Journal Pre-proof

The intrafirm complexity of systemically important financial institutions

R.L. Lumsdaine, D.N. Rockmore, N.J. Foti, G. Leibon, J.D. Farmer

Please cite this article as: { doi: https://doi.org/

This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

The Intrafirm Complexity of Systemically Important Financial Institutions

R.L. Lumsdaine[(a)] robin.lumsdaine@american.edu, *D.N. Rockmore[(b,f)], N. J. Foti[(c)], G. Leibon[(b,d)], J.D. Farmer[(e,f)]*

[(a)]Kogod School of Business, American University; Erasmus University Rotterdam; National Bureau of Economic Research and Center for Financial Stability
[(b)]Department of Mathematics and Computer Science, Dartmouth College
[(c)]Paul G. Allen School of Computer Science & Engineering, University of Washington
[(d)]Coherent Path, Inc.
[(e)]Institute for New Economic Thinking at the Oxford Martin School and Mathematical Institute, University of Oxford
[(f)]The Santa Fe Institute

Corresponding author:
Robin L. Lumsdaine Kogod School of Business, American University 4400 Massashusetts Avenue, NW Washington, DC 20016-8044
+1-202-885-1964

Daniel N. Rockmore
Department of Mathematics, Dartmouth College 6188 Kemeny Hall Hanover, NH 03755
+1-603-646-3260
Daniel.N.Rockmore@Dartmouth.edu

Nicholas J. Foti
Paul G. Allen School of Computer Science & Engineering University of Washington Box 352355 Seattle, WA 98195-2355
+1-603-491-5649
nfoti01@gmail.com

Gregory Leibon
Dartmouth College HB6188 Hanover, NH 03755 +1-603-443-0242
gleibon@gmail.com

J. Doyne Farmer
Institute for New Economic Thinking at the Oxford Martin School Manor Road Building Manor Road Oxford OX1 3UQ United Kingdon doyne.farmer@inet.ox.ac.uk
+44-1865-610-403

In November, 2011, the Financial Stability Board, in collaboration with the International Monetary Fund, published a list of 29 "systemically important financial institutions" (SIFIs, now referred to as "globally systemically important banks" or G-SIBs), institutions whose failure, by virtue of "their

size, complexity, and systemic interconnectedness", could have dramatic negative consequences for the global financial system. While "size" and "interconnectedness" have been the subject of much quantitative analysis, less attention has been paid to measuring "complexity." Yet without a consistent way to measure complexity, there is little guarantee that the designated SIFIs capture the complexity that the FSB is concerned about, and little hope of mitigating the consequences that the FSB warns of. In this paper we propose the structure of an individual firm's *majority-control hierarchy* as a proxy for institutional complexity. We demonstrate as a proof-of-concept how this method might be used by bank supervisors, particularly the Federal Reserve under its authority as consolidated supervisor, using a data set containing information on the majority-control hierarchies of many of the designated SIFIs. Our mathematical intrafirm network representation (and various associated metrics we propose) provides a uniform way to compare firms with often very disparate organizational structures -- one that is distinct from a simple size comparison.

Keywords: SIFI, G-SIB, control hierarchy, macroprudential regulation, bank supervision, consolidated supervision
JEL classification codes: G28, G21, G01, C02

*

1. Introduction

The Financial Stability Board (FSB) describes a systemically important financial

institution, or SIFI, as a financial institution "whose disorderly failure, because of their size,

complexity and systemic interconnectedness, would cause significant disruption to the wider

financial system and economic activity" (Financial Stability Board 2011).[1] Developed in

the aftermath of the 2008 global financial crisis (hereafter the 'post-crisis period'), this

characterization represents an expanded regulatory definition relative to earlier ones based

primarily on size (e.g., the list of "mandatory banks" subject to the Basel II capital

regulations, see 72 FR 69298, December 7, 2007). While size-based thresholds are

appealing from a regulatory perspective in that they produce a dichotomous outcome,

---

[1] The definition used by others, including the Basel Committee on Bank Supervision (BCBS, 2014), includes two additional characteristics, namely the lack of readily available substitutes or financial institution infrastructure, and global (cross-jurisdictional) activity. Since our focus in this paper is on complexity, we motivate our topic using the simpler FSB definition but highlight that we are aware of the other definitions as well. We also note that the BCBS has advocated the use of an indicator for G-SIB designation that is an equally-weighted average of exposures in each of the five categories. For the determinants of complexity, they list: (1) notional amount of OTC derivatives, (2) trading and AFS securities, (3) level 3 assets, each with a 1/15 (6 2/3%) weight. Scores are then distributed into buckets with a capital add-on associated with the different buckets.

