How Much SRISK Is Too Much?*

Robert Engle[†] Tianyue Ruan[‡]

September 27, 2018

Abstract

When financial firms are under-capitalized, they are vulnerable to external shocks. This is commonly measured by stress tests or market-based measures of systemic risk such as SRISK. The natural response to such vulnerability is to raise capital, and this can endogenously start a financial crisis. Excessive credit growth can be interpreted as under-capitalization of the financial sector. Hence, we can assess how much SRISK an economy can stand, and measure the probability of a crisis. Using a crisis severity variable constructed by Romer and Romer (2017), we estimate a Tobit model for 23 developed economies. We develop a *probability of crisis* measure and an *SRISK capacity* measure from the Tobit estimates. These reveal the important global externality whereby the risk of a crisis in one country is strongly influenced by the under-capitalization of the rest of the world.

*The authors would like to thank the Sloan foundation, the National Science Foundation, the Macro Financial Modeling Group under the leadership of Lars Hansen and Andy Lo, and NYU Stern for financial support to the Volatility Institute. Further thanks go to the Armellino Foundation, Blackrock, the Global Risk Institute, the Norwegian Finance Initiative, the Inter-American Development Bank, and ITG for additional support. The authors are indebted to many seminar participants from the Bernoulli Lecture at EPFL, Sapienza University in Rome, SoFiE 10th Annual Conference, the Slovak Economic Association, Central European University, Hong Kong Polytechnic University, Zhejiang University, VINS, IFABS at Ningbo, the Inter-American Development Bank, the NBER/NSF Time Series Conference at Northwestern University, the Royal Danish Academy, the MFM Winter 2018 Meeting, and NYU Stern. Particularly helpful comments from Darrell Duffie, Eric Jondeau, Diane Pierret, Marina Brogi, Lasse Pedersen, Soren Johansen, Katarina Juselius, Mark Gertler, Ed Altman, Richard Berner, Kim Schoenholtz, Viral Acharya, Ralph Koijen, Markus Brunnermeier, Tobias Adrian, Matt Richardson, Thomas Philippon, Philipp Schnabl, and many others are deeply appreciated.

[†]Stern School of Business, New York University, rengle@stern.nyu.edu.

[‡]NUS Business School, National University of Singapore, tianyue.ruan@nus.edu.sg.

Electronic copy available at: https://ssrn.com/abstract=3108269

1 Introduction

When financial institutions are under-capitalized, they are vulnerable to external shocks. Even more importantly, however, they may become the source of these shocks through normal capital raising behavior. We use the popular SRISK data to measure under-capitalization of individual financial institutions and countries. Because it is widely believed that excessive credit growth is a cause of financial crises it is natural to motivate our SRISK measure as a measure of excessive credit growth. When under-capitalization is extreme, these endogenous shocks become sufficient to cause an economic downturn which we call a financial crisis. In this paper, we develop and estimate a model of the impact of under-capitalization on the *probability of crisis*. This model can be solved to obtain a corresponding *SRISK capacity* which would keep this probability below 50% as long as SRISK remains below this level.

The process by which under-capitalization leads to a financial crisis has been widely studied in the theoretical macro-finance literature and to some extent in the empirical literature. These models feature asset fire sales and credit rationing in various forms. We borrow from this literature suggestions on the relevant measures of under-capitalization, and for the first time, come up with quantitative estimates of how much systemic risk it takes to generate a financial crisis.

Using crisis severity measures created by Romer and Romer (2017), we estimate the level of under-capitalization that precipitates a financial crisis. We examine 23 developed economies over time and seek a quantitative measure of the probability of a crisis as a function of the aggregate capital shortfall and other variables. From these estimates can then compute *SRISK capacity*.

This model of systemic risk features two widely recognized externalities. The risk of a financial crisis in a country depends upon the total capital shortfall of the financial

Electronic copy available at: https://ssrn.com/abstract=3108269

sector in this country. Thus, a single risky firm will not be systemic unless it is extremely large. Any firm that reduces its capital shortfall will benefit all firms and the economy as a whole. Similarly, the risk of one country depends upon the aggregate SRISK of the rest of the world. Hence, a country that improves its regulation or reduces its financial sector under-capitalization will benefit other countries and global financial stability. This setting clearly requires regulation and cooperation to achieve optimal performance.

The paper that is closest to estimating aggregate crisis probability is Adrian, Boyarchenko, and Domenico (2016). They examine how financial conditions in the US affect US GDP and discover that weak financial conditions lead to a risk of GDP downturns. However, they do not calculate the probability of a financial crisis or motivate the use of the broad financial conditions index which is composed of more than 100 series. And because they only examine the US and use this index, they do not incorporate either of these externalities.

The paper proceeds as follows: Section 2 motivates our measure of excessive credit growth. Section 3 introduces SRISK and discusses details of its construction. Section 4 describes the mechanism that generates fire sales and suggests measures of the intensity of such a spiral. Section 5 describes SRISK data and the crisis severity measure of Romer and Romer (2017). Section 6 presents the empirical models and estimates on the level of SRISK that induces a financial crisis. Section 7 shows the probability of crisis and SRISK capacity measures constructed from empirical estimates and discusses the global externalities. Section 8 provides robustness analyses. Section 9 concludes.

2 Excessive Credit Growth

It is widely believed that financial crises result from *excessive* credit growth. See, for instance, Reinhart and Rogoff (2009), who claim that this time is *not* different, Borio (2014)

Electronic copy available at: https://ssrn.com/abstract=3108269

and Drehmann, Borio, and Tsatsaronis (2012) on financial cycles, Adrian and Shin (2010, 2014) on leverage cycles, and Schularick and Taylor (2012) on the predictive power of credit growth. However, the challenge is how to measure excessive credit growth. We will argue that credit growth is excessive if the financial sector does not have sufficient capital to cover market value losses in a downturn. This is consistent with the notion that at the end of a credit cycle, increasingly risky credit will be issued and the holders of this credit will be leveraged financial institutions with insufficient capital to cover losses in a downturn. This is how a "credit boom goes bust." We estimate the dollar amount of capital that a financial firm would need to raise in order to function normally if we have another financial crisis. This is our widely used measure called SRISK which is computed weekly and published on NYU Stern's Volatility Laboratory (V-Lab).[1] It measures capital shortfall under stress and is similar to a regulatory stress test. It was initially introduced in Acharya, Brownlees, Farazmand, and Richardson (2011) and Acharya, Engle, and Richardson (2012), and then expanded in Brownlees and Engle (2017). The dollar amount of capital shortfall can be summed for an entire financial sector to provide a measure of systemic risk.

In order to implement this model, we must distinguish between productive credit growth and excessive credit growth. As credit typically grows with output, various measures of excessive growth have been proposed focused on whether it grows much faster than output. Instead, we focus on the quality and risk of firms extending and ultimately holding credit.

Consider a simple example of a bank that holds mortgages. This bank may eventually at the end of a credit cycle, lend to underqualified borrowers and overvalued houses. In this case, the mortgages will actually be worth less than the accounting values and the bank's ratio of market value to book value will fall. This can be seen in the equity value

[1]See https://vlab.stern.nyu.edu/welcome/risk/.

Electronic copy available at: https://ssrn.com/abstract=3108269

of the firm as well as mark-to-market valuation of the loans. If there is a downturn in the housing market, the value of the mortgages will fall further as the collateral loses value and the borrowers become weaker. The firm may have to cover these losses from its capital, and if the capital is inadequate, the value of the assets of the bank may fall below its liabilities. In this case, stock market valuation will collapse, and the bank may seek a bailout in order to continue functioning. Thus excessive credit growth can be measured by the capital shortfall of the financial sector.

From this example, it is natural to examine the valuation of financial sector assets relative to their liabilities. If the assets are undervalued and risky, we would expect to see low market-to-book ratios and high equity volatility. If this is a systemic problem as opposed to a single firm, we would expect to see high correlations with market-wide events. This is measured by the beta of the firm. In a stress scenario, the broad market will decline, and the impact on capital at each firm will depend upon its beta. And in fact, the betas will differ depending upon the asset holdings as the market is well aware of these effects.

Thus the task is to estimate the capital that would be needed under stress for each firm in order to continue to operate.

3 Methodology to Calculate SRISK

Developed by Acharya et al. (2011, 2012) and Brownlees and Engle (2011, 2017), SRISK is the capital that a financial firm would need to raise in order to continue to function normally if we have another financial crisis. Because it is difficult to raise capital in a financial crisis, this capital shortfall will either be met mostly by the taxpayer money or the firm will cease to function normally and may fail. For this reason, the measure is considered to be an indicator of systemic risk in much the same way as are supervisory

Electronic copy available at: https://ssrn.com/abstract=3108269

stress tests.

Normal operation of a financial firm requires that its market capital ratio (its market value of equity divided by the sum of the book value of liabilities and the market value of equity) be above the prudential capital ratio. Let k denote this prudential ratio. The capital shortfall of a financial firm at horizon T can therefore be computed as:

$$Capital\ Shortfall = k\left(Debt_{t+T} - Equity_{t+T}\right) - Equity_{t+T} \tag{1}$$

SRISK is defined as the median capital shortfall conditional on a financial crisis. To estimate SRISK, we use a bivariate daily time series model of equity returns of the firm and of a broad market index that incorporates asymmetric volatility, time-varying correlation, and asynchronous trading. We describe this model in greater detail below. In addition, we use $k = 8\%$ which corresponds to the typical leverage ratio of well managed financial firms in tranquil periods.[2,3] We also make the simplifying assumption that $Debt_t = Debt_{t+T}$ as the notional value of liabilities is not likely to change with the stress but the value of equity will.

To forecast the capital shortfall, we use a standard financial approach, the market regression, which in its simplest form can be expressed as

$$r_t^f = \beta_t^f R_t^M + \varepsilon_t^f \tag{2}$$

where r_t^f is the equity return on day t for firm f. Similarly, R_t^M is the return on a global equity index. This relation captures the market view of the rate at which falling asset values lead to equity declines when the market collapses. It explicitly focuses on the

[2]In June 2017 for example the average of the six largest US bank leverage ratios was 9.1, in June 2014 it was 11.3, in June 2010 it was 15.5, and in June 2005 it was 9.5. The average of these is 11.3 or slightly less leverage than used in SRISK. Users of V-Lab have the option of setting the capital ratio at any desired level.

[3]Firms using IFRS accounting rather than GAAP typically have a bigger balance sheet as there is less netting of derivatives. Consequently, we use 5.5% for all European firms.

Electronic copy available at: https://ssrn.com/abstract=3108269

comovement of this firm and the market, rather than just the volatility of the firm. This is why this is a macro-prudential measure of risk rather than micro-prudential.

Let p_t^M denote the equity price of the firm at the end of day t and P_t^M the market analogue. The cumulative fractional return of the firm over the horizon of T can be approximated with the same form of the 1-day-ahead forecast if the beta has almost a unit root and therefore the beta can be factored out of the sum.[4]

$$\begin{aligned} \frac{p_{t+T}^f - p_t^f}{p_t^f} &= \exp\left(\sum_{j=1}^{T}\left(\beta_{t+j}^f P_{t+j}^M + \varepsilon_{t+j}^f\right)\right) \\ &\approx \exp\left(\left(\beta_t^f\right)\left(\log\left(P_{t+T}^M / P_t^M\right)\right)\sum_{j=1}^{T}\left(\varepsilon_{t+j}^f\right)\right) \end{aligned} \tag{3}$$

The cumulative fractional return is itself a random variable; we consider its median value conditional on a stressed market return. Under the assumption that the idiosyncratic errors have a zero median, the median cumulative fractional return in a crisis can be expressed in the following way, where θ corresponds to the level of stress we consider.

$$\begin{aligned} Median\left(\frac{p_{t+T}^f - p_t^f}{p_t^f}\right) &= \exp\left(\left(\beta_t^f\right)\log\left(1-\theta\right)\right) \\ \frac{P_{t+T}^M - P_t^M}{P_t^M} &= -\theta \end{aligned} \tag{4}$$

We modify (2) to take into account two important considerations: asynchronous trading and time-varying beta.

1. *Asynchronous trading.* The world market index we use is the MSCI ACWI ETF. As this is traded in the US, the closing price is the price at the NY close, and it reflects the information that the market knows at that time. For firms located in different time zones,

[4] An alternative simulation approach to forecast the cumulative firm return has been implemented in Brownlees and Engle (2011, 2017) and in V-Lab's MESSIM.

Electronic copy available at: https://ssrn.com/abstract=3108269

part of the response will appear to be from the previous day return. Hence

$$r_t^f = \beta_t^f R_t^M + \gamma_t^f R_{t-1}^M + \varepsilon_t^f \tag{5}$$

2. *Time-varying beta.* The Dynamic Conditional Beta (DCB) model proposed by Engle (2016) constructs a time-varying beta by allowing volatility and correlation to be time-varying with GARCH and DCC. To allow for both the constant beta and the time-varying beta flexibly, we adopt the following model:

$$r_t^f = \left(\phi_1 + \phi_2 \hat{\beta}_t^f\right) R_t^M + \left(\phi_3 + \phi_4 \hat{\gamma}_t^f\right) R_{t-1}^M + \varepsilon_t^f \tag{6}$$

where $\hat{\beta}$ and $\hat{\gamma}$ are computed from the asymmetric firm and market volatility estimates (GJR-GARCH) and a time-varying correlation estimate (DCC). The four coefficients of this model can be estimated assuming a GJR-GARCH error term.

With these two modifications, the (total) beta and SRISK of the firm are:

$$\tilde{\beta} = \left(\hat{\phi}_1 + \hat{\phi}_2 \hat{\beta}_t^f\right) + \left(\hat{\phi}_3 + \hat{\phi}_4 \hat{\gamma}_t^f\right) \tag{7}$$

$$SRISK_t = kDebt_t - (1-k)Equity_t \exp\left(\tilde{\beta}\log(1-\theta)\right) \tag{8}$$

We consider T to be six months in the future and calibrate the market stress level $\theta$to be 40% as the MSCI index declined approximately 40% over six months during the Global Financial Crisis. By using a one-factor stress, we reduce the possibility that we will miss the cause of the next financial crisis. Whether the crisis is caused by a housing market collapse, sovereign debt collapse, commodity price collapse, exchange rate collapse, government shut down or derivative market failure, the stock market is probably going to fall in anticipation of a decline in the real economy. It is almost inconceivable that a financial crisis could occur without a substantial fall in the stock market.

Electronic copy available at: https://ssrn.com/abstract=3108269

In measuring the value of debt for insurance companies, we make adjustment for separate accounts. Separate accounts are variable annuities and other investments made by insurance clients which are shown both as assets and liabilities of the firm. Consequently, fluctuations in value do not affect the book equity of the firm but they do affect the size of the balance sheet. Insurance companies often argue that the separate accounts assets and liabilities should be excluded for calculating the capital ratio as they do not really belong on the balance sheet. Indeed, the treatment of separate accounts affect the capitalization measures greatly: Separate accounts amount to about 40% of total liabilities of insurance companies. In December 2015, the median leverage among 12 U.S. life insurers was 8.1 excluding separate accounts entirely and 16.4 including them. In assessing the impact of separate accounts, we settle on a partial exclusion of separate accounts in both assets and liabilities due to the penalty associated with early withdrawal and the complicated guarantee structure.[5] Including 40% of separate accounts would make the median leverage 11.6 in December 2015, close to the leverage of conservatively managed banks in ordinary times that we use to calibrate the prudential ratio k. Under the premise that December 2015 is a calm time, we include 40% of separate accounts in total liabilities to calculate SRISK for insurance companies.

Interestingly, the firm betas reflect the characteristics of the crisis and their impact on each firm. For example, in the Great Recession, the beta of Bank of America and Citigroup rose to three and four while neither Goldman Sachs nor BNP Paribas moved much at all.

[5]Although the separate accounts assets belong to the investors rather than the insurance companies, many separate accounts contain explicit guarantees of return, cumulative return, distribution, and other features. Under GAAP, firms are required to report the value of these guarantees as "future policy benefits." These are estimated based on values of derivatives and hedges and management judgment. Data from SNL Financial show that future policy benefits are roughly two-thirds of total separate accounts for the large US insurance companies. There are reasons to be concerned about the accuracy of these estimates of future policy benefits, see Drexler, Plestis, and Rosen (2017) and Koijen and Yogo (2018) for details. That future policy benefits are estimated at the current time rather than at stressed times is relevant for our analysis. Under a stress scenario it is natural that the liabilities would increase. Empirically these appear not to have increased during the financial crisis even though the value of the separate accounts fell substantially. For some companies there is a remarkable stability of these future policy benefits despite rapid growth in separate accounts.

Electronic copy available at: https://ssrn.com/abstract=3108269

They did not have direct balance sheet exposures to subprime mortgages. However, in the European Sovereign Debt Crisis, Credit Agricole and BNP's betas rose almost to three as these banks were very exposed to Greek and other peripheral countries' debt.

For each country, the aggregate SRISK is the sum of all the financial firms with positive values. This is the sum of all the institutions that would have to raise capital in a crisis. We do not net the overcapitalized firms as the capital from these prudent firms is not immediately available to re-capitalize the weak firms. Mergers do happen, but more often, the well-capitalized firms protect their capital as bad events are approaching.

4 Deleveraging Cycles

When SRISK is high, either the regulator or the risk manager of individual companies will compel the firms to strengthen their balance sheet. This can be done in many ways but the most commonly-used approaches involve sales of assets or equity. When such sales are in large volume, it is inevitable that there will be a price impact which is commonly called a "fire sale externality." This leads to a downward spiral of the financial sector and ultimately of the economy that has been written about extensively. See for example Cont and Schaanning (2016), Greenwood, Landier, and Thesmar (2015), Pedersen (2009) among others. The initial conditions that make such a fire sale likely are precisely a large quantity of goods to be sold in a hurry without sufficiently many willing buyers.

Although we focus on the downward spiral in a fire sale, a similar process can be seen in the upward direction. Ruan (2017) calls this the "race to the top." Deregulation in this context allows financial institutions to increase their risk and leverage thereby making the system more likely to have a financial crisis in the future, although less likely in the short run. Thus, the models here are essentially a theory of endogenous financial cycles.

Electronic copy available at: https://ssrn.com/abstract=3108269

There are many ways firms can reduce leverage. In this paper, we consider three. In each approach, there is a corresponding natural measure of the severity of systemic risk. When firms are urged to reduce their systemic risk either by risk managers or by regulators, they can do nothing and wait for growth to increase their equity and if it doesn't, appeal for a bailout. A second alternative is to sell new shares of stock (or equivalently reduce dividends). The third strategy is to sell existing assets and use the proceeds to reduce debt.

If firms choose to do nothing, then the cost to the regulator is the loss of GDP that would be required in a bailout. Thus, the natural measure of the size of the risk is SRISK/GDP. When this ratio is high, the risk to the economy is very high if firms choose this strategy. If firms choose to sell new shares to raise capital or reduce dividends, they may lower the values of existing shares. That is partly a signaling effect whereby the signal that the bank must raise capital conveys information that the firm is in trouble. It may also simply be a supply and demand effect where the more shares are in existence, the lower the price. In either case, the larger the volume of shares that needs to be sold, the bigger this effect. Hence if the value of shares that need to be sold is a large fraction of the shares that are already outstanding, the price impact is likely to be large, and firms will hesitate to use this channel. Hence a natural measure of excessive risk is SRISK/MV where MV is the market capitalization.

If the firm chooses to sell assets and the total amount contemplated is small compared with the stock of assets, then asset sales are likely to be cost-effective. Thus, financial institutions may choose between these approaches depending upon the ratio of assets to market cap in the financial sector as a whole. This measure of leverage would imply that firms would be more likely to raise equity when leverage is low and sell assets when leverage is high. Since financial crises often coincide with periods of high leverage, the sale of assets is a common approach. Furthermore, the well-known debt overhang prob-

Electronic copy available at: https://ssrn.com/abstract=3108269

lem first described in the seminal paper by Myers (1977) becomes more pronounced in these periods: The potential increase in the value of assets should a firm decide to hold on to its assets mostly accrues to existing debt holders. A firm that maximizes equity value may therefore choose to sell assets rather than raise capital to deleverage.

Asset sales have a damaging feature that has been widely discussed. Large asset sales are likely to depress the price of assets which will, in turn, increase the leverage of all financial firms holding similar assets. This is often called a leverage spiral which means that deleveraging by selling assets may even be counterproductive in the extreme but at minimum will require more sales than initially anticipated. Frequently this leverage spiral is called a "fire sale" and may lead to asset sales at prices below their fundamental value. However, the same phenomenon is called price impact in market microstructure and is a description of expectations that are reduced by observations of selling pressure. In this case, the price is depressed because market participants reduce their expectations of future value.

In either context, there are many sellers and insufficient quantities of motivated and well-capitalized buyers to prevent the price from declining. As a consequence, firms experience a capital loss on assets. Investors can see this prospect and will reduce their equity valuations in advance of the mechanical process of selling assets. Similarly, the real economy that depends upon credit extended by the financial sector will find that they are competing with the small number of buyers for financial assets leading to excessive costs of credit. Furthermore, the negative expectations of the economic future extracted from the asset selling will also make borrowers look increasingly risky. Just when enterprises need capital, the financial markets will deny it.

The natural risk measure if firms choose to deleverage by selling assets is, therefore, the ratio of assets for sale over total assets. This can be computed analytically for the financial sector as a whole if there is no price impact. In order to reduce SRISK to zero,

Electronic copy available at: https://ssrn.com/abstract=3108269

A^{sales} of assets are sold and debt is retired. If there is no price impact, there would be no effect on equity, therefore

$$0 = k(Debt_t - A^{sales}) - (1-k)Equity_t \exp\left(\tilde{\beta}\log(1-\theta)\right) \tag{9}$$

$$A^{sales} = SRISK/k \tag{10}$$

Thus the natural measure of the size of SRISK which is dangerous is SRISK/(Total Assets/k) which is interpreted as A^{sales}/A^{total} and will be denoted by SRISK/(TA*k). Consistent with our partial inclusion of separate accounts for calculating SRISK, only 40% of separate accounts are included in total assets.

Interestingly, in the Ruan (2017) framework, this leverage spiral can work in reverse. When financial firms have excess capital, they can build leverage by buying assets and incurring price impact. In this case price impact is a positive contribution to firm value and investors will presumably recognize this in advance as well. In this situation, firms appear well capitalized but have rising leverage. This is an explanation for the market valuations of financial firms before the financial crisis and maybe also after the Trump election.

5 Data

Data on SRISK/GDP, SRISK/MV, and SRISK/TA varies across countries and over time. Recent snapshots of these aggregate statistics are given in Figures 1, 2 and 3 which show the countries with the highest values of each measure. Notice that Japan is the highest for all three measures, although this has not generally been the case. We see that the capital shortfall of Japan is about 15% of GDP so the cost of bailing out the banks in another financial crisis would be enormous.

Electronic copy available at: https://ssrn.com/abstract=3108269

In Figure 2, the new capital that would be needed by these institutions is roughly 100% of the current market value of the firms. Hence, the stock market value of these firms would need to double through selling new shares in order to eliminate the capital shortfall. Selling this volume of new shares would surely depress the price substantially especially if undertaken in a crisis.

Finally, the ratio of SRISK to total assets in Japan is a little over 5%. Since the capital ratio is 8% for Japan, this means that 5/8 of the financial sector assets would need to be sold in order to reduce SRISK to zero. Again, this is likely to start a severe fire sale in assets which will reduce their market value substantially through the price impact of the sales. Furthermore, firms that are selling assets such as corporate loans or mortgages will be unlikely to issue new credit, and therefore there is likely to be substantial credit disruption.

To compare these three measures as indications of how much SRISK is too much, we use the Romer and Romer (2017) crisis measure. This carefully constructed measure not only records whether there is a financial crisis but also assesses how severe is the crisis. For each country and for every six months, the two authors read the OECD Economic Outlook and classify the nature of any financial crisis. According to their classification, the primary feature underlying a financial crisis is a disruption in credit supply. Their measure of crisis severity is on a scale of 1 to 15 which goes from a "credit disruption minus" to "extreme crisis plus." This measure is constructed for 24 developed economies for the period from 2000 to 2012. If there is no crisis, the measure is zero; if the measure is greater than 3, it becomes more than a "minor credit disruption."

We use these crisis measures for 23 of their countries. We do not include Iceland because it still does not have any publicly traded banks. During the 2007—09 Global Financial Crisis, all of them failed. We average the monthly SRISK figures over each six month period for each of the countries. The crisis severity measure is shown for all

Electronic copy available at: https://ssrn.com/abstract=3108269

countries in Figure 4. As can be seen there are substantial differences across countries. For most countries, there is a peak in 2008—9 but in some it continues to rise and for others there was a financial crisis in 2002—3 as well. For some, the crisis is more severe than for others.

6 How Much SRISK Is Too Much?

The goal is to explain the cross-section and time-series of crisis severity with these systemic risk measures. In this way, we can learn how much SRISK is associated with a crisis.

The first model is a panel regression of this crisis measure on the three SRISK measures including time and country fixed effects. In Table 1 the first column shows that the most useful variables are SRISK/(TA*k) and SRISK/MV as the measure divided by GDP is negative. The negative sign could be due to regulatory forbearance when the costs of a bailout are very high. When regulators exercise "regulatory forbearance" and don't force the deleveraging to begin, they could delay a crisis although possibly making it more serious when it happens.

The remaining columns replace the time fixed effects with various measures of world capital shortfall. The results are similar to the fixed effects model and show that some measure of world SRISK is correlated with the time fixed effects.

Running the same regression as a predictive regression (Table 2) reinforces this conclusion. With a lagged dependent variable and fixed effects, SRISK/(TA*k) remains significant. However, with measures of world SRISK, its significance is reduced.

The regression does not take account of the fact that more than half of the dependent variable observations are zero. The relation between crisis severity and SRISK is naturally

Electronic copy available at: https://ssrn.com/abstract=3108269

a hockey stick rather than a straight line. In fact, the preferred estimator for this model is a Tobit which recognizes that the dependent variable is truncated at zero.

A Tobit model is defined in terms of a latent variable y_l which depends upon explanatory variables X and a disturbance. The observed dependent variable, y, is a truncated version of y_l. Under the assumption that the error term follows a standard normal distribution, the model can be expressed by two equations as follows:

$$y = \begin{cases} y_l \text{ if } y_l > 0 \\ 0 \text{ otherwise} \end{cases}$$

$$y_l = X\beta + \sigma\varepsilon,\ \varepsilon \sim \mathcal{N}(0, 1)$$

We estimate the Tobit model with country fixed effects to allow the possibility that countries will differ in the tolerable level of SRISK. This may be due to institutional markets for selling assets and pools of investors who might be willing to step in even as a crisis is approaching. It may also be due to differences in the likelihood of a government rescue that would protect both financial firms and those buying assets.

We consider a domestic model that only uses country-level SRISK variables to explain crisis severity and a global model that expands the set of explanatory variables with world SRISK variables. The motivation of the global model comes from the observed co-movement in crisis severity across countries in Figure 4 and the correlation between world SRISK variables and the estimated time fixed effects in Table 1. To better capture the externality aspect of financial crises, we modify how these world variables are constructed. For each country, the world SRISK variables are calculated using the sum of the respective country-level variables across all other countries, which we refer to as leave-out-sums. This modification also facilitates the SRISK capacity measure developed later.

Electronic copy available at: https://ssrn.com/abstract=3108269

The estimation results are reported in Table 3. The SRISK/(TA*k) variable is highly significant in either the domestic model or the global model. Columns (1) and (2) are the specifications with the best Schwarz criterion among many specifications including many not reported here for the domestic and global models, respectively.

Estimates of country fixed effects, omitted from Table 3 for brevity, are tabulated in Table 4. For the global model, the fixed effects reflect individual countries' resistance to a crisis with equal values of SRISK. These range from -17.9 for Belgium to -5.4 for Turkey. Thus, with equal characteristics, Turkey is much more likely to have a crisis than Belgium. Interestingly, Japan is in the middle at -10.1.

Based on the Tobit model, we can assess the distance from a financial crisis quantitatively. We propose two measures for this quantitative assessment. The first one is a probability of a crisis. Following the Tobit model, such measure can be expressed as the probability that the dependent variable will exceed a value $q > 0$ conditional on X for a meaningful value of q. We choose $q = 4$ which corresponds to a mild crisis under the Romer and Romer (2017) classification.

$$ProbCrisis = \Pr(y > 4|X) = \Pr(y_l > 4|X) = \Pr\left(\varepsilon > \frac{4 - X\beta}{\sigma}\middle| X\right) = 1 - \Phi\left(\frac{4 - X\beta}{\sigma}\right) \tag{11}$$

The second measure gauges whether there is a level of SRISK that makes the probability of a crisis just 50%. From 11 we see that the probability of a crisis is a half when $X\hat{\beta} = 4$. We can solve for a *SRISK capacity* that corresponds to this level of risk, holding everything else constant. Here $\hat{\beta_1}$ is the estimated (combined) coefficient on SRISK/(TA*k) in the Tobit estimation. In the domestic model, since both SRISK/(TA*k) and its lag are included in the domestic model, $\hat{\beta_1}$ is the sum of their coefficients or 24.917. In the global model, $\hat{\beta_1}$

Electronic copy available at: https://ssrn.com/abstract=3108269

is the coefficient of country SRISK/(TA*k) or 13.165.

$$SRISK\_CAPACITY = SRISK + \frac{4 - X\hat{\beta}}{\hat{\beta}_1} \times k \times TA \tag{12}$$

We compute and analyze these two measures at the monthly frequency. To reconcile with the half-yearly frequency of the estimation sample, we use the six-month moving average of the explanatory variables to construct these two measures.

7 Results and Discussions

From these regressions, we compute the *probability of crisis* and the *SRISK capacity*. The *probability of crisis* is computed from both the domestic model (Column 1) and the global model (Column 2) from Table 3. Similarly, the *SRISK capacity* is computed from both equations. Figure 5 and 6 plot these two measures for the US. Similar figures are included in the Appendix A: Measures for Other Countries for all countries studied by Romer and Romer (2017) with the exception of Iceland.

From Figure 5, we can see that the domestic and global models give rather similar estimates for the probability that the US is in a crisis over the 17-year period although the peaks are a little higher in the global model. In 2008 the probability rose to 80% or 90% whereas it was only about 60% in the European sovereign debt crisis and in the recent period has fallen to less than 10%. In contrast, many of the European countries had a greater peak in 2011 than in 2008.

Figure 6 gives more information on the systemic risk in the US. The solid line is the SRISK for the US. Whenever this exceeds the capacity, the probability of crisis rises above 50%. Notice that the dashed line is rising over time indicating that the capacity of the US to avoid a financial crisis is rising, and in this case it is a result of the increase in total

Electronic copy available at: https://ssrn.com/abstract=3108269

assets in the financial sector.

Notably, the *SRISK capacity* obtained from the global model dives in 2008–09 making a crisis in the US more likely. Similarly, it declines dramatically in 2011-12 and 2016. In a similar vein, the probability of crisis obtained from both models peaks at the same time. These three incidences, not unique to the US, correspond to the three crisis episodes since 2000 and reflect the global nature of a financial crisis. Let us briefly review these three episodes and summarize their common features to emphasize the importance of capturing this nature.

The time series dynamics of world SRISK since 2000 (Figure 7) reveals three peaks. The magnitude of the peak is close to $4 trillion in each case and greatly exceeds the SRISK during the first seven years of this century. The first two peaks correspond to two well-known crisis episodes, the Global Financial Crisis and the European sovereign debt crisis. The third peak in 2016–17 might be called the Asian debt crisis with Japan and China as the two biggest contributors. What is common in all three episodes is that banks massively increase their holdings of perceived riskless debt and subsequently experience stress in one or at most a handful of countries, and such financial stress spreads from these countries to other countries.

The first episode, the Global Financial Crisis, is widely perceived to be tied to the housing sector which experienced a rapid increase in the five years before the crisis. The housing price boom was fueled by a rapidly growing mortgage market that employed financial engineering and securitization to fund ever declining quality of mortgages from an international pool of investment capital. The magic of CDO (collateral debt obligation) securitization was that investors and ratings agencies and regulators all regarded the senior tranches as nearly riskless, regardless of the quality of the component mortgages. As the housing price started to fall, the previously neglected risks became apparent. SRISK of all US financial firms increased from the end of 2006 to August 2008, and such an increase

Electronic copy available at: https://ssrn.com/abstract=3108269

was most substantial for the firms that were big participants in the mortgage market. European and Asian banks also needed capital as the crisis in the US escalated. In August 2008, there are 9 European banks with SRISK higher than Lehman.

In the second episode, the US equity markets responded to the emerging Greek sovereign debt crisis with the flash crash on May 6, 2010. SRISK in Europe increased more than 30% from the low of $1,562 billion in the summer of 2009 to the peak of $2,045 billion in Jan 2012. This rise corresponded to a collapse of sovereign bond prices of several Eurozone economies, particularly the peripheral countries, Greece, Italy, Portugal, Spain, and Ireland. These bonds were regarded as riskless by investors, rating agencies, and regulators. In the early European stress tests, there was no stress considered on sovereign debt, and it continued to have zero risk weights. As many of the banks held large positions in these bonds, their equity valuation fell rapidly. As the local economies declined, the bank asset positions fell further and SRISK rose substantially. As the European financial sector deteriorated, the US continued to strengthen in part to the regulatory reform of the Dodd-Frank Act. By January 2012, the US SRISK was down to $689 billion. On the other hand, Asian SRISK continued to rise to $1,183 billion in January 2012.

As SRISK has fallen dramatically in the US and more slowly in Europe, they have been rising in Asia. Japan and China are the two biggest contributors; they have pushed the world SRISK to a level similar to the previous two peaks. The rapid growth of Japan's SRISK from 2013 to 2017 corresponds roughly to the monetary stimulus by Prime Minister Shinzo Abe. Japanese banks hold sizable positions in Japanese Government bonds. These are discounted by financial markets leading to the large measure of under-capitalization. In China, banks are mostly state-owned, and they extend credit to state-owned enterprises and local government agencies. It is worth noting that most Japanese and Chinese banks have low betas. That is, they do not appear to be particularly risky. The substantial leverage resulting from the low market-to-book ratios is the key driver of the high SRISK.

Electronic copy available at: https://ssrn.com/abstract=3108269

In fact, when we decompose their change in SRISK into three components—the change in debt, the change in equity, and the change in risk—we can see that the increase in liabilities is more than 100% of the increase in SRISK with some equity increase offsetting the increase in risk. Although the mechanisms are different, both China and Japan have rapidly rising debt levels, and much of this debt is naturally viewed as riskless. In both cases, the majority of the debt is explicitly or implicitly guaranteed by the government. In neither country is there a crisis that looks like the Global Financial Crisis or the European sovereign debt crisis. Both countries, however, have economic stress. Furthermore, the slowdown of the Chinese economy leads to dramatic declines in natural resource prices around the world and a commensurate increase in bank stress in natural resource-rich economies.

The global model captures the important global externalities that the risk of a crisis in one country is strongly influenced by the rest of the world. Under-capitalization in one country will increase the probability of a crisis in another. The financial stability of countries are interconnected, and each has a stake in the regulation of the rest of the world.

8 Robustness

Parameter Stability

As a robustness check, we run a diagnostic analysis estimating the domestic and global models, excluding one country from the sample at a time. Figure 8 shows the point estimates and the 95% confidence intervals associated with each main regressor for the domestic model. The horizontal axis lists the country that is dropped from the sample. For instance, the first point corresponds to the estimation without Australia, and the last

Electronic copy available at: https://ssrn.com/abstract=3108269

point corresponds to the estimation without the United States. For each regressor, we also draw a horizontal line at the value of the coefficient obtained from the full sample of 23 countries (see Column 1 of Table 3) for the sake of comparison. We observe that the significance of both SRISK/(TA*k) and its lag remains strong no matter which country is taken out of the sample. The magnitude of the coefficients remains stable as well. The full-sample point estimate of SRISK/(TA*k) falls in the 95% confidence interval of this variable estimated regardless of which country is dropped from estimation, so does the full-sample estimate of lagged SRISK/(TA*k).

Figure 9 shows the analog for the global model. There are three main regressors in the global model: country-level SRISK/(TA*k), world SRISK/(TA*k), and lagged world SRISK/(TA*k). Once again, the significance of all three variables, as well as the magnitude, remains relatively unchanged no matter which country is excluded from the sample.

This approach also allows us to investigate why the SRISK/GDP variable has a negative coefficient. We estimate a diagnostic Tobit model which includes SRISK/(TA*k), SRISK/GDP, and SRISK/MV as regressors, as well as country fixed effects, using the full sample of 23 countries and excluding one country from the sample at a time. The results are shown in Figure 10. Similar as in Table 1, SRISK/GDP variable has a significant negative coefficient in the full sample. In the samples where a country other than Japan is dropped, we also obtain such a significant negative coefficient. When we exclude Japan, however, the SRISK/GDP coefficient turns positive and has a much larger magnitude.

Varying Tuning Parameters

From Equation (8), we see that SRISK is calculated under a given stress level and a given prudential capital ratio. For insurance companies, there is another assumption on the included fraction of separate accounts. In our analysis so far, we have fixed these factors

Electronic copy available at: https://ssrn.com/abstract=3108269

and estimate the coefficients of SRISK measures calculated under the fixed values. In this section, we relax the assumption on these factors and examine the impact of varying them. It turns out that the stress level, the prudential capital ratio, and the included fraction of separate accounts can also be thought of as parameters whose different values give rise to different models ("tuning parameters").

The firm-level SRISK depends on the prudential capital ratio and the included fraction of separate accounts linearly, and on the stress level nonlinearly. When we aggregate firm-level SRISK into country-level SRISK, we sum up firms with positive SRISK, which makes country-level SRISK nonlinear on all these three tuning parameters. Another source of nonlinearity comes from the fact that we also need to adjust total assets, which appear in the denominator of SRISK/(TA*k) variables, by only including a fraction of total separate accounts. In sum, the Tobit model is a highly nonlinear system of three tuning parameters. In this case, we consider a grid of values of the tuning parameters and iterate through the grid to find the best model. Specifically, we consider:

- Stress level (θ) ranging from 30% to 60% in a 5% increment (7 different values)
- Included fraction of separate accounts (s) ranging from 0% to 100% in a 20% increment (6 different values)
- Capital ratios (k) can take 5 different pairs of value. The baseline model assumes k2=5.5% for European countries and k1=8% for all other countries. The other 4 values assume the same treatment for all countries: 4%, 5.5%, 8%, and 10%.

Overall, there are $7 \cdot 6 \cdot 5 = 210$ combinations. For each combination, we re-calculate firm-level SRISK using the formula and get country-level SRISK by summing all positive-SRISK firms for each point in time, and re-estimate both the domestic model and the global model.

A natural criterion we can use to select the best model is the highest (maximized) log-

Electronic copy available at: https://ssrn.com/abstract=3108269

likelihood.[6] Based on this criterion, the best domestic model is achieved with a stress level of 50%, an included fraction of separate accounts of 0%, and a capital ratio of 5.5% for all countries; the best global model is achieved with a stress level of 60%, an included fraction of separate accounts of 0%, and a capital ratio of 4% for all countries.

In both cases, the selected best model is different from the baseline model. This "in sample" victory, however, can be driven by the realized sample of data or be truly indicative of a difference "in population". To tell these two possibilities apart, we also need to compare the relative performance of the models more rigorously by assessing whether the difference in performance is statistically significant based on the work of Diebold and Mariano (1995).

Consider now the log-likelihood differential under set A of tuning parameter values relative to set B for country i and time t

$$z_{i,t} = LL_{i,t}^{A} - LL_{i,t}^{B}$$

The null hypothesis is that the mean of log-likelihood differential z is zero. In a time-series context where the DM test is originally developed, the DM statistic can be calculated by regressing the log-likelihood differential z on an intercept, using heteroscedasticity and autocorrelation consistent (HAC) standard errors. In a panel data context such as ours, both the Newey-West standard errors which consider autocorrelation of the moving average type with lag length q and the clustered standard errors at the level of the panel identifier are HAC. Neither of them allows for cross-sectional correlation, however. To allow for this possibility, we also calculate the Driscoll and Kraay (1998) standard errors which are robust to very general forms of cross-sectional as well as temporal dependence.

[6]Since the number of parameters estimated in the Tobit models and the sample period stay the same when we vary the values of these tuning parameters, selecting the best model with AIC or BIC yields the same result.

Electronic copy available at: https://ssrn.com/abstract=3108269

Loosely speaking, the Driscoll and Kraay methodology applies a Newey-West type correction to the sequence of cross-sectional averages of the moment conditions.

Both the Newey-West and the Driscoll-Kraay standard errors allow for residual correlation of the moving average type with lag length q. This assumption is not overly restrictive as autoregressive (AR) processes can normally be well approximated by finite order MA processes. Let $m(T)$ denote the maximum lag length up to which the residuals may be autocorrelated. Newey and West (1987) show that for the case of using modified Bartlett weights, their estimator is consistent if $m(T)$ increases with T but at a rate slower than $T^{1/4}$. Therefore, it is not advisable to select an $m(T)$ which is close to the maximum lag length $T-1$. We adopt the plug-in estimators developed by Andrews (1991), Newey and West (1994), and others. These are automated procedures that use an asymptotic mean squared error criterion to deliver the optimum lag length.

In our application, we use the following simple rule of thumb based on the Newey and West (1994) plug-in procedure to determine the maximum lag length m(T):

$$m(T) = \text{floor}\left[4\,(T/100)^{2/9}\right]$$

We conduct the Diebold and Mariano test for testing the difference between an alternative model (model A) and the baseline model (model B) for all alternative models that have a higher (maximized) log-likelihood than the baseline model. These are seemingly better models. By construction, the DM test statistic will be positive. For the domestic model, 36 out of 209 sets of alternative tuning parameter values yield a seemingly better model. They are listed in Table 5 in the descending order of log-likelihood. Regardless of which HAC standard errors are used, the difference from the baseline model is insignificant for any alternative model. Therefore, we conclude that the baseline model is adequate.

For the global model, 106 out of 209 sets of alternative tuning parameter values yield a

Electronic copy available at: https://ssrn.com/abstract=3108269

seemingly better model. For brevity, Table 6 only includes the DM test for 36 highest log-likelihood models. The complete result for all 106 seemingly better models is reported in the long table in Appendix B. There is evidence that better stress tests can be found than the baseline when using the simple Newey-West standard errors but these are not significantly different when using measures which take the panel structure into account.

9 Conclusion

We have estimated a model of systemic risk which is designed to show both the probability of a crisis and the distance between current measures of systemic risk and the level which makes the probability of crisis equal to a half. The model builds on the theory that deleveraging will have a price impact and the greater the magnitude of the deleveraging the more dangerous the adjustment. In its most extreme case, the real economy has restricted access to credit as the financial sector experiences a fire sale thus endogenously generating a financial crisis.

This paper quantifies this process with a simple model that incorporates systemic externalities both within countries and between countries. These externalities reinforce the potential benefits of financial regulation and coordination on a country and international level.

Countries in this model can and do differ in their tolerance of financial under-capitalization, although we do not explore the economic or political origins of these differences. The main results are insensitive to dropping any of the countries in the sample and to alternative parameters of the stress tests. Thus we hope that this research will provide a reliable guide to how much systemic risk is really too much.

Electronic copy available at: https://ssrn.com/abstract=3108269

References

Acharya, Viral V., Christian T. Brownlees, Farhang Farazmand, and Matthew Richardson, 2011, Measuring Systemic Risk, in *Regulating Wall Street: The Dodd-Frank Act and the New Architecture of Global Finance*, chapter 4 (Wiley Publisher).

Acharya, Viral V., Robert F. Engle, and Matthew Richardson, 2012, Capital Shortfall: A New Approach to Ranking and Regulating Systemic Risks, *American Economic Review: Papers and Proceedings* 102, pp. 59–64.

Adrian, Tobias, Nina Boyarchenko, and Giannone Domenico, 2016, Vulnerable Growth, *Federal Reserve Bank of New York Staff Reports* 794.

Adrian, Tobias, and Hyun Song Shin, 2010, Liquidity and leverage, *Journal of Financial Intermediation* 19, 418–437.

Adrian, Tobias, and Hyun Song Shin, 2014, Procyclical Leverage and Value-at-Risk, *Review of Financial Studies* 27, 373–403.

Andrews, Donald W K, 1991, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation, *Econometrica* 59, 817–858.

Borio, Claudio, 2014, The Financial Cycle and Macroeconomics: What Have We Learnt?, *Journal of Banking and Finance* 45, 182–198.

Brownlees, Christian T., and Robert F. Engle, 2011, Volatility, Correlation and Tails for Systemic Risk Measurement.

Brownlees, Christian T., and Robert F. Engle, 2017, SRISK: A Conditional Capital Shortfall Index for Systemic Risk Measurement, *Review of Financial Studies* .

Cont, Rama, and Eric Schaanning, 2016, Fire sales, indirect contagion and systemic stress-testing.

Electronic copy available at: https://ssrn.com/abstract=3108269

Diebold, F X, and R S Mariano, 1995, Comparing predictive accuracy, *Journal Of Business & Economic Statistics* 13.

Drehmann, Mathias, Claudio Borio, and Kostas Tsatsaronis, 2012, Characterising the Financial Cycle: Don't Lose Sight of the Medium Term!

Drexler, Alejandro, Thanases Plestis, and Richard J. Rosen, 2017, How much risk do variable annuity guarantees pose to life insurers?, *Chicago Fed Letter* .

Driscoll, John C, and Aart C Kraay, 1998, Consistent Covariance Matrix Estimation with Spatially Dependent Panel Data, *The Review of Economics and Statistics* 80, 549–560.

Engle, Robert F., 2016, Dynamic Conditional Beta, *Journal of Financial Econometrics* 14, 643–667.

Greenwood, Robin, Augustin Landier, and David Thesmar, 2015, Vulnerable banks, *Journal of Financial Economics* 115, 471–485.

Koijen, Ralph S. J., and Motohiro Yogo, 2018, The Fragility of Market Risk Insurance.

Myers, Stewart C., 1977, Determinants of corporate borrowing, *Journal of Financial Economics* 5, 147–175.

Newey, Whitney K, and Kenneth D West, 1987, A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix, *Econometrica* 55, 703–708.

Newey, Whitney K, and Kenneth D West, 1994, Automatic Lag Selection in Covariance Matrix Estimation, *The Review of Economic Studies* 61, 631–653.

Pedersen, Lasse Heje, 2009, When Everyone Runs for the Exit, *International Journal of Central Banking* 177–196.

Electronic copy available at: https://ssrn.com/abstract=3108269

Reinhart, Carmen, and Kenneth Rogoff, 2009, *This Time is Different: Eight Centuries of Financial Folly* (Princeton University Press).

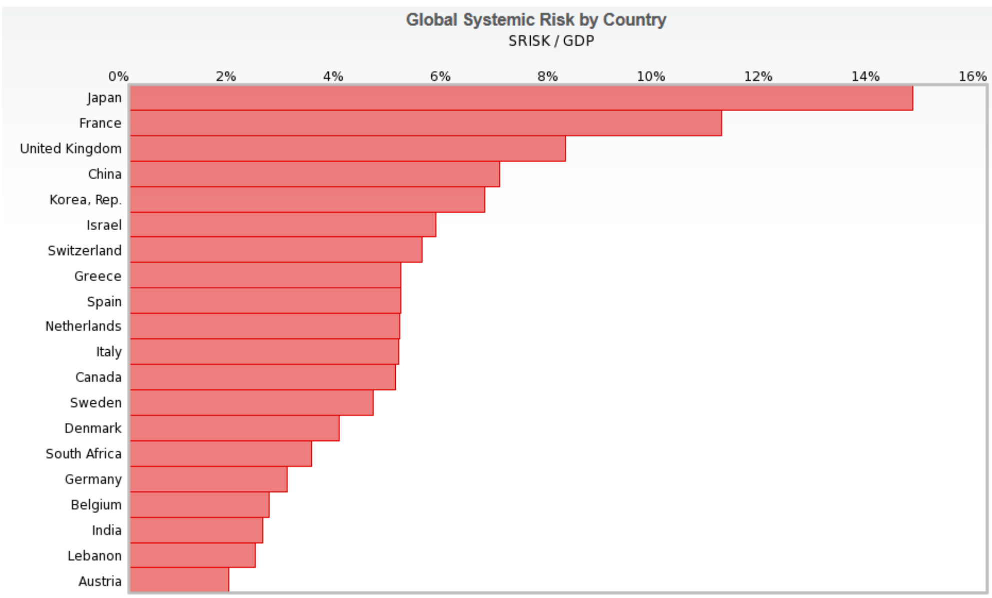
Romer, Christina D, and David H Romer, 2017, New Evidence on the Impact of Financial Crises in Advanced Countries, *American Economic Review* 107, 3072–3118.

Ruan, Tianyue, 2017, The Race to the Top: Implications for Systemic Risk.

Schularick, Moritz, and Alan M Taylor, 2012, Credit Booms Gone Bust: Monetary Policy, Leverage Cycles and Financial Crises, 1870–2008, *American Economic Review* 102, 1029–1061.

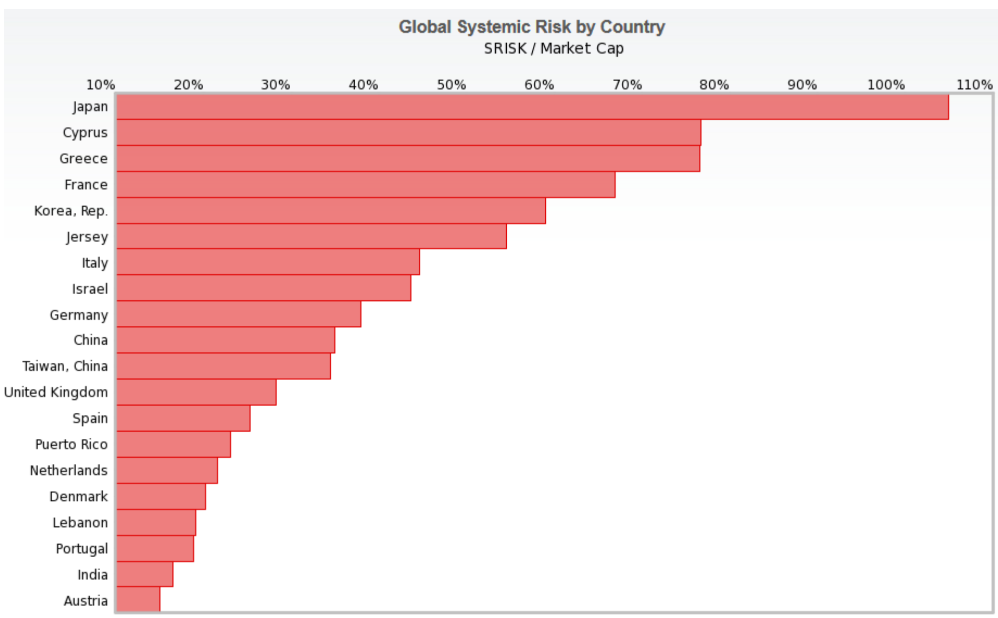
Electronic copy available at: https://ssrn.com/abstract=3108269

Figure 1: **SRISK/GDP by country in July 2017**



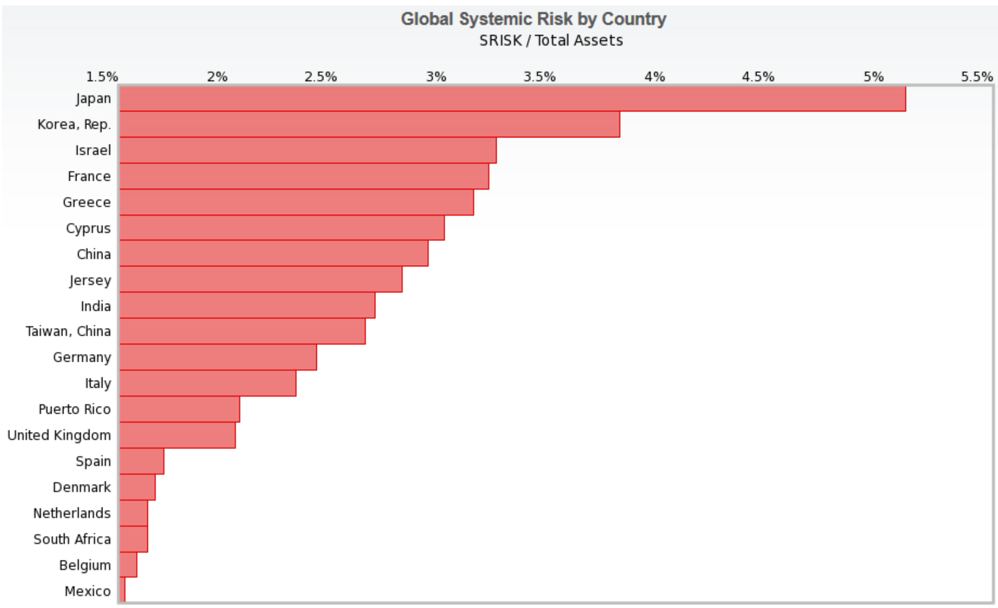
Electronic copy available at: https://ssrn.com/abstract=3108269

Figure 2: **SRISK/MV by country in July 2017**



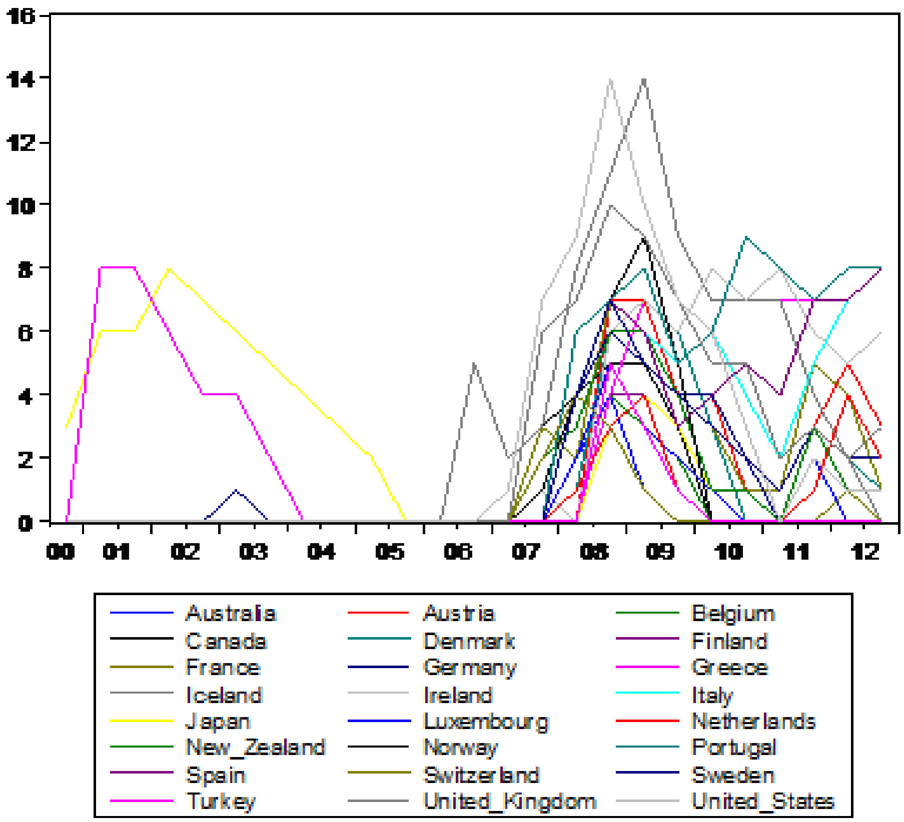
Electronic copy available at: https://ssrn.com/abstract=3108269

Figure 3: **SRISK/TA by country in July 2017**



Electronic copy available at: https://ssrn.com/abstract=3108269

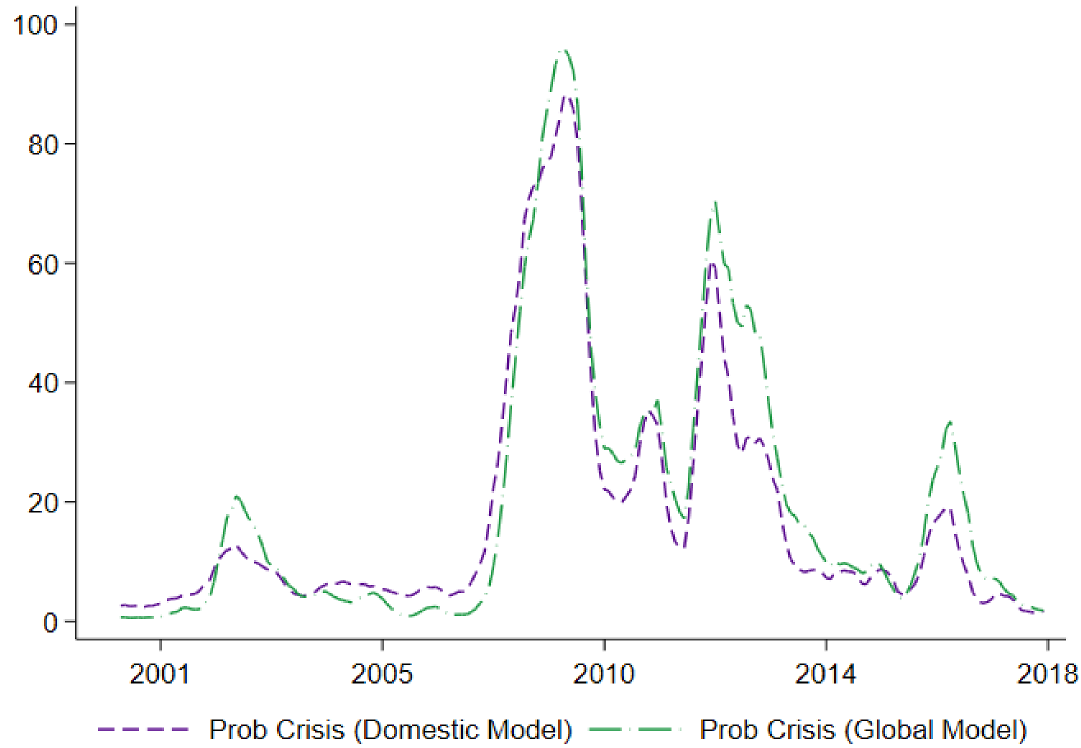
Figure 4: **Romer and Romer (2017) crisis severity over time by country**



Electronic copy available at: https://ssrn.com/abstract=3108269

Figure 5: **Probability of Crisis (%): United States**

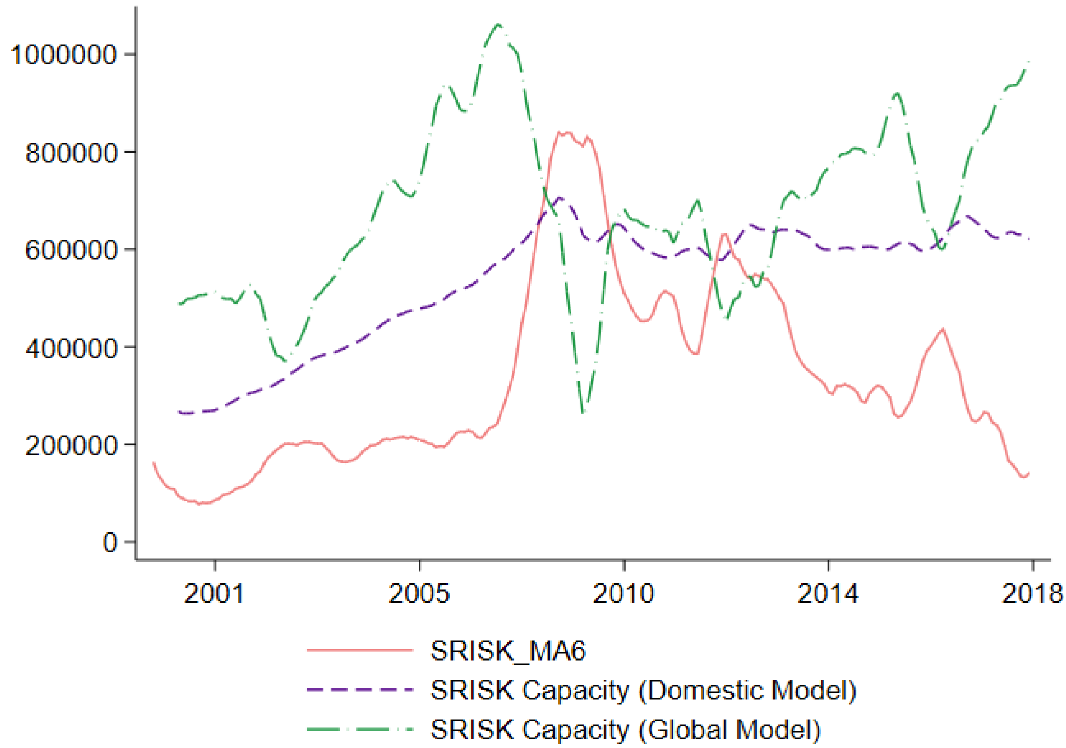
This figure plots the *Probability of Crisis* measures of the United States obtained from the domestic and global models. See the text for how the *Probability of Crisis* measure is constructed.



Electronic copy available at: https://ssrn.com/abstract=3108269

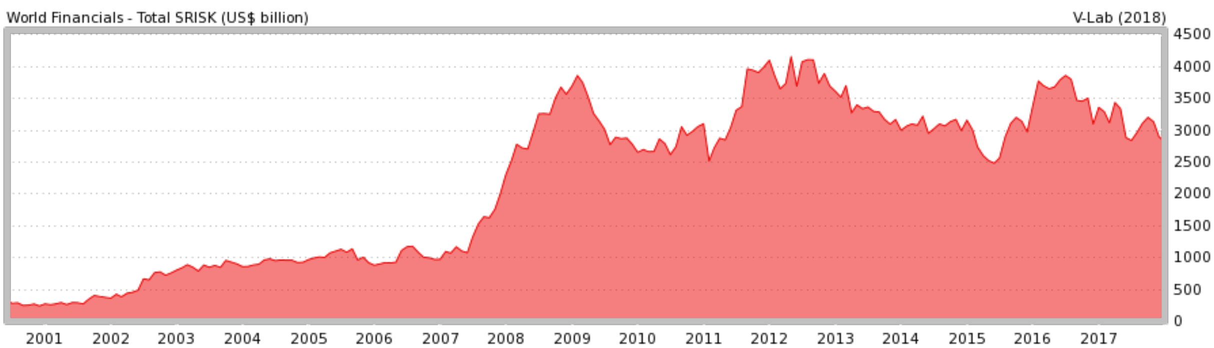
Figure 6: **SRISK Capacity (USD Million): United States**

This figure plots the *SRISK Capacity* measures of the United States obtained from the domestic and global models relative to the six-month moving-average of SRISK. See the text for how the *SRISK Capacity* measure is constructed.



Electronic copy available at: https://ssrn.com/abstract=3108269

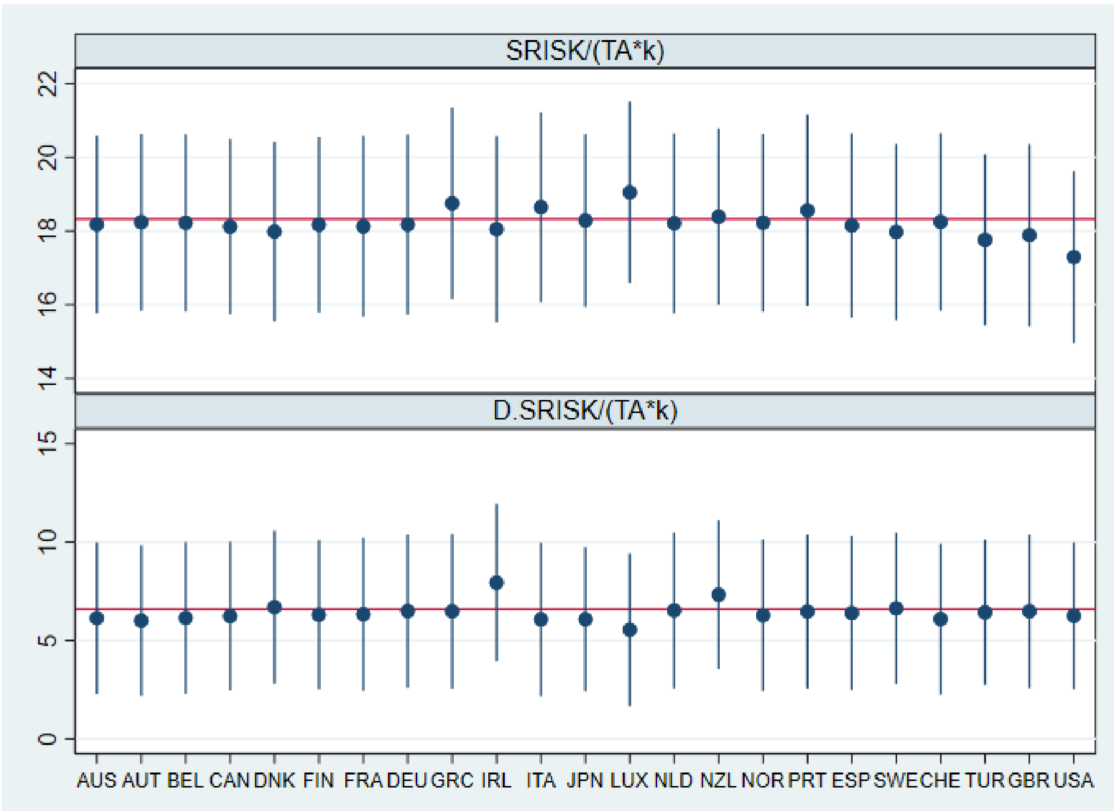
Figure 7: **World SRISK over time**



Electronic copy available at: https://ssrn.com/abstract=3108269

Figure 8: **Estimates for domestic model: Dropping one country at a time**

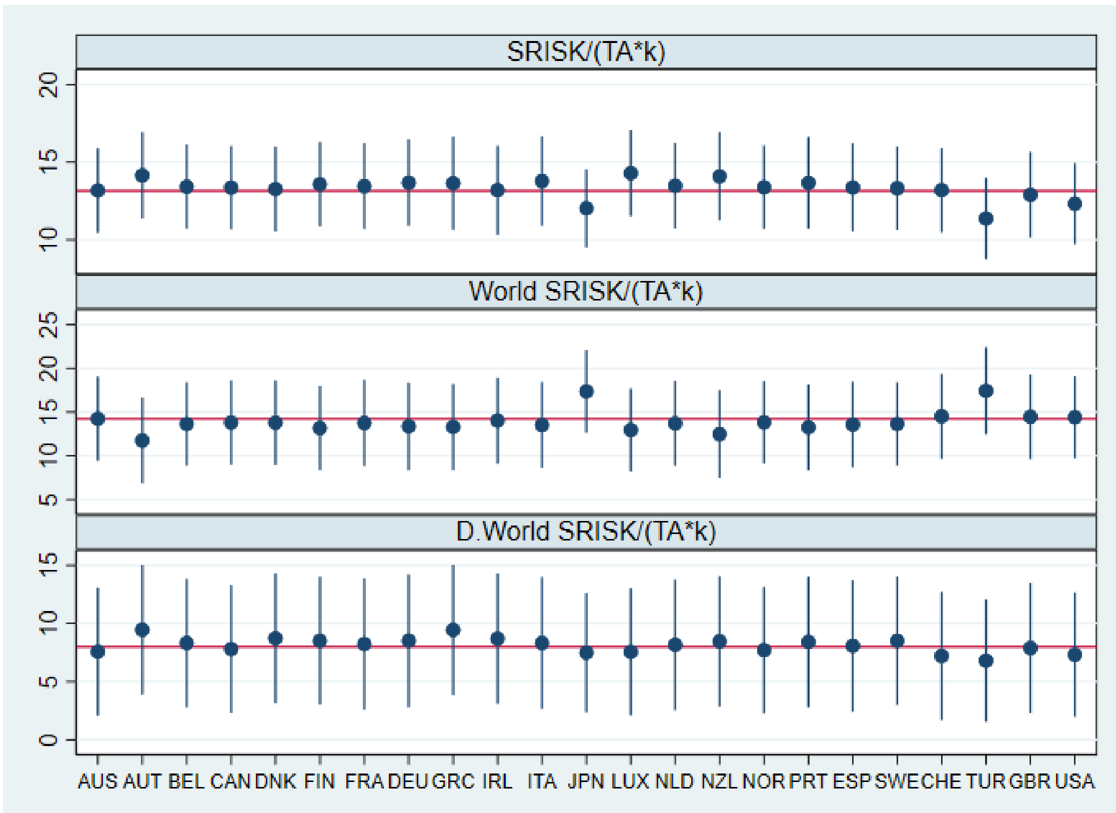
This graph shows the stability of coefficient estimates in the domestic model by plotting the coefficients and the 95% confidence intervals estimated with one country dropped from the sample at a time. The horizontal axis lists the country that is dropped. The horizontal line in each subgraph indicates the level of the coefficient obtained from the full sample of 23 countries.



Electronic copy available at: https://ssrn.com/abstract=3108269

Figure 9: **Estimates for global model: Dropping one country at a time**

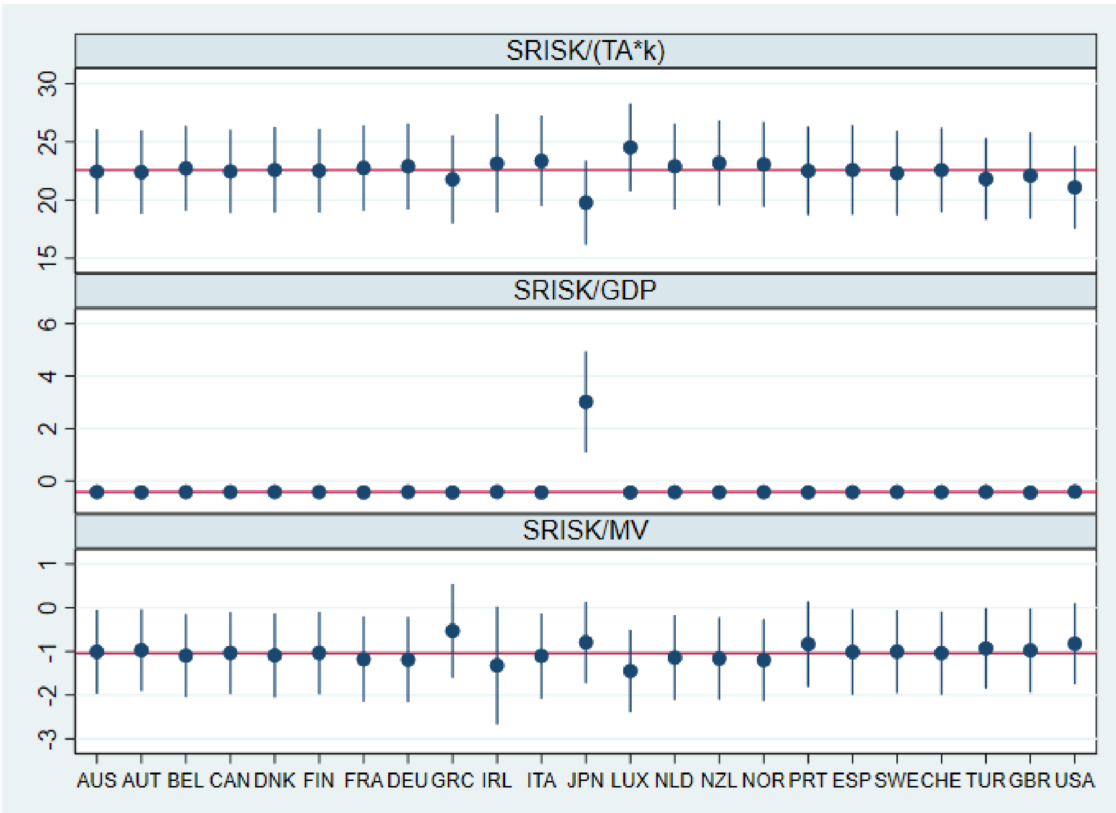
This graph shows the stability of coefficient estimates in the global model by plotting the coefficients and the 95% confidence intervals estimated with one country dropped from the sample at a time. The horizontal axis lists the country that is dropped. The horizontal line in each subgraph indicates the level of the coefficient obtained from the full sample of 23 countries.



Electronic copy available at: https://ssrn.com/abstract=3108269

Figure 10: **Estimates for diagnostic Tobit model: Dropping one country at a time**

This graph shows the stability of coefficient estimates in the diagnostic Tobit model by plotting the coefficients and the 95% confidence intervals estimated with one country dropped from the sample at a time. This diagnostic model includes SRISK/(TA*k), SRISK/GDP, and SRISK/MV as regressors, as well as country fixed effects. The horizontal axis lists the country that is dropped. The horizontal line in each subgraph indicates the level of the coefficient obtained from the full sample of 23 countries.



Electronic copy available at: https://ssrn.com/abstract=3108269

Table 1: **Crisis severity and systemic sisk measures (OLS)**

This table reports the OLS estimates of how systemic risk measures are contemporaneously associated with crisis severity. The sample includes all countries studied by Romer and Romer (2017) with the exception of Iceland from the second half of 2000 to the second half of 2012. The unit of observation is country-halfyear. The country fixed effects are included in all specifications. Standard errors are reported in parentheses. ***, **, and * represent 1%, 5%, and 10% significance, respectively.

	Dep Var: CRISIS				
	(1)	(2)	(3)	(4)	(5)
SRISK/(TA*k)	5.494***	4.253***	3.808***	3.797***	5.309***
	(0.693)	(0.750)	(0.745)	(0.705)	(0.750)
SRISK/GDP	-0.228***	-0.278***	-0.297***	-0.291***	-0.279***
	(0.053)	(0.058)	(0.059)	(0.058)	(0.061)
SRISK/MV	0.457**	0.480**	0.707***	0.594***	0.724***
	(0.207)	(0.230)	(0.227)	(0.224)	(0.234)
World SRISK/(TA*k)		-8.803*	6.058***		
		(4.557)	(0.993)		
World SRISK/GDP		13.370			
		(20.709)			
World SRISK/MV		10.866***		5.518***	
		(2.617)		(0.776)	
World log SRISK					0.373***
					(0.133)
Time FE	Yes	No	No	No	No
Country FE	Yes	Yes	Yes	Yes	Yes
Within Country R^2	0.611	0.493	0.477	0.489	0.449
Overall R^2	0.652	0.547	0.532	0.543	0.481
Observations	564	564	564	564	564

Electronic copy available at: https://ssrn.com/abstract=3108269

Table 2: **Predictive power of systemic risk measures**

This table reports the OLS estimates of how systemic risk measures predict crisis severity. The sample includes all countries studied by Romer and Romer (2017) with the exception of Iceland from the second half of 2000 to the second half of 2012. The sample size decreases relative to Table 1 due to the use of lagged variables. The unit of observation is country-halfyear. The country fixed effects are included in all specifications. Standard errors are reported in parentheses. ***, **, and * represent 1%, 5%, and 10% significance, respectively.

	Dep Var: CRISIS			
	(1)	(2)	(3)	(4)
L.CRISIS	0.789***	0.844***	0.871***	0.825***
	(0.031)	(0.035)	(0.035)	(0.033)
L.SRISK/(TA*k)	1.369***	0.262	0.856*	-0.794
	(0.425)	(0.500)	(0.479)	(0.482)
L.World SRISK/(TA*k)		-0.805		
		(0.837)		
L.World SRISK/MV			-2.188***	
			(0.664)	
L.World log SRISK				0.296***
				(0.104)
Time FE	Yes	No	No	No
Country FE	Yes	Yes	Yes	Yes
Within Country R^2	0.815	0.674	0.681	0.679
Overall R^2	0.836	0.710	0.716	0.714
Observations	541	541	541	541

Electronic copy available at: https://ssrn.com/abstract=3108269

Table 3: **Crisis severity and systemic risk measures (Tobit)**

This table reports the Tobit estimates of how systemic risk measures are contemporaneously associated with crisis severity. The sample includes all countries studied by Romer and Romer (2017) with the exception of Iceland from the second half of 2000 to the second half of 2012. The unit of observation is country-halfyear. The world SRISK variables are calculated using leave-one-out sums of respective SRISK variables. The country fixed effects are included in all specifications. Standard errors are reported in parentheses. ***, **, and * represent 1%, 5%, and 10% significance, respectively.

	Dep Var: CRISIS			
	(1)	(2)	(3)	(4)
SRISK/(TA*k)	18.325***	13.165***	12.872***	15.467***
	(1.213)	(1.366)	(1.311)	(1.385)
D.SRISK/(TA*k)	6.592***		3.958**	
	(1.931)		(1.874)	
World SRISK/(TA*k)		14.249***		
		(2.387)		
D.World SRISK/(TA*k)		7.987***		
		(2.759)		
World SRISK/MV			10.128***	
			(1.576)	
World log SRISK				1.855***
				(0.360)
D.World log SRISK				3.988***
				(0.977)
var(e.CRISIS)	11.102***	9.852***	9.829***	10.573***
	(1.263)	(1.110)	(1.109)	(1.197)
Country FE	Yes	Yes	Yes	Yes
Pseudo R^2	0.261	0.291	0.287	0.283
Observations	561	561	561	561

Electronic copy available at: https://ssrn.com/abstract=3108269

Table 4: **Country fixed effects in Tobit Model**

This table reports the estimates of country fixed effects corresponding to models reported in Table 3. Standard errors are reported in parentheses. ***, **, and * represent 1%, 5%, and 10% significance, respectively.

	Dep Var: CRISIS			
	(1)	(2)	(3)	(4)
Australia	-7.950***	-12.812***	-10.118***	-34.752***
	(1.406)	(1.655)	(1.454)	(5.405)
Austria	-7.261***	-10.436***	-7.943***	-32.844***
	(0.991)	(1.152)	(0.978)	(5.066)
Belgium	-15.849***	-17.913***	-15.321***	-41.048***
	(1.555)	(1.590)	(1.503)	(5.154)
Canada	-5.634***	-10.000***	-7.367***	-32.027***
	(0.996)	(1.287)	(1.047)	(5.229)
Denmark	-10.622***	-13.378***	-10.695***	-36.355***
	(1.236)	(1.336)	(1.191)	(5.127)
Finland	-6.365***	-10.960***	-8.588***	-32.832***
	(1.175)	(1.499)	(1.284)	(5.350)
France	-11.727***	-13.364***	-10.736***	-36.680***
	(1.170)	(1.207)	(1.112)	(4.957)
Germany	-11.640***	-13.317***	-10.677***	-36.580***
	(1.131)	(1.171)	(1.080)	(4.928)
Greece	-5.175***	-8.367***	-5.990***	-30.906***
	(1.111)	(1.231)	(1.051)	(5.161)
Ireland	-7.774***	-10.441***	-7.816***	-33.462***
	(1.119)	(1.205)	(1.058)	(5.077)
Italy	-5.670***	-8.902***	-6.372***	-31.476***
	(0.997)	(1.157)	(0.963)	(5.110)
Japan	-8.699***	-10.125***	-7.749***	-32.984***
	(1.012)	(1.013)	(0.949)	(4.771)
Luxembourg	-9.918***	-13.308***	-10.740***	-35.972***
	(1.299)	(1.415)	(1.254)	(5.231)
Netherlands	-14.455***	-16.393***	-13.738***	-39.603***
	(1.371)	(1.419)	(1.314)	(5.076)
New Zealand	0.366	-5.285***	-2.647**	-27.204***
	(0.940)	(1.306)	(1.029)	(5.369)
Norway	-7.903***	-11.441***	-8.872***	-34.021***
	(1.157)	(1.320)	(1.133)	(5.214)
Portugal	-3.938***	-7.358***	-4.922***	-29.841***
	(0.992)	(1.167)	(0.960)	(5.148)
Spain	-3.827***	-7.622***	-5.159***	-29.906***
	(0.938)	(1.169)	(0.942)	(5.168)
Sweden	-9.151***	-12.416***	-9.772***	-35.051***
	(1.129)	(1.296)	(1.120)	(5.152)
Switzerland	-10.815***	-13.978***	-11.282***	-36.714***
	(1.200)	(1.351)	(1.185)	(5.162)
Turkey	-0.846	-5.425***	-3.039***	-27.164***
	(0.791)	(1.102)	(0.847)	(5.137)
United Kingdom	-6.423***	-9.348***	-6.726***	-32.054***
	(1.028)	(1.145)	(0.974)	(5.056)
United States	-4.345***	-8.153***	-5.595***	-29.941***
	(0.883)	(1.102)	(0.888)	(5.011)
Pseudo R^2	0.261	0.291	0.287	0.283
Observations	561	561	561	561

Electronic copy available at: https://ssrn.com/abstract=3108269

Table 5: **Tuning parameters for the domestic model**

This table reports the t-statistics for the Diebold and Mariano (DM) test comparing the domestic model obtained from an alternative set of tuning parameter values against the baseline model. We construct the DM test statistic with three types of heteroscedasticity and autocorrelation consistent (HAC) standard errors—the Newey-West, clustered by country, and the Driscoll-Kraay standard errors. In specifying the maximum lag for the residual autocorrelation for both the Newey-West and the Driscoll-Kraay standard errors, we use a heuristic based on the Newey and West (1994) plug-in procedure. See the main text for details. We include all alternative models that have a higher (maximized) log-likelihood, which result in a positive DM test statistic by construction, here.

	Tuning parameters				DM statistic using HAC standard errors		
	Stress (%)	SA incl. (%)	k1 (%)	k2 (%)	Newey-West	Clustered by country	Driscoll-Kraay
Model							
1	50	0	5.5	5.5	0.799	0.576	0.832
2	50	20	5.5	5.5	0.799	0.574	0.822
3	50	40	5.5	5.5	0.749	0.535	0.775
4	45	0	5.5	5.5	0.644	0.515	0.700
5	45	20	5.5	5.5	0.643	0.513	0.699
6	45	40	5.5	5.5	0.617	0.490	0.678
7	50	60	5.5	5.5	0.686	0.489	0.715
8	45	60	5.5	5.5	0.579	0.457	0.645
9	50	80	5.5	5.5	0.620	0.442	0.651
10	45	80	5.5	5.5	0.529	0.415	0.595
11	50	100	5.5	5.5	0.559	0.399	0.590
12	45	100	5.5	5.5	0.481	0.376	0.544
13	60	0	4.0	4.0	0.269	0.190	0.305
14	60	20	4.0	4.0	0.266	0.188	0.294
15	60	40	4.0	4.0	0.236	0.167	0.259
16	45	0	8.0	5.5	0.680	0.472	0.645
17	60	60	4.0	4.0	0.194	0.137	0.211
18	45	20	8.0	5.5	0.630	0.439	0.568
19	60	80	4.0	4.0	0.152	0.108	0.164
20	55	0	5.5	5.5	0.208	0.140	0.199
21	45	40	8.0	5.5	0.465	0.329	0.403
22	55	20	5.5	5.5	0.190	0.128	0.177
23	60	100	4.0	4.0	0.107	0.076	0.115
24	55	40	5.5	5.5	0.129	0.088	0.120
25	45	60	8.0	5.5	0.244	0.179	0.209
26	40	0	5.5	5.5	0.072	0.059	0.078
27	40	20	5.5	5.5	0.070	0.058	0.076
28	40	40	5.5	5.5	0.060	0.050	0.066
29	55	60	5.5	5.5	0.064	0.044	0.060
30	40	0	8.0	5.5	0.463	0.537	0.416
31	40	20	8.0	5.5	0.484	0.612	0.438
32	40	60	5.5	5.5	0.028	0.023	0.031
33	50	0	8.0	5.5	0.030	0.021	0.027
34	45	80	8.0	5.5	0.030	0.023	0.026
35	55	80	5.5	5.5	0.007	0.005	0.007
36	40	80	5.5	5.5	0.002	0.002	0.003

Electronic copy available at: https://ssrn.com/abstract=3108269

Table 6: **Diebold and Mariano test for the global model**

This table reports the t-statistics for the Diebold and Mariano (DM) test comparing the global model obtained from an alternative set of tuning parameter values against the baseline model. We construct the DM test statistic with three types of heteroscedasticity and autocorrelation consistent (HAC) standard errors—the Newey-West, clustered by country, and the Driscoll-Kraay standard errors. In specifying the maximum lag for the residual autocorrelation for both the Newey-West and the Driscoll-Kraay standard errors, we use a heuristic based on the Newey and West (1994) plug-in procedure. See the main text for details. We conduct the DM test for the alternative models that have a higher (maximized) log-likelihood, which result in a positive DM test statistic by construction. For brevity, we include the result for 36 highest log-likelihood models here. The complete result for all 106 seemingly better models is reported in the long table in the appendix.

