Dynamic Systemic Risk: *Networks in Data Science*

SANJIV R. DAS, SEOYOUNG KIM, AND DANIEL N. OSTROV

SANJIV R. DAS
is the William and Janice Terry professor of finance and data science at Santa Clara University in Santa Clara, CA.
srdas@scu.edu

SEOYOUNG KIM
is an associate professor of finance at Santa Clara University in Santa Clara, CA.
srkim@scu.edu

DANIEL N. OSTROV
is a professor of mathematics at Santa Clara University in Santa Clara, CA.
dostrov@scu.edu

Systemic risk arises from the confluence of two effects. First, individual financial institutions (FIs) experience increases in the likelihood of default. Second, these degradations in credit quality are transmitted through the connectedness of these institutions. The framework in this article explicitly models the contributions of both of these drivers of systemic risk. By embedding these constructs in a data science model drawn from the field of social networks, we are able to construct a novel measure of systemic risk.

The Dodd–Frank Act (2010) defined a *systemically important FI* (SIFI) as any FI that is (1) large, (2) complex, (3) connected to other FIs, and (4) critical, in that it provides hard-to-substitute services to the financial system.[1] The Act did not recommend a systemic risk-scoring approach. This article provides objective models to determine SIFIs and to calculate a composite systemic risk score.

The Merton (1974) model provides an elegant way to use option pricing theory to determine the credit quality of a single firm (i.e., its term structure of credit spreads and the term structure of the probability of default [PD] for different horizons). We demonstrate how the model may be extended to a network of connected FIs, including a metric for the systemic risk of these firms that evolves over time. Therefore, this article provides an example of the power of combining mathematical finance with network science.

Our systemic risk measure has two primary attributes: (1) aggregation—that is, our metric combines risk across all firms and all connections between firms in the system to produce a summary systemic risk number that may be measured and tracked over time; and (2) attribution—how systemic risk can be mathematically analyzed to measure the sources that contribute to overall system risk. The primary way we want to understand attribution is through an *institution risk measure*, which determines the risk contributions from each firm so that the extent to which a single firm contributes to systemic risk at any point in time is quantifiable. A secondary way to look at attribution is to compute a *connectedness risk measure*, which determines the risk contributions from each pairwise link between two firms at any point in time.

CONTRAST WITH EXTANT APPROACHES

Current approaches to measuring systemic risk include the systemic expected shortfall (SES) measure of Acharya et al. (2017)[2]; the conditional value at risk (CoVaR)

[1] See also the literature analysis of Silva, Kimura, and Sobreiro (2017) for a conceptual overview and definition of systemic financial risk.

[2] See the extensive research in this class of models at Rob Engle's V-Lab at NYU: https://vlab.stern.nyu.edu/.

measure of Adrian and Brunnermeier (2016); the construction of FI networks using bivariate Granger causality regressions from Billio et al. (2012) (and a more general framework from Merton et al. 2013); the distressed insurance premium measure of Huang, Zhou, and Zhu (2012) and Black et al. (2016); the absorption ratio of Kritzman et al. (2011); the system value at risk of Bluhm and Krahnen (2014); the credit default swap (CDS)-based metric of interconnectedness used by Abbass et al. (2016); and the calculation of capital charges required to insure against unexpected losses as from Avramidis and Pasiouras (2015).

These approaches predominantly employ the correlation matrix of equity returns to develop their measures. A recent comprehensive article by Giglio, Kelly, and Pruitt (2016) examines 19 systemic risk metrics for the US economy and finds that these measures collectively are predictive of heightened left-tail economic outcomes. Furthermore, a dimension reduction approach creates a composite systemic risk measure that performs well in forecasts. Unlike the measure in this article, these 19 metrics do not exploit network analysis. All measures cited are mostly return based, and these have been criticized by Löffler and Raupach (2018) as being subject to gaming in that a bank may cause the systemic risk measure to rise, while, at the same time, having its own contribution fall. These spillover issues do not appear to be a problem in this article.

In contrast, Burdick et al. (2011) used semistructured archival data from the Securities and Exchange Commission and Federal Deposit Insurance Corporation to construct a co-lending network and then used network analysis to determine which banks pose the greatest risk to the system. Finally, Das (2016) combined credit and network information to construct aggregate systemic risk metrics that are decomposable and may be measured over time. The unifying theme across these models is to offer static snapshots of the network of FIs at various points in time. This article is a stochastic dynamic extension of the Das (2016) model.

STOCHASTIC DYNAMICS IN A NETWORK MODEL

We extend these static network models by including stochastic dynamics for the assets of the financial firms in the model. This is where the Merton (1974) model becomes useful. We give this model the moniker *Merton on a network*. This model uses geometric Brownian motion as the stochastic process for each FI's underlying assets. That is, for the n FIs in the system, we have

$$da_i = \mu_i a_i \, dt + v_i a_i \, dB_i, \qquad i = 1, 2, \ldots, n \tag{1}$$

$$da_i \, da_j = \rho_{ij} \, dt, \qquad i = 1, 2, \ldots, n; \quad j = 1, 2, \ldots, n \tag{2}$$

Here μ_i is the ith FI's expected growth rate, and v_i is its volatility (both annualized). The asset movement of FIs i and j are correlated through the coefficient ρ_{ij}.

Assuming that the ith FI has a face value of debt D_i with maturity T, Merton's model established that the FI's equity, E_i, is a call option on the assets:

$$E_i = a_i \Phi(d_{1,i}) - D_i e^{-r_f T} \Phi(d_{2,i}) \tag{3}$$

$$d_{1,i} = \frac{\ln(a_i / D_i) + (r_f + v_i^2 / 2)T}{v_i \sqrt{T}} \tag{4}$$

$$d_{2,i} = d_1 - v_i \sqrt{T} = \frac{\ln(a_i / D_i) + (r_f - v_i^2 / 2)T}{v_i \sqrt{T}} \tag{5}$$

where r_f is the risk-free rate of interest (annualized), and $\Phi(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx$ is the cumulative standard normal distribution function. Merton's model also shows that the volatility of equity is

$$\sigma_i = v_i \frac{\partial E_i}{\partial a_i} \frac{a_i}{E_i} \tag{6}$$

Because a_i and v_i are not directly observable in the market, but E_i and σ_i are, the pair of Equations 3 and 6 may be solved simultaneously to determine the values of a_i and v_i for each i at any time, t. These values, as we will see later, allow us to obtain the one-year probability of default (PD) for each financial firm, denoted λ_i, at any given point in time.[3]

Our measure for systemic risk captures the size and PD of all FIs (from the Merton model) and combines this with a network of FI connectedness to construct

[3] In implementing our model as Merton on a network, our approach is distinct from those that infer risk-neutral PDs from CDS spreads on the referenced banks (e.g., as by Huang, Zhou, and Zhu 2012). We are also afforded greater flexibility in inferring what the PDs may be under varying market conditions.

one composite system-wide value. We exploit the stochastic structure of the asset movements of all FIs via Equations 1 and 2 to create a variety of constructions of the connectedness (network) matrix. Because the underlying assets are stochastic and correlated, so is the network; as a consequence, the systemic risk score is dynamic. In sum, we have a systemic risk measure that captures, over time, the size, risk, and connectedness of firms in the financial system.

The contagion literature has attempted to capture stochastic systemic risk by other means. Simulation of contagion networks is one approach; see Espinosa-Vega and Sole (2010), Upper (2011), and Hüser (2015). Bivalent networks of banks and assets have been simulated on data from Venezuela in another approach by Levy-Carciente et al. (2015). In our complementary approach, network and firm risk are endogenously generated through the underlying Merton (1974) model, which also offers a direct empirical implementation. To illustrate, we will later provide an example using a 20-year data sample from large, publicly traded FIs.

PRACTICAL VALUE OF THE MODEL

The models developed here have many features of interest to risk managers and regulators. First, each model produces a single number for the systemic risk in the economy. Second, the risk contribution of each institution in the system enables a risk ranking of these institutions. This ranking and the measures that determine them can help determine whether an institution is systemically important, the extent of additional supervision the institution should require, and how much the capital charge should be for the risks the institution poses to the system. Third, the risk contribution of each pairwise connection between two FIs can be measured. This allows regulators to determine which relationships between FIs are of greatest concern to the overall health of the system. Fourth, the models display several useful mathematical properties that we develop to indicate a good measure of systemic risk, as discussed in the next section. Fifth, the model's rich comparative statics may be used to examine various policy prescriptions for mitigating systemic risk.

In the next section, we introduce our general framework for systemic risk and the institution risk measure. This section also introduces four desirable properties for a systemic risk model. The following section introduces three models within the general framework that have similar structures. We discuss the institution risk measure for the three models and then show that each model possesses all four desirable properties. In the next section, we introduce our fourth model, which takes a different, although intuitive, structure from the first three models. Here we discuss both the institution risk measure and the connectedness risk measure for the model, although in this case we show that the model possesses only three of the four desirable properties. The data section provides a discussion of the data, spanning two decades (from 1995 to 2015), to which we apply our four models. The empirical section describes applications of our four models and demonstrates the general consistency of their results. We close with a concluding discussion and extensions.

A GENERAL FRAMEWORK FOR SYSTEMIC RISK

Dependence

For our general framework, the systemic risk, $\mathcal{S}$, for a system of n FIs depends on the following three sets of variables:

1. $\boldsymbol{\lambda}$, an n-vector whose components, λ_i, represent the annual probability that the ith FI will default.
2. $\mathbf{a}$, an n-vector whose components, a_i, represent the market value of assets in the ith FI.
3. $\boldsymbol{\Sigma}$, an $n \times n$ matrix whose components, Σ_{ij}, represent the financial connection from the ith FI to the jth FI. Depending on the model for these connections, $\boldsymbol{\Sigma}$ may or may not be symmetric.

In other words, our systemic risk measures take the following functional form

$$\mathcal{S} = f(\boldsymbol{\lambda}, \mathbf{a}, \boldsymbol{\Sigma}) \tag{7}$$

where a specific systemic risk model corresponds to a specific function f and specific definition for the connection matrix $\boldsymbol{\Sigma}$.

Our approach complements the ideas laid out by De Nicolo, Favara, and Ratnovski (2012), who offered a class of externalities that lead to systemic risk. First, externalities from strategic complementarities are captured through asset ($\mathbf{a}$) correlations in our model.

Second, externalities related to fire sales are embedded in the default probabilities ($\boldsymbol{\lambda}$). Third, externalities from interconnectedness are captured through network structures ($\boldsymbol{\Sigma}$) in the model. These features connect the financial sector to systemic risk and the macroeconomy.

THE INSTITUTION RISK MEASURE, CONNECTEDNESS, AND THE CONNECTEDNESS RISK MEASURE

It is important that the impact of each institution on the overall systemic risk, $\mathcal{S}$, can be measured. For example, consider the case in which $\mathcal{S}$ is homogeneous in its default risks, $\boldsymbol{\lambda}$, which means, for any scalar $\alpha > 0$,

$$\alpha f(\boldsymbol{\lambda}, \mathbf{a}, \Sigma) = f(\alpha\boldsymbol{\lambda}, \mathbf{a}, \Sigma) \tag{8}$$

In this case one way to measure the impact of each institution on $\mathcal{S}$ is to decompose $\mathcal{S}$ into the sum of n components by differentiating Equation 8 with respect to α, yielding the result of Euler's theorem

$$\mathcal{S} = \frac{\partial \mathcal{S}}{\partial \boldsymbol{\lambda}} \boldsymbol{\lambda} = \sum_{i=1}^{n} \frac{\partial \mathcal{S}}{\partial \lambda_i} \lambda_i \tag{9}$$

This result clearly suggests using each component, $\frac{\partial \mathcal{S}}{\partial \lambda_i} \lambda_i$, of the sum to define the corresponding *institution risk measure* of institution i.

Systemic risk is also impacted by the *connectedness* of the institutions via pairwise links between the institutions. These links may be directed or undirected, depending on the model. One way to measure the connection from institution i to institution j is to use Σ_{ij}. In this case, if $\boldsymbol{\Sigma}$ is symmetric, it corresponds to undirected links; otherwise, there is at least one $\Sigma_{ij} \neq \Sigma_{ji}$, which corresponds to a directed link. Graphically, these links can be shown for a directed or undirected network by using a binary network adjacency matrix $\mathbf{B}$ whose components, B_{ij}, are derived from Σ_{ij} by selecting a threshold value K and then defining $B_{ij} = 1$ if $\Sigma_{ij} > K$ and $i \neq j$; otherwise, $B_{ij} = 0$. Links are then shown in an edge graph only when $B_{ij} = 1$, noting that the threshold value K can be altered as desired.

The strength of the connections described in the last paragraph do not necessarily correspond to measurements of the risk that the connection from institution i to institution j poses to the overall systemic risk. In the cases in which it does, we can refer to the strength of the connection as the *connectedness risk measure* from institution i to institution j. Connectedness risk measures are important to regulators who wish to determine which relationships between institutions are of primary concern to the overall health of the system.

FOUR FINANCIAL PROPERTIES

Ideally, from a practical viewpoint, the definition of $\boldsymbol{\Sigma}$ and the definition of the function f that defines systemic risk, $\mathcal{S}$, conforms to the following four financial properties:

- **Property 1: All other things being equal, $\mathbb{S}$ should be minimized by dividing risk equally among the *n* FIs and maximized by putting all the risk into one institution.** That is, the more the risk is spread out, the lower $\mathcal{S}$ should be. The definition of risk will depend on the model. This is a standard property emanating from diversification but is also applicable in the case of contagion. If all risk is concentrated in one entity, then contagion is instantaneous; therefore, if risk is spread out, a useful property is that the systemic score should be correspondingly lower.
- **Property 2: $\mathcal{S}$ should increase as the FIs become more entwined.** That is, if any of the off-diagonal elements of $\boldsymbol{\Sigma}$ increase, then $\mathcal{S}$ should increase. The more connected the institutions are, the greater the likelihood of contagion and systemic risk.
- **Property 3: If all the assets, a_i, are multiplied by a common factor, $\alpha > 0$, they should have no effect on $\mathcal{S}$.** If a country's FIs' assets all grow or all shrink in the same way, it should not affect the systemic risk of the country's financial system. That is, we want $f(\boldsymbol{\lambda}, \alpha\mathbf{a}, \boldsymbol{\Sigma}) = f(\boldsymbol{\lambda}, \mathbf{a}, \boldsymbol{\Sigma})$. This property is useful because it enables comparison of systemic risk scores across countries, and even for the same country, across time.
- **Property 4: Substanceless partitioning of a bank into two banks has no effect on $\mathcal{S}$.** If institution i's assets are artificially divided into two institutions of size γa_i and $(1-\gamma)a_i$ for some $\gamma \in [0, 1]$, where both of these new institutions are completely connected to each other and both have the same connections with the other banks that the original institution did, then this division is without substantive meaning, so it should not affect the value of $\mathcal{S}$. Splitting a large bank into two fully connected components with the same

connections as before should not change $\mathcal{S}$ because such a split is mere window dressing. To bring down the value of $\mathcal{S}$ by breaking up a bank, the metric states that it is important to either disconnect the two components or reduce the connectivity for each one. In fact, the metric $\mathcal{S}$ enables a regulator to assess the effect of different kinds of bank splits on reducing systemic risk.

SYSTEMIC RISK NETWORK MODELS THAT ARE HOMOGENOUS IN DEFAULT RISKS

We first examine three models that are homogenous in default risks, each using different empirical approaches and notions of risk. All three of these models satisfy all four of the financial properties listed earlier. The proof that they are satisfied is contained in the Appendix.

Models C, D, and G

We define $\mathbf{\Sigma} = \mathbf{M}$, an $n \times n$ matrix where $M_{ij} \in [0, 1]$ for all i and j and $M_{ii} = 1$ for all i. We consider three examples of $\mathbf{M}$ matrices with this property:

1. Model $\mathbf{C}$, a correlation-based model. In this case, $M_{ij} = \frac{1}{2}(\rho_{ij} + 1)$, where ρ_{ij} is the correlation between the daily asset returns of institutions i and j. Here, $\mathbf{M}$ defines an undirected network for connectedness.
2. Model $\mathbf{D}$, a conditional default model. In this case, M_{ij} is the annual conditional probability that institution j defaults if institution i fails. In this case, $\mathbf{M}$ defines a directed network. We note that even though the model is composed of default probabilities, we are using the Merton model only to define connectedness over the long term and thereafter assume this is independent of day-to-day changes in default risk.
3. Model $\mathbf{G}$, a Granger causality model. This model is based on the methodology in Billio et al. (2012). For each pair of FIs (i, j), a pair of lagged value regressions of daily asset returns, r, is run to determine whether i Granger causes j and whether j Granger causes i.

$$r_i(t) = \delta_1 + \delta_2 \cdot r_i(t-1) + \delta_3 \cdot r_j(t-1) + \epsilon_i$$

$$r_j(t) = \delta_4 + \delta_5 \cdot r_j(t-1) + \delta_6 \cdot r_i(t-1) + \epsilon_j$$

The connectedness matrix is defined as follows: $M_{ij} = 1 - p(\delta_6)$ and $M_{ji} = 1 - p(\delta_3)$, where $p(x)$ is the p-value for the hypothesis that the coefficient $x = \delta_6$ or δ_3 is equal to zero in the regressions. When $i = j$, we set $M_{ii} = 1$. In this case, $\mathbf{M}$ defines a directed network.

Next, define $\mathbf{c}$ to be the n-vector whose components, c_i, represent institution i's credit risk. Specifically, we define

$$\mathbf{c} = \mathbf{a} \circ \boldsymbol{\lambda}$$

where $\circ$ represents the Hadamard (or Schur) product, meaning that we have element-wise multiplication: $c_i = a_i\lambda_i$.[4]

With these definitions of $\mathbf{M}$ and $\mathbf{c}$, we can define the *systemic risk*, $\mathcal{S}$, by

$$\mathcal{S} = \frac{\sqrt{\mathbf{c}^T \mathbf{M} \mathbf{c}}}{\mathbf{1}^T \mathbf{a}} \tag{10}$$

where $\mathbf{1}$ is an n-vector of ones, and the superscript T denotes the transpose of the vector. Note that the numerator is the weighted norm of the vector $\mathbf{c}$, and the denominator $\mathbf{1}^T\mathbf{a} = \sum_{i=1}^{n} a_i$ represents the total assets in the n FIs. Also note that $\mathbf{M}$ is unitless in models $\mathbf{C}$, $\mathbf{D}$, and $\mathbf{G}$; therefore, because of the presence of assets, both the numerator and denominator in Equation 10 have monetary units that cancel each other, so $\mathcal{S}$ is a unitless measure of systemic risk.

The Institution Risk Measure and Connectedness

Our model is homogeneous in $\boldsymbol{\lambda}$, so, from Equation 9, we have that

$$\mathcal{S} = \frac{\partial \mathcal{S}}{\partial \boldsymbol{\lambda}} \boldsymbol{\lambda} = \sum_{i=1}^{n} \frac{\partial \mathcal{S}}{\partial \lambda_i} \lambda_i$$

where, from differentiating our system risk definition in Equation 10, we obtain the n-dimensional vector

$$\frac{\partial \mathcal{S}}{\partial \boldsymbol{\lambda}} = \frac{1}{2} \frac{\mathbf{a} \circ [(\mathbf{M} + \mathbf{M}^T)\mathbf{c}]}{\mathbf{1}^T \mathbf{a} \sqrt{\mathbf{c}^T \mathbf{M} \mathbf{c}}} \tag{11}$$

[4] We note that this definition of credit risk is qualitatively similar in nature to replacing $\mathbf{a}$ with the quantity of debt. That is, FIs tend to uniformly maximize along the imposed capital adequacy ratio, which results in the low cross-sectional variation in leverage across the institutions in question. Exhibit 1 presents various examples of the range in leverage across institutions at different points in time.

This decomposition of $\mathcal{S}$ gives the risk measure of each institution. The off-diagonal elements of $\mathbf{M}$ give the connectedness, although this notion of connectedness is not a connectedness risk measure.

A SYSTEMIC RISK NETWORK MODEL THAT IS NOT HOMOGENOUS IN DEFAULT RISKS

The network model in this section corresponds to a different financial view of constituting risk. As explained in the Appendix, this section's model satisfies our first three financial properties, but not the fourth.

Model R (Internal Risk Plus External Risks Model)

For this model we define $\mathbf{\Sigma} = \mathbf{M}$, where M_{ij} is the annual probability that FIs i and j both default. Next, we consider the following view of defining the risk to the system from institution i: Institution i has internal risk, which measures the chance that it will collapse and via the impact of that collapse, hurts the system directly; and it has external risk, the chance that its collapse will cause other FIs to collapse, hurting the system further. The internal risk for FI i is defined simply as the credit risk, $c_i = \lambda_i a_i$, that we had previously. Note that we can also write this as $c_i = M_{ii} a_i$ because, by definition, $M_{ii} = \lambda_i$. The external risk from FI i to FI j is defined as the probability that FI i will default multiplied by the probability that FI j will default given that FI i defaults multiplied by the assets in FI j. Because this is equal to the probability that both FI i and FI j default multiplied by the assets in FI j, we can write this as $M_{ij} a_j$.

We thus can define ρ_i, which is the internal risk from FI i plus the sum of the external risks from FI i to each of the other FIs, by

$$\rho_i = \sum_{j=1}^{n} M_{ij} a_j \tag{12}$$

Defining $\boldsymbol{\rho}$ to be the n-vector with components ρ_i, we can define the systemic risk to be

$$\mathcal{S} = \frac{\sqrt{\boldsymbol{\rho}^T \boldsymbol{\rho}}}{\mathbf{1}^T \mathbf{a}} \tag{13}$$

Note again that $\mathcal{S}$ is unitless, as was the case in the previous section when we defined $\mathcal{S}$ in Equation 10 for models $\mathbf{C}$, $\mathbf{D}$, and $\mathbf{G}$.

The Institution Risk Measure and the Connectedness Risk Measure

These measures are straightforward. Institution i's risk measure in this case is the value of ρ_i defined earlier. Note here that $\sum_{i=1}^{n} \rho_i \neq \mathcal{S}$, unlike the case in which $\mathcal{S}$ is homogeneous in $\boldsymbol{\lambda}$, for which this equality holds because of Equation 9. This model, unlike the three models from the previous section, has a connectedness risk measure from bank i to bank j, which is the external risk, $M_{ij} a_j$.

DATA SOURCES AND DESCRIPTION OF VARIABLES

All four models are easy to implement using publicly available data. We describe our data sources and present key summary statistics. The data used are extensive and publicly available. Hence, the approach is amenable to many data science methods applied to big data.

Sources

Our sample period spans January 1992 to December 2015 and consists of publicly traded FIs under major Standard Industrial Classification (SIC) groups 60 (depository institutions), 61 (nondepository credit institutions), and 62 (security and commodity brokers, dealers, exchanges, and services).[5] We obtain daily stock returns, stock prices, and shares outstanding for each of these firms, as well as the daily market returns, from the Center for Research in Securities Prices. We obtain applicable Treasury rates (i.e., the constant-maturity rates) on a monthly basis from the Federal Reserve Bank reports, and we obtain quarterly balance-sheet and income-statement data from Compustat. Our final sample consists of a panel dataset of 2,066,868 firm-days for 1,171 distinct FIs, from which we select the 20 largest institutions by total assets at various points across time. Working with more institutions does not pose computational difficulty; we choose only 20 institutions

[5] For a detailed breakdown of the SIC division structure, see https://www.osha.gov/pls/imis/sic_manual.html.

for clarity. The top 20 institutions consistently represent over 70% of the total worth of the assets in the 1,171 FIs.

Key Definitions and Data-Generating Computations

We solve for the ith FI's market value of assets, $a_i(t)$, and the annualized volatility of asset returns, $\nu_i(t)$ on day t, based on the Merton (1974) model for calculating equity value and equity return volatility. Recall Equations 3 and 6. Given market capitalization, $E_i(t)$; annualized equity return volatility,[6] $\sigma_i(t)$; total face value of debt, $D_i(t)$; and the annualized risk-free rate of return,[7] $r_f(t)$, we can use a simultaneous nonlinear equation root finder to simultaneously solve Equations 3 and 6 and determine the values of $a_i(t)$ and $\nu_i(t)$ for any i and t.[8]

Once we have our panel of daily asset values, $a_i(t)$, and volatilities, $\nu_i(t)$, we can calculate the daily asset returns, $r_i(t)$. The daily asset returns allow us to run the Granger regressions that determine M_{ij} in model **G** and to determine ρ_{ij}, the correlation of the daily asset returns of institutions i and j, which defines M_{ij} in model **C**. Furthermore, the daily asset returns allow us to compute asset betas, $\beta_i(t)$, which we do on a daily, rolling basis, based on a three-year (i.e., 750-day) lookback period for $r_i(t)$. Using this information, we can then calculate expected asset returns, $\mu_i(t)$, using the capital asset pricing model as follows

$$\mu_i(t) = \beta_i(t)\cdot(\mu_{MKT}(t) - r_f(t)) + r_f(t) \tag{14}$$

where $\mu_{MKT}(t)$ represents the annualized expected return on the market portfolio on day t. For the illustrative purposes of this article, we simply set $\mu_{MKT}(t)$ equal to a constant value of 10%.

The expected asset returns are used to determine $\lambda_i(t)$, the annualized PD, which is the probability that the market value of the FI's assets, a_i, governed by the geometric Brownian motion in Equation 1, will become smaller than the FI's current debt, D_i, in a year. That is

$$\lambda_i(t) = \Phi\left(-\hat{d}_{2,i}\right) \tag{15}$$

where $\Phi(\cdot)$ is the cumulative standard normal distribution function,

$$\hat{d}_{2,i} = \frac{\ln\left(\frac{a_i(t)}{D_i(t)}\right) + \left(\mu_i(t) - \left(\frac{(\nu_i(t))^2}{2}\right)\right)T}{\nu_i(t)\sqrt{T}}, \tag{16}$$

and $T = 1$ year. Note that $\hat{d}_{2,i}$ has the same definition as $d_{2,i}$ in Equation 5, but with $r_f(t)$ in that equation replaced by $\mu_i(t)$. That is, $\hat{d}_{2,i}$ corresponds to $d_{2,i}$ in the physical, instead of the risk-neutral, measure.

To determine the joint probability that both FIs i and j will default, which is the M_{ij} for model **R**, we have that

$$M_{ij} = \Phi_2(-\hat{d}_{2,i}, -\hat{d}_{2,j}, \rho_{ij})$$

where $T = 1$ year and $\Phi_2(\cdot,\cdot,\cdot)$ is the bivariate cumulative standard normal distribution function defined by

$$\Phi_2(z_1, z_2, \rho) = \int_{-\infty}^{z_1}\int_{-\infty}^{z_2} \frac{1}{2\pi\sqrt{\det(\mathbf{S})}} \exp\left(-\frac{1}{2}\mathbf{x}^T\mathbf{S}^{-1}\mathbf{x}\right) dx_1 dx_2$$

where $\mathbf{x}$ is a column vector with entries x_1 and x_2, and $\mathbf{S}$ is a 2 × 2 matrix with ones on the diagonal and ρ in the two off-diagonal entries.[9] Finally, to determine the conditional default probability M_{ij} for model **D**, we simply divide the M_{ij} for model **R** by λ_i.

Exhibit 1 shows the evolution of these basic summary statistics over time. We note as a reality check for our calculations that the total book value of assets tracks our calculated implied market value of assets in each exhibit. For instance, as of the end of June 1995, we see that our 20 FIs held an average of approximately \$120.1 billion in total assets, which grows considerably to \$354.3 billion by the end of June 2000 and then grows further to \$1,313 billion by the end of June 2007. However, as a result of the financial crisis of 2008, this average is only moderately greater, at \$1,546 billion, by the end of June 2015. The average leverage stays approximately constant at 0.9407, 0.9475, and 0.9521 in June of 1995, 2000, and 2007, respectively. Some deleveraging to an average ratio of 0.9265 happens by the end of June 2015. The dominance of the 20 largest FIs over the field of all FIs fluctuates over the years, from

[6] We calculate equity-return volatility based on a 130-day (i.e., six-month) lookback period, which we then multiply by $\sqrt{252}$.

[7] We use the three-month constant maturity T-bill rate.

[8] We use the multiroot function for finding roots, which is included in R's rootSolve package.

[9] We use the pmvnorm function, which is included in R's mvtnorm package, to calculate $\Phi_2(\cdot,\cdot,\cdot)$.

EXHIBIT 1
20 Largest FIs at Various Times

	Mean	P25	P50	P75
Panel A: June 30, 1995				
Book Value of Assets	120,061	66,302	107,548	125,522
Leverage	0.9407	0.9290	0.9390	0.9529
Market Capitalization, E	7,622	2,718	6,814	9,548
Equity Volatility, σ	0.2370	0.1941	0.2230	0.2711
Implied Market Value of Assets, a	114,788	65,456	108,927	124,465
Implied Volatility of Assets, v	0.0171	0.0101	0.0175	0.0224
% of all FIs' Total Assets:	77.34%			
Panel B: June 30, 2000				
Book Value of Assets	354,319	235,274	276,039	417,851
Leverage	0.9475	0.9342	0.9518	0.9655
Market Capitalization, E	40,353	11,354	26,272	58,439
Equity Volatility, σ	0.4485	0.3522	0.4662	0.4999
Implied Market Value of Assets, a	357,125	243,392	265,194	418,995
Implied Volatility of Assets, v	0.0531	0.0159	0.0394	0.0922
% of all FIs' Total Assets:	73.83%			
Panel C: June 29, 2007				
Book Value of Assets	1,313,221	796,235	1,191,233	1,541,156
Leverage	0.9521	0.9443	0.9558	0.9673
Market Capitalization, E	61,228	3,472	39,833	88,437
Equity Volatility, σ	0.1956	0.1646	0.1999	0.2240
Implied Market Value of Assets, a	1,254,163	777,503	1,132,813	1,468,445
Implied Volatility of Assets, v	0.0092	0.0006	0.0073	0.0181
% of all FIs' Total Assets:	77.51%			
Panel D: June 30, 2015				
Book Value of Assets	1,546,362	1,068,044	1,525,649	1,883,981
Leverage	0.9265	0.9063	0.9393	0.9432
Market Capitalization, E	70,998	1,823	52,161	88,712
Equity Volatility, σ	0.2287	0.1969	0.2156	0.2583
Implied Market Value of Assets, a	1,500,492	1,093,359	1,433,546	1,826,187
Implied Volatility of Assets, v	0.0096	0.0004	0.0104	0.0191
% of all FIs' Total Assets:	76.83%			

Note: All dollar amounts are all in millions.

77.34% of all FIs' total assets in June 1995, to 73.83% in June 2000, and then to 77.51% in June 2007. Interestingly, even with global concern over FIs deemed too big to fail during the financial crisis of 2008, this number only dips slightly to 76.83% by June 2015.

Summary Statistics

We present basic summary statistics for the 20 largest FIs at various points in time. These summary statistics, given in Exhibit 1, consist of

1. *Book value of assets*, the total book value of each of the 20 FI's assets (in millions of dollars).
2. *Leverage*, the total face value of debt scaled by the total book value of the assets.
3. *Market capitalization*, E, the total market value of equity (in millions), calculated as the price per share times the number of shares outstanding.
4. *Equity volatility*, σ, the equity-return volatility based on a 130-day (i.e., six-month) lookback period.

5. *Implied market value of assets*, a, the implied market value of assets (in millions) based on the Black–Scholes formula for options valuation.
6. *Implied volatility of assets*, ν, the implied assets' return volatility based on the Black–Scholes formula for options valuation.
7. The total book value of the assets held by the 20 largest FIs as a percentage of the total book value of the assets held by all FIs.

EMPIRICAL ILLUSTRATIONS

We test our network risk framework on the financial data mined in the previous section. Recall that we have four models for systemic risk (models **C**, **D**, **G**, and **R**) within our overall framework. We compare these models in this section.

We determine systemic risk under each of our four models every six months (at the end of June and December) between 1995 and 2015. At each of these six-month intervals, we extract and analyze data for the top 20 FIs by total book value of assets, which, as we have noted, consistently accounts for approximately 75% of the aggregate assets of the more than 1,000 FIs we had available. For each of the four models, we plot the value of systemic risk over time, with each time series normalized to be in the range [0, 1], in Exhibit 2. First, this plot confirms that systemic risk spiked in the financial crisis of 2008. We also see smaller conflagrations of systemic risk in 2000 and 2011. Second, we see that all the models generate time series that track each other closely, with pairwise correlations ranging from 90%–97% (with a mean of 95%). Therefore, even though the four models are derived in uniquely different ways, time variation in the systemic risk score in these models is very much the same, implying that our systemic risk framework is robust to model choice.

It is also useful to look at the institution risk measure to see which FIs contributed the most to systemic risk. This is shown in Exhibit 3 using model **G** in 2007 and 2014. We can see that in 2007 mortgage-related FIs such as RBS Holdings (discontinued ticker ABNYY), Banco Santander (SAN), Federal Home Loan Mortgage Corp (FMCC), Fannie Mae (FNMA), Mitsubishi Trust (MTU), and Lehman Brothers (LEHMQ) were the top systemically risky firms. In 2014, the top systemic risk contributors were Mizuho Financial Group (ticker MFG), Lloyds Banking Group (LYG), Royal Bank of Scotland (RBS), Mitsubishi Trust (MTU), Sumitomo Mitsui Financial Group (SMFG), and Barclays (BCS). From both plots, we see that risk contributions are concentrated in a few banks. Furthermore, mortgage-related firms were more systemically risky in 2007, whereas in 2014, the traditional large banks were salient contributors of systemic risk.

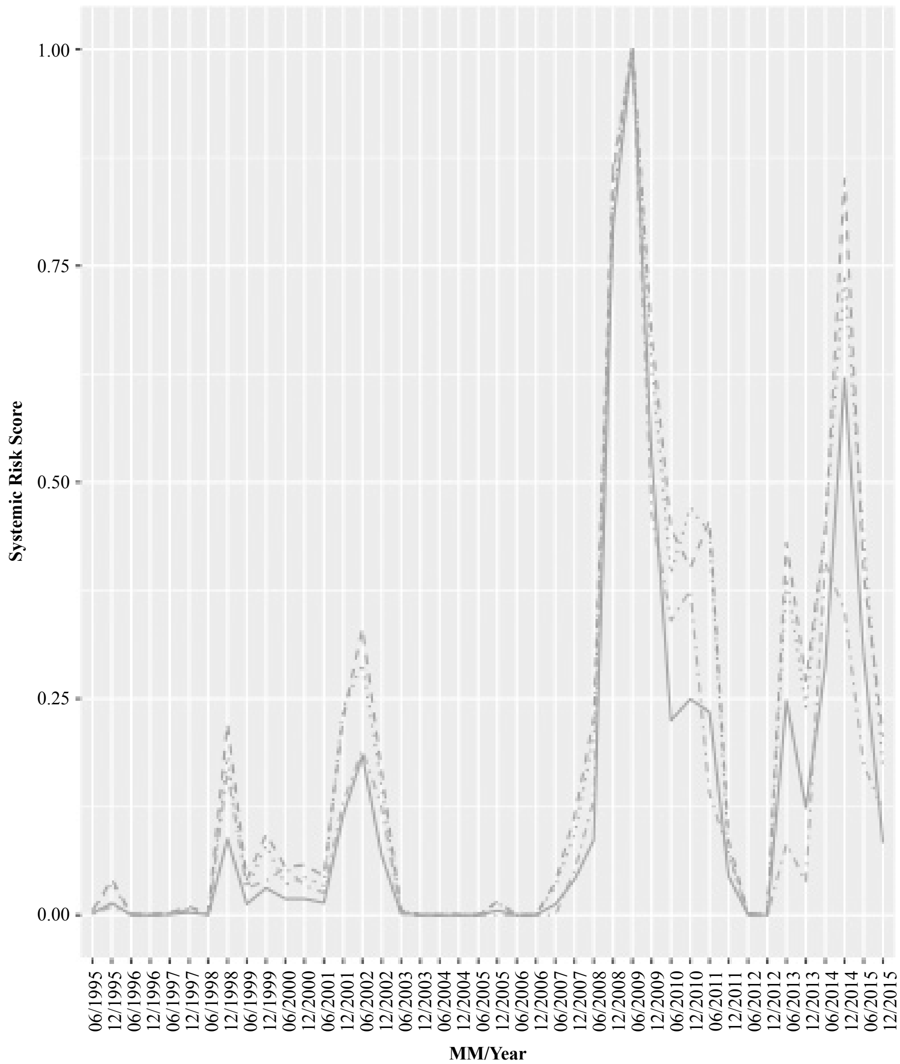
We checked that the institution risk measure rankings are similar across the four models. The top few names remain very much the same, irrespective of which model is used. In particular, the top five systemically risky FIs are the same in all four models, although not in the same order. These are Royal Bank of Scotland (RBS), Lloyds (LYG), Mizuho (MFG), Mitsubishi (MTU), and Sumitomo Mitsui (SMFG). Thus, there are two UK banks and three Japanese banks. Post-crisis measures in the United States may have reduced these banks' systemic risk levels.

Exhibit 4 extends this consistency check by displaying the union of the four models' top five risky institutions in each six-month interval. We note that in each interval there are between 5 and 13 FIs, where 5, of course, represents complete agreement among the four models and 20, of course, is the maximum possible number of FIs in the union. The average number of FIs is 6.45, showing considerable consistency among the four models in determining the top risky FIs.

We see Lehman Brothers (LEHMQ) appear consistently as a top systemically risky institution up until its demise in 2008. Around the time of the financial crisis in 2008, we also see Fannie Mae (FNMA) and the Federal Mortgage Credit Corporation (FMCC) show up as key contributors to systemic risk. Interestingly, though, these institutions were beginning to appear in the top risky list in 2003, suggesting that our methodology may have been able to provide an early warning about these mortgage-related institutions and their role in the systemic risk of the financial system.

In the latter time periods from our sample, we see Lloyds (LYG), Royal Bank of Scotland (RBS), Bank of America (BAC), and Deutsche Bank (DB) appear consistently, reflecting the fact that these institutions have been troubled in the last few years. Other large US banks that appear regularly, as is to be expected, are Citigroup (C), J.P. Morgan (JPM), and Morgan Stanley (MS). Many Japanese banks also appear, such

Exhibit 2
Systemic Risk over Time (1995–2015)

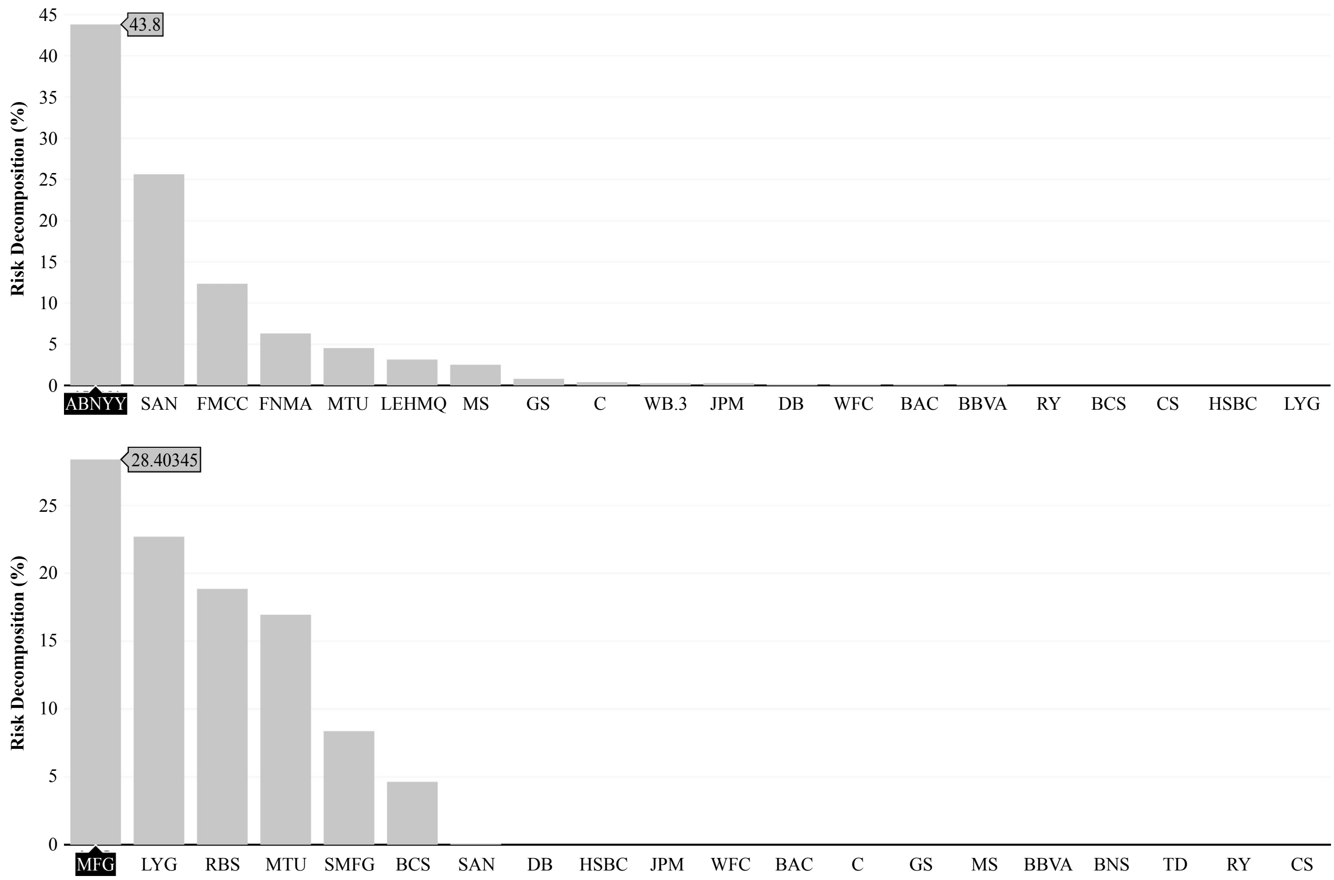


Notes: The plot shows systemic risk computed from data for the top 20 FIs (by assets). All four models, C (dashed line), D (dotted line), G (dotted dashed line), and R (solid line), are represented. The average correlation between all four models' time series is 95%.

as Mitsubishi (MTU), Mizuho (MFG), and Sumitomo Mitsui (SMFG).

We can also investigate the links between institutions that contribute the most to systemic risk in each six-month interval. Exhibit 5 illustrates this for model R. We see the same SIFIs that show up in Exhibit 4, but in this graphic, we show links (pairs of FIs) rather than individual FIs. As expected, up to the crisis we see Lehman (LEHMQ) appear on a regular basis, both as affecting other FIs and being affected by others. Santander (SAN) appears on both sides of links throughout the sample. Morgan Stanley (MS) seems to be at the receiving end of most links in which it appears. In the latter third of the sample, Mitsubishi (MTU) and Mizuho (MFG), both Japanese banks, demonstrate mutual systemic spillover risk to each other. They are also connected to another Japanese FI, Sumitomo Mitsui (SMFG). These examples illustrate that, in addition to designating individual SIFIs,

Exhibit 3
Institution Risk Measures



Notes: We display the institution risk measure using model G. This decomposes the systemic risk by institution. The upper plot is for December 2007, and the lower is for December 2014.

our model may also be used to designate systemically risky relationships.

We may wish to explore how sensitive the systemic risk measure is both, to changes in the financial strength of the FIs and to changes in the strength of the connections between the FIs. Specifically, we explore the changes in our systemic risk measures when we impose a blanket-wide increase in all the PD values (i.e., all PDs, λ_i) and when we impose a blanket-wide decrease or increase in all the pairwise correlations (i.e., all the ρ_{ij}, subject, of course, to remaining within the interval $[-1,1]$). In Exhibit 6, we demonstrate the effect of these changes at two snapshots in time: December 29, 2000 and December 31, 2007. We see from the exhibit that reasonable changes in either the PD values or in the correlation values affect the systemic risk score, mirroring the importance of considering the strength of both the individual FIs and the interconnections between the FIs in calculating systemic risk.

Finally, we consider the deficiency of return-based models highlighted by Löffler and Rapauch (2018). They showed that many of these popular models permitted banks to take on more risk, thereby raising overall systemic risk but at the same time reducing their own risk contribution relative to others, sometimes to the extent that their systemic risk contribution would even decline. We examine whether our model suffers from such a deficiency by increasing an FI's PD by 1% while holding all the other FIs' PD values frozen and then calculating how much the FI's institution risk

Exhibit 4
Top SIFIs

19950630	C	BT.2	FFB	CMB.1	GWF.	BK	MS						
19951229	BPC.2	C	CMB.1	JPM	GWF.								
19960628	C	FNMA	MS	JPM	FMCC	GWF.							
19961231	FNMA	MS	LEHMQ	C	FMCC								
19970630	GWF.	LEHMQ	MS	C	FMCC								
19971231	NW.1	LEHMQ	MS	SAN	C	BSC.1							
19980630	SAN	C	BSC.1	LEHMQ	BMO								
19981231	NW.1	LEHMQ	SAN	MS	JPM	C							
19990630	SAN	LEHMQ	MS	BSC.1	C								
19991231	NW.1	SAN	LEHMQ	BSC.1	FMCC	MS							
20000630	SAN	AFS	LEHMQ	FMCC	MS	FNMA							
20001229	SAN	LEHMQ	MS	BSC.1	JPM								
20010629	SAN	MS	LEHMQ	BSC.1	C	BAC							
20011231	ABNYY	SAN	MS	LEHMQ	AXP								
20020628	IMI.2	ABNYY	SAN	JPM	MS	C							
20021231	ABNYY	SAN	IMI.2	JPM	C								
20030630	ABNYY	JPM	MS	FNMA	LEHMQ	C	GS						
20031231	FMCC	FNMA	MS	JPM	GS	C							
20040630	DB	LEHMQ	BBVA	IMI.2	SAN	FNMA	BCS	ABNYY	MS	GS	C	JPM	
20041231	MS	NMR	BBVA	ABNYY	CS	C	FNMA	JPM	DB	LEHMQ	GS	FMCC	BCS
20050630	FNMA	FMCC	MS	BCS	ABNYY	C	JPM	LEHMQ	GS				
20051230	ABNYY	FNMA	FMCC	BCS	SAN	CS	LEHMQ	GS	C	JPM			
20060630	FNMA	LEHMQ	DB	GS	SAN	C	JPM						
20061229	LEHMQ	BSC.1	ABNYY	SAN	BBVA	MS	C	FNMA	DB	BAC	BCS		
20070629	MTU	LEHMQ	MS	GS	ABNYY	FNMA							
20071231	ABNYY	SAN	FNMA	FMCC	MTU								
20080630	MTU	FNMA	FMCC	DB	C	MS							
20081231	MTU	FMCC	FNMA	C	BAC	DB							
20090630	MFG	FNMA	MTU	C	BAC	DB							
20091231	MFG	MTU	NMR	FNMA	FMCC	C							
20100630	MTU	MFG	NMR	CS	BBVA	DB	SAN						
20101231	MFG	MTU	BBVA	SAN	CS	DB							
20110630	MFG	MTU	MS	BBVA	SAN								
20111230	BAC	C	MS	DB	BBVA	WFC							
20120629	BAC	MS	C	JPM	BBVA								
20121231	BBVA	DB	BAC	C	MS	SAN							
20130628	LYG	MFG	SMFG	RBS	MS	BBVA							
20131231	LYG	SMFG	MFG	SAN	BCS	HSBC	JPM	DB	BAC	C			
20140630	LYG	BCS	SMFG	BBVA	BAC								
20141231	RBS	LYG	MFG	MTU	SMFG								
20150630	RBS	LYG	SAN	MTU	SMFG								
20151231	RBS	SAN	BCS	DB	MS	BAC							

Notes: The graphic shows the FIs that contribute the most to systemic risk every half year in the sample across all four models. Each row displays the union of each of the four models' top five FIs that contribute the most risk. If the FIs are the same across all models, we will see exactly five FIs listed in a row; if not, then a few more will appear. One can see high agreement across models because the average number of firms in the rows is only 6.45.

EXHIBIT 5
Top Risky Links

19950630	BT.2:C	C:BT.2	FFB:C	C:FFB	C:CMB.1
19951229	BPC.2:C	C:BPC.2	BPC.2:CMB.1	BPC.2:JPM	CMB.1:BPC.2
19960628	C:FNMA	FNMA:C	C:JPM	JPM:C	FMCC:JPM
19961231	LEHMQ:MS	MS:LEHMQ	C:FNMA	FMCC:FNMA	FNMA:C
19970630	GWF.:C	GWF.:FMCC	C:FNMA	GWF.:FNMA	BK:FNMA
19971231	LEHMQ:MS	MS:LEHMQ	SAN:NW.1	NW.1:SAN	BSC.1:SAN
19980630	SAN:C	BSC.1:C	C:SAN	BSC.1:FMCC	FMCC:C
19981231	LEHMQ:NW.1	SAN:NW.1	NW.1:LEHMQ	NW.1:SAN	NW.1:MS
19990630	BSC.1:SAN	LEHMQ:SAN	LEHMQ:MS	SAN:LEHMQ	SAN:C
19991231	SAN:NW.1	NW.1:SAN	LEHMQ:NW.1	NW.1:LEHMQ	BSC.1:NW.1
20000630	AFS:SAN	SAN:FNMA	SAN:FMCC	LWHMQ:MS	FMCC:FNMA
20001229	LEHMQ:MS	LEHMQ:SAN	SAN:LEHMQ	BSC.1:SAN	SAN:JPM
20010629	SAN:MS	MS:SAN	LEHMQ:MS	LEHMQ:SAN	SAN:LEHMQ
20011231	SAN:ABNYY	ABNYY:SAN	ABNYY:MS	MS:ABNYY	SAN:MS
20020628	IMI.2:ABNYY	SAN:ABNYY	ABNYY:SAN	IMI.2:SAN	ABNYY:IMI.2
20021231	SAN:ABNYY	IMI.2:SAM	ABNYY:SAN	SAN:JPM	SAN:IMI.2
20030630	ABNYY:JPM	JPM:ABNYY	ABNYY:FNMA	FNMA:ABNYY	LEHMQ:MS
20031231	FMCC:FNMA	FNMA:FMCC	JPM:FNMA	GS:MS	FNMA:JPM
20040630	DB:ABNYY	LEHMQ:MS	FNMA:DB	DB:FNMA	MS:LEHMQ
20041231	MS:GS	MS:LEHMQ	LEHMQ:MS	GS:MS	MS:FNMA
20050630	FMCC:FNMA	FNMA:FMCC	LEHMQ:MS	MS:GS	MS:LEHMQ
20051230	FMCC:ABNYY	ABNYY:FMCC	ABNYY:BCS	ABNYY:JPM	ABNYY:CS
20060630	FNMA:DB	DB:FNMA	LEHMQ:GS	GS:LEHMQ	FNMA:C
20061229	BSC.1:LEHMQ	LEHMQ:BSC.1	LEHMQ:MS	DB:C	DB:ABNYY
20070629	LEHMQ:MTU	MS:MTU	MTU:MS	LEHMQ:GS	MTU:LEHMQ
20071231	SAN:ABNYY	ABNYY:SAN	FMCC:ABNYY	FNMA:ABNYY	FNMA:SAN
20080630	FNMA:MTU	FMCC:MTU	MTU:FNMA	MTU:FMCC	FNMA:FMCC
20081231	FMCC:MTU	FNMA:MTU	C:MTU	MTU:DB	FMCC:DB
20090630	MFG:MTU	FNMA:MTU	FNMA:MFG	MTUIMFG	C:MTU
20091231	MFG:MTU	NMR:MTU	FNMA:MTU	FMCC:MTU	MTU:MFG
20100630	MFG:MTU	MTU:MFG	NMR:MTU	NMR:MFG	BBVA:MTU
20101231	MFG:MTU	MTU:MFG	BBVA:MTU	BBVA:MFG	SAN:MTU
20110630	MFG:MTU	MTU:MFG	MS:MTU	MS:MFG	BBVA:MTU
20111230	C:BAC	BAC:C	MS:BAC	BAC:DB	MS:DB
20120629	MS:BAC	C:BAC	BAC:C	BAC:JPM	JPM:BAC
20121231	BBVA:BAC	C:DB	MS:BAC	DB:C	BAC:BBVA
20130628	LYG:MFG	SMFG:MFG	MFG:SMFG	LYG:SMFG	MFG:LYG
20131231	LYG:SMFG	SMFG:LYG	SMFG:MFG	LYG:MFG	MFG:SMFG
20140630	LYG:BCS	SMFG:BCS	LYG:SMFG	BCS:SMFG	BCS:LYG
20141231	RBS:MTU	LYG:MTU	MFG:MTU	SMFG:MTU	RBS:MFG
20150630	LYG:RBS	SAN:RBS	RBS:SAN	LYG:SAN	RSB:LYG
20151231	RBS:SAN	SAN:RBS	SAN:BCS	RBS:BCS	BCS:SAN

Notes: The graphic shows the five links with the highest connectedness risk measure in each six-month interval according to model R. The links are listed in the form i:j for a directed link from institution i to institution j.

EXHIBIT 6

Percentage Changes in Systemic Risk Measures

20 Largest FIs: December 29, 2000

Weighted Average Pd = 1.83%				
Weighted Average Pairwise Correlation (Corr) = 0.2286				
	PD + 0.1%	**PD + 1%**	**Corr – 0.2**	**Corr + 0.2**
Model R	+4.91%	+55.55%	–7.18%	+9.23%
Model C	+4.24%	+42.51%	–6.68%	+3.24%
Model E	+2.62%	+27.87%	–3.37%	+3.91%
Model G	+9.62%	+84.81%	–	–

20 Largest FIs: December 31, 2007

Weighted Average Pd = 3.76%				
Weighted Average Pairwise Correlation (Corr) = 0.3065				
	PD + 0.1%	**PD + 1%**	**Corr – 0.2**	**Corr + 0.2**
Model R	+3.10%	+36.64%	–4.70%	+5.01%
Model C	+2.22%	+22.29%	–5.98%	+1.77%
Model E	+1.46%	+15.24%	–1.86%	+1.74%
Model G	+8.57%	+73.18%	–	–

Notes: This exhibit demonstrates how the systemic risk score changes with changes in the PD or changes in the strength of the network structure.

measure changes compared to each of the other FIs. Exhibit 7 shows this effect for the top 20 FIs in 2007 and for the top 20 FIs in 2014. At both times, for each of the 20 FIs, we see from the exhibit that the FI's own institution risk measure increases more than that of the other FIs, because, for each row, the values on the diagonal are higher than the other values. A closer analysis of the data used to create the exhibit shows that the other FIs' institution risk measure actually decreases generally, and the highest increase in the data is only about half of the increase of the FI whose PD is increased. This indicates that our metric is not susceptible to gaming by any one bank.

CONCLUDING COMMENTS

Using data science and modeling tools from the social networks arena, we capture the systemic risk of a financial system in a Merton-on-a-network model that includes three important determining elements: (1) connectedness (via banking networks), (2) joint default risk (from an extension of the Merton 1974 model), and (3) size (i.e., the market value of a bank's assets, also implied from the Merton model). We define and analyze four important properties of our systemic risk measure and develop four different models that generally have these properties.

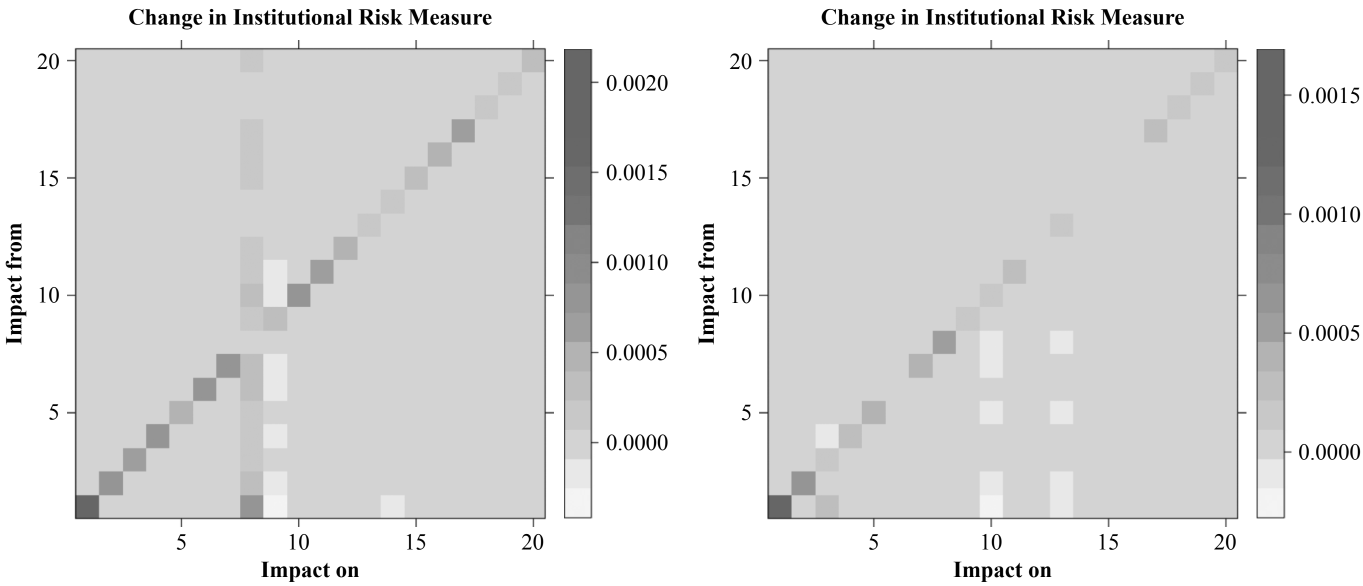
Empirical examination demonstrates that systemic risk, as well as the risk assigned to individual banks within the system, are similar across these four models, suggesting that the framework is robust to implementation design, in contrast to conflicting findings about other systemic risk measures, as shown by Benoit et al. (2013).[10] The metric also does not appear to suffer from the deficiency noted by Löffler and Rapauch (2018).

The current model supports many theoretical and empirical extensions. For example, whereas the model setting is that of the financial system, we may embed this model within a broader general equilibrium model of the entire economy, either by adding other sectors or by making the financial system variables functions of the broader macroeconomy.

[10] This article found systemic risk results to vary markedly across the four models they surveyed, namely marginal expected shortfall and SES, both from Acharya et al. (2017); the systemic risk measure from Acharya, Engle, and Richardson (2012) and Brownlees and Engle (2012); and the ΔCoVaR from Adrian and Brunnermeier (2016).

EXHIBIT 7

Spillover Risk–Change in Institutional Risk Measures



Notes: We see how much a single bank's increase in its PD affects its institution risk measure (i.e., its contribution to systemic risk) in comparison to that of the other banks. The left panel is for 2007 and the right for 2014. This experimental analysis was done for the case of model G. The largest numbers are on the diagonal, indicating that an increase to a bank's own PD increases its institution risk measure more than it increases any of the other 19 banks' institution risk measures. The diagonal values are higher than the off-diagonal values, which are mostly indistinguishable from zero. Also note that the difference in increases are more marked for 2007 before the crisis than they were for 2014.

Furthermore, we are able to extract the time series for systemic risk, which may be related to macroeconomic variables and events. Our framework supports objective real-time measurement of systemic risk, identification of SIFIs, and identification of systemically important connections between FIs so that the system may be analyzed, monitored, and controlled by regulators. The article demonstrates the efficacy of open big data in conjunction with data science techniques in risk management.

APPENDIX

PROOFS OF MODEL PROPERTIES

Financial Properties for the Homogenous Models C, D, and G

All four desired financial properties for $\mathcal{S}$ hold in models **C**, **D**, and **G**, as we next proceed to establish.

Property 1: All other things being equal, $\mathbb{S}$ is minimized by dividing the credit risk equally among the n FIs and is maximized by putting all the credit risk into one institution. To make all other things be equal, we set the total assets, $\Sigma_{i=1}^{n} a_i = \mathbf{1}^T \mathbf{a}$, constant; set the total credit risk, $\Sigma_{i=1}^{n} c_i = \mathbf{1}^T \mathbf{c}$, equal to a constant, c_{total}; and set M_{ij} equal to the same number, m, if $i \neq j$ while, of course, keeping $M_{ii} = 1$ for all i. For the singular case in which $m = 1$, all the institutions act like a single institution, and so it makes no difference to $\mathcal{S}$ how the credit risk is spread among the institutions. For the general case in which $m < 1$, from the definition of $\mathcal{S}$ in Equation 10, we see that maximizing or minimizing $\mathcal{S}$ now corresponds to maximizing or minimizing $\mathbf{c}^T \mathbf{M} \mathbf{c} = \Sigma_{i=1}^{n} c_i^2 + m \Sigma_{i=1}^{n} \Sigma_{j \neq i} c_i c_j$, subject to the restriction that $\mathbf{1}^T \mathbf{c} = \Sigma_{i=1}^{n} c_i = c_{total}$.

Because $m < 1$, it is clear that $\Sigma_{i=1}^{n} c_i^2 + m \Sigma_{i=1}^{n} \Sigma_{j \neq i} c_i c_j \leq \Sigma_{i=1}^{n} c_i^2 + \Sigma_{i=1}^{n} \Sigma_{j \neq i} c_i c_j = (\Sigma_{i=1}^{n} c_i)^2 = c_{total}^2$. However, if all the credit risk is put into one institution, we have $\Sigma_{i=1}^{n} c_i^2 + m \Sigma_{i=1}^{n} \Sigma_{j \neq i} c_i c_j = c_{total}^2$, the highest possible value, and so $\mathcal{S}$ is maximized when all the credit risk is concentrated into one FI.

On the other hand, the Lagrange multiplier method tells us that we have minimized $\Sigma_{i=1}^{n} c_i^2 + m \Sigma_{i=1}^{n} \Sigma_{j \neq i} c_i c_j$ subject to the restriction $\Sigma_{i=1}^{n} c_i = c_{total}$ when (denoting the Lagrange multiplier by λ),

$$\frac{\partial}{\partial c_k}\left(\sum_{i=1}^{n} c_i^2 + m\sum_{i=1}^{n}\sum_{j\neq i}^{n} c_i c_j\right) = \lambda \frac{\partial}{\partial c_k}\sum_{i=1}^{n} c_i \quad \text{where } k = 1, 2, \ldots\ n$$

and

$$\sum_{i=1}^{n} c_i = c_{total}$$

The first n equations give us that $c_1 = c_2 = \ldots = c_n = \frac{\lambda - 2mc_{total}}{2(1-m)}$. That is, when $\mathcal{S}$ is minimized, all c_i have the same value. The second equation then tells us that each $c_i = \frac{c_{total}}{n}$, and so we have that $\mathcal{S}$ is minimized by dividing the credit risk equally among the n institutions.

Property 2: $\mathbb{S}$ should increase as the institutions' defaults become more connected. Consider the case in which $\mathbf{a}$ and $\mathbf{c}$ are both held constant so that $\mathcal{S}$ only depends on $\mathbf{M}$, specifically through the expression

$$\mathbf{c}^T\mathbf{M}\mathbf{c} = \sum_{i=1}^{n}\sum_{j=1}^{n} c_i M_{ij} c_j$$

in the numerator of our model's definition of $\mathcal{S}$. Clearly, the bigger the values of M_{ij}, the larger $\mathcal{S}$ becomes. Because M_{ii} must always equal 1, $\mathcal{S}$ is minimized when $\mathbf{M} = \mathbf{I}$, the identity matrix, and is maximized when the components of the $\mathbf{M}$ matrix are all ones. We note that when $\mathbf{M} = \mathbf{I}$ $\sqrt{\mathbf{c}^T\mathbf{M}\mathbf{c}} = \sqrt{\Sigma_{i=1}^{n} c_i^2} = ||c||_2$, the 2-norm of the vector $\mathbf{c}$, whereas when $\mathbf{M}$ is all ones, $\sqrt{\mathbf{c}^T\mathbf{M}\mathbf{c}} = \Sigma_{i=1}^{n} c_i = ||c||_1$, the 1-norm of the vector $\mathbf{c}$.

Property 3: If all the assets, a_i, are multiplied by a common factor, $\alpha > 0$, it should have no effect on $\mathbb{S}$. In our model, if we replace each a_i with αa_i, we then replace $\sqrt{\mathbf{c}^T\mathbf{M}\mathbf{c}}$ by $\alpha\sqrt{\mathbf{c}^T\mathbf{M}\mathbf{c}}$ and replace $\mathbf{1}^T\mathbf{a}$ with $\alpha\mathbf{1}^T\mathbf{a}$. Because the α then cancel in the expression for $\mathcal{S}$ from Equation 10, we have the desired property that systemic risk is unchanged.

Property 4: Substanceless partitioning of an institution into two institutions should have no effect on $\mathbb{S}$. If institution i's assets are artificially divided into two institutions of size γa_i and $(1 - \gamma)a_i$ for some $\gamma \in [0, 1]$, where both of these new institutions are completely connected to each other and both have the same connections with the other banks that the original institution did, then this division is without substantive meaning and should not affect the value of $\mathcal{S}$. Without loss of generality, we can let the index of the divided institution $i = n$, so, in our model, the new $(n + 1)$-vector $\mathbf{c}$ is

$$\mathbf{c} = \begin{bmatrix} c_1 \\ \vdots \\ c_{n-1} \\ \gamma c_n \\ (1-\gamma)c_n \end{bmatrix}$$

and the new $(n + 1) \times (n + 1)$ matrix $\mathbf{M}$ is

$$\mathbf{M} = \begin{bmatrix} 1 & M_{12} & \cdots & M_{1(n-1)} & M_{1n} & M_{1n} \\ M_{21} & \ddots & \ddots & \vdots & \vdots & \vdots \\ \vdots & \ddots & \ddots & M_{(n-2)(n-1)} & \vdots & \vdots \\ M_{(n-1)1} & \cdots & M_{(n-1)(n-2)} & 1 & M_{(n-1)n} & M_{(n-1)n} \\ M_{n1} & \cdots & \cdots & M_{n(n-1)} & 1 & 1 \\ M_{n1} & \cdots & \cdots & M_{n(n-1)} & 1 & 1 \end{bmatrix}$$

where we note that $M_{(n+1)n} = M_{n(n+1)} = 1$ to reflect the fact that both of the new institutions are completely connected to each other. A quick computation shows that the new $\sqrt{\mathbf{c}^T\mathbf{M}\mathbf{c}}$ is equal to the old $\sqrt{\mathbf{c}^T\mathbf{M}\mathbf{c}}$, and because $a_1 + \ldots + a_n = a_1 + \ldots + a_{(n-1)} + \gamma a_n + (1 - \gamma)a_n$, we also have that the new $\mathbf{1}^T\mathbf{a}$ is equal to the old $\mathbf{1}^T\mathbf{a}$. Therefore, the value of $\mathcal{S}$ in Equation 10 is unchanged, and our model has this desired property.

Financial Properties for the Nonhomogeneous Model R

Property 1: All other things being equal, $\mathbb{S}$ is minimized by dividing the risk equally among the n FIs and is maximized by putting all the risk into one institution. Paralleling our approach in the previous section, we hold the total assets, $\Sigma_{i=1}^{n} a_i = \mathbf{1}^T\mathbf{a}$, constant and hold the total risk, $\Sigma_{i=1}^{n}\rho_i = \mathbf{1}^T\boldsymbol{\rho}$, equal to a constant. If we replace $\mathbf{c}$ and $\mathbf{M}$ in the model from the previous section for $\mathcal{S}$ given in Equation 10 with $\boldsymbol{\rho}$ and the identity matrix $\mathbf{I}$, we get our new model for $\mathcal{S}$ in Equation 13. Therefore, the proof of Property 1 from the previous section with $m = 0$ also establishes Property 1 for the model of $\mathcal{S}$ in Equation 13.

We note that if the numerator in the definition of $\mathcal{S}$ in Equation 13 were $\Sigma_{i=1}^{n}\rho_i$, the 1-norm of $\boldsymbol{\rho}$, instead of $\sqrt{\boldsymbol{\rho}^T\boldsymbol{\rho}}$, the 2-norm of $\boldsymbol{\rho}$, we would lose Property 1.

Property 2: $\mathbb{S}$ should increase as the institutions' defaults become more connected. An increasing connection means M_{ij} is increasing, which, from Equation 12, means that ρ_i increases. As any ρ_i increases, we have from Equation 13 that $\mathcal{S}$ increases, assuming, as we also did in the previous section, that $\mathbf{a}$ is held constant.

Property 3: If all the assets, a_i, are multiplied by a common factor, $\alpha > 0$, it should have no effect on $\mathcal{S}$. In our model, if we replace each a_i with αa_i, we replace $\sqrt{\rho^T \rho}$ by $\alpha\sqrt{\rho^T \rho}$, and we replace $\mathbf{1}^T\mathbf{a}$ with $\alpha\mathbf{1}^T\mathbf{a}$. Because the α then cancel in the expression for $\mathcal{S}$ given in Equation 13, we have the desired property that systemic risk is unchanged.

Property 4: Substanceless partitioning of an institution into two institutions should have no effect on $\mathbb{S}$. This property does not hold. Let's say we artificially divide institution n's assets into two institutions, call them institution n_{new} and institution $(n + 1)_{new}$, of size γa_n and $(1 - \gamma)a_n$. Because the division is artificial, $M_{n_{new}n_{new}} = M_{(n_{new}+1)n_{new}} = M_{n_{new}(n_{new}+1)} = M_{(n_{new}+1)(n_{new}+1)}$, which all equal M_{nn}, where n again represents the divided institution before it was divided, and, for any $i < n$, $M_{n_{new}i} = M_{(n_{new}+1)i} = M_{in_{new}} = M_{i(n_{new}+1)}$ equals $M_{ni} = M_{in}$.

From Equation 12, we see that the ρ_i are unchanged for $i = 1, 2, ..., n$. However, an extra $(n + 1)$th component now has been added to the vector $\boldsymbol{\rho}$, where $\rho_{n+1} = \rho_n$, which must increase the norm of $\boldsymbol{\rho}$, which must increase the systemic risk $\mathcal{S}$ in Equation 13. Therefore, artificial division of a FI increases $\mathcal{S}$ instead of having no effect on it.

ACKNOWLEDGMENTS

We are grateful for comments and input from Frank Fabozzi; Deborah Lucas; Robert Merton; Hersh Shefrin; and participants in talks at MIT's Center for Financial Policy, the Big Data in Finance Conference at University of Michigan, Box, CapitalOne, BlackRock, Berkeley's Center for Data Analytics and Risk, RFinance Conference Chicago, EDHEC Business School, McGill University, University of Delaware, AI and Machine Learning Conference (London), and the Open Data Science Conference (ODSC West 2017, San Francisco).

REFERENCES

Abbass, P., C. Brownlees, C. Hans, and N. Polich. 2016. "Credit Risk Interconnectedness: What Does the Market Really Know." *Journal of Financial Stability* 29: 1–12.

Acharya, V., R. F. Engle, and M. Richardson. 2012. "Capital Shortfall: A New Approach to Ranking and Regulating Systemic Risks." *American Economic Review* 102 (3): 59–64.

Acharya, V., L. H. Pedersen, T. Philippon, and M. Richardson. 2017. "Measuring Systemic Risk." *The Review of Financial Studies* 30 (1): 2–47.

Adrian, T., and M. Brunnermeier. 2016. "CoVaR." *American Economic Review* 106 (7): 1705–1741.

Avramidis, P., and F. Pasiouras. 2015. "Calculating Systemic Risk Capital: A Factor Model Approach." *Journal of Financial Stability* 16: 138–150.

Benoit, S., J. E. Colliard, C. Hurlin, and C. Perignon. 2013. "A Theoretical and Empirical Comparison of Systemic Risk Measures." *Working Paper #FIN-2014-1030*, HEC Paris.

Billio, M., M. Getmansky, A. Lo, and L. Pelizzon. 2012. "Econometric Measures of Connectedness and Systemic Risk in the Finance and Insurance Sectors." *Journal of Financial Economics* 104 (3): 536–559.

Black, L., R. Correa, X. Huang, and H. Zhou. 2016. "The Systemic Risk of European Banks during the Financial and Sovereign Debt Crises." *Journal of Banking and Finance* 63: 107–125.

Bluhm, M., and J. P. Krahnen. 2014. "Systemic Risk in an Interconnected Banking System with Endogenous Asset Markets." *Journal of Financial Stability* 13: 75–94.

Brownlees, T. C., and R. F. Engle. 2012. "Volatility, Correlation and Tails for Systemic Risk Measurement." Working paper, New York University.

Burdick, D., S. Das, M. A. Hernandez, H. Ho, G. Koutrika, R. Krishnamurthy, L. Popa, I. Stanoi, and S. Vaithyanathan. 2011. "Extracting, Linking and Integrating Data from Public Sources: A Financial Case Study." *IEEE Data Engineering Bulletin* 34 (3): 60–67.

Das, S. 2016. "Matrix Metrics: Network-Based Systemic Risk Scoring." *The Journal of Alternative Investments* 18 (4): 33–51.

De Nicolo, G., G. Favara, and L. Ratnovski. 2012. "Externalities and Macroprudential Policy." Discussion note 12/05, International Monetary Fund.

Espinosa-Vega, M., and J. Sole. 2010. "Cross-Border Financial Surveillance: A Network Perspective." Working paper WP/10/105, International Monetary Fund.

Giglio, S., B. Kelly, and S. Pruitt. 2016. "Systemic Risk and the Macroeconomy: An Empirical Evaluation." *Journal of Financial Economics* 119 (3): 457–471.

Huang, X., H. Zhou, and H. Zhu. 2012. "Systemic Risk Contributions." *Journal of Financial Services Research* 42 (1): 55–83.

Hüser, A. C. 2015. "Too Interconnected to Fail: A Survey of the Interbank Networks Literature." *Journal of Network Theory in Finance* 1 (3): 1–50.

Kritzman, M., Y. Li, S. Page, and R. Rigobon. 2011. "Principal Components as a Measure of Systemic Risk." *The Journal of Portfolio Management* 37 (4): 112–126.

Levy-Carciente, S., D. Y. Kenett, A. Avakian, H. E. Stanley, and S. Havlin. 2015. "Dynamical Macroprudential Stress Testing Using Network Theory." *Journal of Banking and Finance* 59: 164–181.

Löffler, G., and P. Rapauch. 2018. "Pitfalls in the Use of Systemic Risk Measures." *Journal of Financial and Quantitative Analysis* 53 (1): 269–298.

Merton, R. C. 1974. "On the Pricing of Corporate Debt: The Risk Structure of Interest Rates." *The Journal of Finance* 29 (2): 449–470.

Merton, R. C., M. Billio, M. Getmansky, D. Gray, A. W. Lo, and L. Pelizzon. 2013. "On a New Approach for Analyzing and Managing Macrofinancial Risks." *Financial Analysts Journal* 69 (2): 22–33.

Silva, W. H. Kimura, and V. A. Sobreiro. 2017. "An Analysis of the Literature on Systemic Financial Risk: A Survey." *Journal of Financial Stability* 28: 91–114.

Upper, C. 2011. "Simulation Methods to Assess the Danger of Contagion in Interbank Markets." *Journal of Financial Stability* 7: 111–125.