thereby being transparent and easy to understand and implement, they are overly simplistic in the presumption that risk can be evaluated via a single value.[2]

Although historically regulatory emphasis has been on the risk of a given organization, the collapse of Lehman Brothers in September 2008 highlighted the extensive interconnectedness of the financial system and the importance of considering not just the risk to a single firm but the risk to the entire financial system, i.e., the risk to financial stability.[3] Much of the research on interconnectedness has been formulated mathematically in terms of *networks.* The importance of such research for policymakers was highlighted in a May 10, 2013 speech by then-Fed Chairman Ben Bernanke: "Network analysis, yet another promising tool under active development, has the potential to help us better monitor the interconnectedness of financial institutions and markets" (Bernanke 2013).

Yet despite a large literature addressing the interrelationships among economic and financial network participants (e.g., Elsinger et al., 2006, Cohen-Cole et al., 2010, Haldane and May 2011, Adamic et al., 2012, Battiston et al., 2012, Billio et al., 2012, Hautsch et al., 2012, 2013, Kapadia et al., 2013, Squartini et al., 2013, Caccioli et al., 2014), there has been comparatively little development of metrics concerning the *complexity of the individual firms* that comprise the system – the other key attribute highlighted in the FSB's definition. Alessandri et al. (2015) point to the Lehman collapse as an example of how widespread financial instability resulted from the failure of a complex organization and go on to argue

---

[2] In their survey of metrics for quantifying systemic risk, Bisias et al. (2012) refer to reliance on a single number to assess risk as a "Maginot Line Strategy". They argue that risk is multifaceted and that therefore multiple metrics are required to capture risk.

[3] The link between SIFI designation of a single institution and global financial stability is well-summarized by Alessandri et al., 2015: "Systemic importance relates to the damage that the failure of a financial institution may cause to global financial stability, whereas systemic risk relates to the probability of default of an institution."

that a firm's complexity is a much more critical determinant of systemic importance than its size. Hampering the development of research related to a financial institution's complexity, in addition to data limitations (National Academy of Sciences 2010), is that "we lack a clear consensus on how to assess an entity's complexity" (Cetorelli, McAndrews, and Traina 2014, Cetorelli and Goldberg 2014). Failing any direct definition, one view of an individual firm's complexity comes from the lens of governance: "high complexity" would be interpreted as a corporate control structure rife with governance challenges for a firm's management, resulting in a lack of oversight that in turn poses significant operational, reputational, and balance sheet risk (Adams 2010, Vitali et al., 2011). This same complexity might present challenges to bank regulators, tasked with gathering information from a variety of sources to assess the systemic importance of the firm. This is particularly the case for the Federal Reserve, as it has the unique role of being the consolidated supervisor.[4] Organizational complexity contributes to the possibility that subsidiaries act in relative obscurity within the organization, thus hindering the regulators' ability to carry out effective consolidated supervision. In this context, complexity therefore poses risk to an accurate assessment of the organization's systemic importance and hence to financial stability.[5]

One way to describe the organizational structure of a firm is via its *control hierarchy* (Vitali et al., 2011), consisting of a (parent) company and all of its subsidiaries, considered in its natural hierarchical and networked arrangement. This is a standard representation of

---

[4] In providing guidance as to how to effectively carry out this role, the Bank Holding Company Supervision Manual (Board of Governors of the Federal Reserve System, 2017) notes that "The Federal Reserve's conduct of consolidated supervision is central to and dependent on the coordination with, and reliance on, the work of other relevant primary supervisors and functional regulators" (page 374, §2060.05.6.8).

[5] The idea that imperfect or incomplete information reduces the ability to mitigate risk is discussed in Battiston and Martinez-Jaramillo (2018).

the *intra*connectedness of a firm, along the lines of Coase (1937) who described a firm as a "system of relationships".[6] While *interfirm* financial network complexity is a well-studied subject (see e.g., Arinaminpathy et al. 2012, Caballero and Simsek 2013, