3.0	0	7
Spread duration	3.2	2.5	2.6	1	7

Table 3: This table provides regressions of future GDP growth on credit spreads at the 1 and 5 year horizon. We run these regressions unconditionally in the top panel. The middle panel includes dummies for crisis vs non-crisis times. The lower panel includes dummies for recessions. Standard errors clustered by year. Crisis and recession dates come from Schularick and Taylor (ST) unless otherwise noted.

Panel A: Unconditional

$$\ln\left(y_{i,t+k}/y_{i,t}\right) = \alpha_i + b \times spreadnorm_{i,t} + c' x_{i,t} + \varepsilon_{i,t}$$

Unconditional	b	se(b)	AdjR^2	N
1 year ($k=1$)	-0.23	0.11	1.4%	844
5 years ($k=5$)	-0.09	0.04	5.7%	844

Panel B: Crisis Dummies

$$\ln\left(y_{i,t+k}/y_{i,t}\right) = \alpha_i + b_1 \times 1_{nocrisis,t} \times spreadnorm_{i,t} + b_2 \times 1_{crisis,t} \times spreadnorm_{i,t} + b_3 \times 1_{crisis,t} + c' x_{i,t} + \varepsilon_{i,t}$$

	b_1	se(b)	b_2	se(b)	b_3	se(b)	AdjR^2	N
ST dates								
1 year ($k=1$)	-0.17	0.11	-0.71	0.32	-3.91	0.74	7.2%	844
5 years ($k=5$)	-0.07	0.05	-0.37	0.15	-0.91	0.35	8.4%	844
RR dates								
1 year ($k=1$)	-0.23	0.11	-0.46	0.38	-0.40	0.77	1.6%	844
5 years ($k=5$)	-0.08	0.05	-0.31	0.18	0.18	0.36	6.3%	844

Panel C: Recession Dummies

$$\ln\left(y_{i,t+k}/y_{i,t}\right) = \alpha_i + b_1 \times 1_{norecess,t} \times spreadnorm_{i,t} + b_2 \times 1_{recess,t} \times spreadnorm_{i,t} + b_3 \times 1_{recess,t} + c' x_{i,t} + \varepsilon_{i,t}$$

	b_1	se(b)	b_2	se(b)	b_3	se(b)	AdjR^2	N
All recessions								
1 year ($k=1$)	-0.18	0.10	-0.47	0.22	-5.10	0.74	2.8%	844
5 years ($k=5$)	-0.06	0.05	-0.26	0.11	-1.02	0.21	11.3%	844
Non-financial								
1 year ($k=1$)	-0.27	0.10	-0.18	0.32	-5.10	0.53	1.6%	844
5 years ($k=5$)	-0.10	0.05	-0.11	0.10	-1.04	0.26	8.8%	844

Table 4: This table provides regressions of our normalized spread measure on lagged spreads at the 1 and 5 year horizon conditional on crises and normal times. Standard errors clustered by year.

$$spreadnorm_{i,t+1} = \alpha_i + b \times spreadnorm_{i,t} + c'x_{i,t} + \varepsilon_{i,t}$$

	b	se(b)			AdjR^2	N
	0.58	0.03			33%	844

$$spreadnorm_{i,t+1} = \alpha_i + b_1 \times spreadnorm_{i,t} + b_2 \times 1_{crisis,t} \times spreadnorm_{i,t} + c'x_{i,t} + \varepsilon_{i,t}$$

	b_1	se(b)	b_2	se(b)	AdjR^2	N
ST	0.58	0.03	-0.17	0.09	34%	844
RR	0.58	0.03	-0.13	0.10	34%	844

$$\Delta spreadnorm_{i,t} = \alpha_i + b1_{crisis,t} + c'x_{i,t} + \varepsilon_{i,t}$$

	b	se(b)			AdjR^2	N
ST	0.42	0.26			1.0%	844
RR	0.22	0.20			1.0%	844

Table 5: Robustness. We provide various robustness checks of our results. We add time fixed effects and control for the change in leverage measure provided by Schularick and Taylor, where leverage measures the total amount of private sector credit in the economy.

Panel A: Controls and fixed effects in unconditional regressions

$$\ln(y_{i,t+k}/y_{i,t}) = \alpha_i + b \times spreadnorm_{i,t} + c' x_{i,t} + \varepsilon_{i,t}$$

	$k = 1$ year				$k = 5$ years			
Variables	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
spreadnorm	-0.22*	-0.23**	-0.50***	-0.50***	-0.16***	-0.17***	-0.17***	-0.17**
ΔSTLev		-0.05		-0.02		-0.09***		-0.07***
Country FE	Y	Y	Y	Y	Y	Y	Y	Y
Time FE			Y	Y			Y	Y
AdjR^2	0.02	0.02	0.32	0.32	0.13	0.13	0.45	0.46
N	648	648	648	648	648	648	648	648

Panel B: Controls conditional on crises

$$severity_{i,t} = \alpha + b \times spreadnorm_{i,t} + c' x_{i,t} + \varepsilon_{i,t}$$

	All Recessions		Fin Crises (RR)		Fin Crises (ST)		Non-Fin Recession	
Variables	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
spreadnorm	-2.0***	-1.9***	-1.9*	-2.1**	-2.3***	-2.1***	-1.5	-21.5
ΔSTLev		-0.17		-0.36***		-0.40**		0.17
AdjR^2	16%	17%	34%	44%	40%	52%	6%	7%
N	124	124	38	38	40	40	87	87

Table 6: Spreads vs lagged spreads. This table explores the information contained in contemporaneous vs lagged spreads for output. It repeats our main regressions using lags of spreads and changes in spreads.

	$\ln\left(y_{i,t+k}/y_{i,t}\right) = \alpha_i + b \times spreadnorm_{i,t} + c'x_{i,t} + \varepsilon_{i,t}$							
	$k = 1$ year				$k = 5$ years			
Variables	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
$spreadnorm_{i,t}$	-0.23**		-0.20*	-0.48***	-0.09**		-0.15**	-0.15**
$spreadnorm_{i,t-1}$	0.19		0.18	0.19	0.00		0.01	0.05
$spreadnorm_{i,t-2}$			-0.21	-0.15			0.01	0.02
$spreadnorm_{i,t-3}$			0.01	0.01			-0.03	-0.03
$\Delta spreadnorm_{i,t}$		-0.21**				-0.05		
Δlev			-0.10**	-0.02			-0.09***	-0.06***
Country FE	Y	Y	Y	Y	Y	Y	Y	Y
Time FE				Y				Y
AdjR^2	2.1%	1.4%	5.3%	35.2%	6.0%	5.7%	13.0%	49.0%
N	844	844	598	598	844	844	598	598

	$severity_{i,t} = \alpha + b \times spreadnorm_{i,t} + c'x_{i,t} + \varepsilon_{i,t}$							
	All Recessions		Fin Crises (RR)		Fin Crises (ST)		Non-Fin Recession	
Variables	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
$spreadnorm_{i,t}$	-1.6**		-1.3*		-2.5**		-1.1	
$spreadnorm_{i,t-1}$	-0.1		0.0		0.4		-0.2	
$\Delta spreadnorm_{i,t}$		-0.4		-1.3*		-2.6*		-0.3
AdjR^2	13%	1%	12%	3%	35%	11%	4%	1%
N	172	172	89	89	51	51	121	121

Table 7: Regressions based on crises in credit spreads. We define crises using our continuous credit spread variable. This table shows the ability of spreads to forecast output outcomes during these credit spread crises.

Panel A: Unconditional

$$\ln\left(y_{i,t+k}/y_{i,t}\right) = \alpha_i + b_1 \times 1_{nocrisis,t} \times spreadnorm_{i,t} + b_2 \times 1_{crisis,t} \times spreadnorm_{i,t} + b_3 \times 1_{crisis,t} + c' x_{i,t} + \varepsilon_{i,t}$$

Our crisis dates	b_1	se(b)	b_2	se(b)	b_3	se(b)	AdjR^2	N
1 year ($k = 1$)	-0.07	0.18	-0.33	0.16	0.29	1.04	1.6%	844
5 years ($k = 5$)	0.10	0.09	-0.16	0.08	-0.16	0.49	7.0%	844

Panel B: Spread crisis

$$severity_{i,t} = \alpha + b \times spreadnorm_{i,t} + \varepsilon_{i,t}$$

Outcome	b	se(b)	AdjR^2	N
Peak to trough	-1.14	.28	16%	86
Flexible depth	-1.06	.30	13%	86
5 yr growth	-0.20	.07	7%	86
1 yr growth	-0.21	.19	2%	86

Panel C: Overlap in Dating Measures

	ST	RR	Spread
ST	51	31	15
RR	31	50	17
Spread	15	17	56

Table 8: Quantile Regressions. We run quantile regressions of output growth on spreads at both 1 and 5 year horizons. We vary quantiles and find spreads primarily forecast lower quantiles of GDP growth.

Quantile Regressions				
$\ln\left(y_{i,t+k}/y_{i,t}\right) = \alpha_i + b \times spreadnorm_{i,t} + \varepsilon_{i,t}$				
	$k = 1$ year		$k = 5$ years	
Quantile	b	se(b)	b	se(b)
0.9	0.14	0.36	0.10	0.33
0.75	-0.02	0.14	0.06	0.06
0.5	-0.11	0.14	-0.06	0.05
0.25	-0.49	0.15	-0.10	0.05
0.1	-0.77	0.28	-0.40	0.12