	Tuning parameters				DM statistic using HAC standard errors		
	Stress (%)	SA incl. (%)	k1 (%)	k2 (%)	Newey-West	Clustered by country	Driscoll-Kraay
Model							
1	60	0	4.0	4.0	2.205	1.538	1.752
2	60	20	4.0	4.0	2.197	1.527	1.767
3	60	40	4.0	4.0	2.122	1.467	1.746
4	50	0	5.5	5.5	2.509	1.658	1.674
5	55	0	5.5	5.5	2.053	1.342	1.439
6	50	20	5.5	5.5	2.531	1.652	1.712
7	55	20	5.5	5.5	2.023	1.314	1.448
8	60	60	4.0	4.0	2.012	1.382	1.703
9	50	40	5.5	5.5	2.454	1.581	1.718
10	55	40	5.5	5.5	1.922	1.243	1.432
11	60	80	4.0	4.0	1.895	1.293	1.648
12	50	60	5.5	5.5	2.339	1.489	1.710
13	45	0	5.5	5.5	2.384	1.704	1.635
14	60	0	5.5	5.5	1.362	0.907	1.025
15	55	60	5.5	5.5	1.815	1.168	1.417
16	45	20	5.5	5.5	2.419	1.707	1.689
17	60	100	4.0	4.0	1.778	1.205	1.584
18	60	20	5.5	5.5	1.303	0.866	1.007
19	50	80	5.5	5.5	2.223	1.400	1.697
20	55	80	5.5	5.5	1.715	1.096	1.401
21	55	0	4.0	4.0	1.776	1.257	1.729
22	45	40	5.5	5.5	2.381	1.655	1.709
23	55	20	4.0	4.0	1.754	1.232	1.734
24	55	100	5.5	5.5	1.621	1.029	1.383
25	50	100	5.5	5.5	2.113	1.314	1.680
26	60	40	5.5	5.5	1.211	0.803	0.975
27	45	60	5.5	5.5	2.289	1.560	1.701
28	55	40	4.0	4.0	1.692	1.182	1.696
29	60	60	5.5	5.5	1.125	0.743	0.944
30	55	60	4.0	4.0	1.581	1.094	1.612
31	45	80	5.5	5.5	2.157	1.444	1.660
32	60	80	5.5	5.5	1.049	0.690	0.916
33	45	100	5.5	5.5	2.028	1.333	1.599
34	55	80	4.0	4.0	1.456	0.997	1.497
35	60	100	5.5	5.5	0.973	0.637	0.882
36	55	100	4.0	4.0	1.318	0.895	1.351

Electronic copy available at: https://ssrn.com/abstract=3108269

Appendices

A Measures for Other Countries

This appendix contains the *Probability of Crisis* and the *SRISK Capacity* measures for all 23 countries that are studied by Romer and Romer (2017) and have SRISK data.

List of Figures

Electronic copy available at: https://ssrn.com/abstract=3108269

B Additional Results for the Diebold and Mariano Test

Table B.1: Diebold and Mariano test for the global model

This table reports the t-statistics for the Diebold and Mariano (DM) test comparing the global model obtained from an alternative set of tuning parameter values against the baseline model. We construct the DM test statistic with three types of heteroscedasticity and autocorrelation consistent (HAC) standard errors—the Newey-West, clustered by country, and the Driscoll-Kraay standard errors. In specifying the maximum lag for the residual autocorrelation for both the Newey-West and the Driscoll-Kraay standard errors, we use a heuristic based on the Newey and West (1994) plug-in procedure. See the main text for details. We include all alternative models that have a higher (maximized) log-likelihood, which result in a positive DM test statistic by construction, here.

	Tuning parameters				DM statistic using HAC standard errors		
	Stress (%)	SA incl. (%)	k1 (%)	k2 (%)	Newey-West	Clustered by country	Driscoll-Kraay
Model							
1	60	0	4.0	4.0	2.205	1.538	1.752
2	60	20	4.0	4.0	2.197	1.527	1.767
3	60	40	4.0	4.0	2.122	1.467	1.746
4	50	0	5.5	5.5	2.509	1.658	1.674
5	55	0	5.5	5.5	2.053	1.342	1.439
6	50	20	5.5	5.5	2.531	1.652	1.712
7	55	20	5.5	5.5	2.023	1.314	1.448
8	60	60	4.0	4.0	2.012	1.382	1.703
9	50	40	5.5	5.5	2.454	1.581	1.718
10	55	40	5.5	5.5	1.922	1.243	1.432
11	60	80	4.0	4.0	1.895	1.293	1.648
12	50	60	5.5	5.5	2.339	1.489	1.710
13	45	0	5.5	5.5	2.384	1.704	1.635
14	60	0	5.5	5.5	1.362	0.907	1.025
15	55	60	5.5	5.5	1.815	1.168	1.417
16	45	20	5.5	5.5	2.419	1.707	1.689
17	60	100	4.0	4.0	1.778	1.205	1.584
18	60	20	5.5	5.5	1.303	0.866	1.007
19	50	80	5.5	5.5	2.223	1.400	1.697
20	55	80	5.5	5.5	1.715	1.096	1.401
21	55	0	4.0	4.0	1.776	1.257	1.729
22	45	40	5.5	5.5	2.381	1.655	1.709
23	55	20	4.0	4.0	1.754	1.232	1.734
24	55	100	5.5	5.5	1.621	1.029	1.383
25	50	100	5.5	5.5	2.113	1.314	1.680
26	60	40	5.5	5.5	1.211	0.803	0.975

Electronic copy available at: https://ssrn.com/abstract=3108269

Table B.1: Diebold and Mariano test for the global model (continued)

	Tuning parameters				DM statistic using HAC standard errors		
	Stress (%)	SA incl. (%)	k1 (%)	k2 (%)	Newey-West	Clustered by country	Driscoll-Kraay
Model							
27	45	60	5.5	5.5	2.289	1.560	1.701
28	55	40	4.0	4.0	1.692	1.182	1.696
29	60	60	5.5	5.5	1.125	0.743	0.944
30	55	60	4.0	4.0	1.581	1.094	1.612
31	45	80	5.5	5.5	2.157	1.444	1.660
32	60	80	5.5	5.5	1.049	0.690	0.916
33	45	100	5.5	5.5	2.028	1.333	1.599
34	55	80	4.0	4.0	1.456	0.997	1.497
35	60	100	5.5	5.5	0.973	0.637	0.882
36	55	100	4.0	4.0	1.318	0.895	1.351
37	40	0	5.5	5.5	1.551	1.271	1.208
38	55	0	8.0	5.5	1.363	0.984	1.173
39	40	20	5.5	5.5	1.550	1.261	1.242
40	40	40	5.5	5.5	1.511	1.221	1.244
41	50	0	8.0	5.5	1.988	1.433	1.608
42	55	20	8.0	5.5	1.236	0.898	1.101
43	50	20	8.0	5.5	1.839	1.344	1.565
44	40	60	5.5	5.5	1.395	1.116	1.182
45	60	0	8.0	5.5	0.770	0.557	0.679
46	55	40	8.0	5.5	1.055	0.769	0.982
47	40	80	5.5	5.5	1.264	0.993	1.091
48	50	40	8.0	5.5	1.537	1.133	1.394
49	60	20	8.0	5.5	0.667	0.483	0.604
50	55	60	8.0	5.5	0.883	0.645	0.860
51	45	0	8.0	5.5	2.691	1.916	1.926
52	50	0	4.0	4.0	0.751	0.527	0.942
53	40	100	5.5	5.5	1.108	0.855	0.957
54	50	60	8.0	5.5	1.233	0.914	1.187
55	50	20	4.0	4.0	0.706	0.493	0.886
56	45	20	8.0	5.5	2.516	1.810	1.984
57	60	40	8.0	5.5	0.546	0.396	0.512
58	55	80	8.0	5.5	0.732	0.534	0.740
59	35	0	8.0	8.0	0.569	0.395	0.464
60	50	40	4.0	4.0	0.660	0.459	0.816
61	40	0	8.0	8.0	0.507	0.344	0.448
62	35	20	8.0	8.0	0.536	0.367	0.452
63	30	0	8.0	8.0	0.551	0.398	0.421
64	40	20	8.0	8.0	0.464	0.311	0.425

Electronic copy available at: https://ssrn.com/abstract=3108269

Table B.1: Diebold and Mariano test for the global model (continued)

	Tuning parameters				DM statistic using HAC standard errors		
	Stress (%)	SA incl. (%)	k1 (%)	k2 (%)	Newey-West	Clustered by country	Driscoll-Kraay
Model							
65	55	100	8.0	5.5	0.612	0.446	0.635
66	45	40	8.0	5.5	2.027	1.485	1.794
67	50	80	8.0	5.5	0.943	0.701	0.946
68	30	20	8.0	8.0	0.526	0.375	0.413
69	60	60	8.0	5.5	0.443	0.321	0.429
70	50	60	4.0	4.0	0.562	0.390	0.678
71	45	0	8.0	8.0	0.382	0.253	0.363
72	35	40	8.0	8.0	0.432	0.293	0.381
73	45	20	8.0	8.0	0.324	0.213	0.320
74	60	80	8.0	5.5	0.361	0.262	0.360
75	40	40	8.0	8.0	0.365	0.244	0.351
76	30	40	8.0	8.0	0.421	0.297	0.344
77	50	100	8.0	5.5	0.696	0.517	0.714
78	50	80	4.0	4.0	0.450	0.310	0.521
79	45	60	8.0	5.5	1.385	1.031	1.328
80	60	100	8.0	5.5	0.307	0.222	0.314
81	35	60	8.0	8.0	0.330	0.222	0.303
82	35	0	5.5	5.5	0.411	0.356	0.372
83	40	60	8.0	8.0	0.280	0.186	0.281
84	45	40	8.0	8.0	0.237	0.156	0.245
85	50	0	8.0	8.0	0.199	0.130	0.196
86	35	20	5.5	5.5	0.377	0.326	0.352
87	50	100	4.0	4.0	0.331	0.226	0.366
88	30	60	8.0	8.0	0.300	0.210	0.256
89	35	40	5.5	5.5	0.331	0.286	0.319
90	35	80	8.0	8.0	0.250	0.167	0.239
91	45	80	8.0	5.5	0.801	0.604	0.769
92	40	80	8.0	8.0	0.212	0.140	0.221
93	45	60	8.0	8.0	0.167	0.109	0.180
94	50	20	8.0	8.0	0.132	0.087	0.135
95	30	80	8.0	8.0	0.202	0.140	0.179
96	35	100	8.0	8.0	0.176	0.117	0.174
97	40	0	8.0	5.5	1.318	1.204	0.798
98	40	100	8.0	8.0	0.143	0.094	0.155
99	35	60	5.5	5.5	0.205	0.177	0.201
100	45	80	8.0	8.0	0.101	0.066	0.113
101	45	100	8.0	5.5	0.343	0.260	0.318
102	30	100	8.0	8.0	0.118	0.081	0.108

Electronic copy available at: https://ssrn.com/abstract=3108269

Table B.1: Diebold and Mariano test for the global model (continued)

	Tuning parameters				DM statistic using HAC standard errors		
	Stress (%)	SA incl. (%)	k1 (%)	k2 (%)	Newey-West	Clustered by country	Driscoll-Kraay
Model							
103	40	20	8.0	5.5	1.349	1.233	0.836
104	50	40	8.0	8.0	0.055	0.036	0.059
105	45	100	8.0	8.0	0.034	0.022	0.040
106	35	80	5.5	5.5	0.059	0.050	0.057

Electronic copy available at: https://ssrn.com/abstract=3108269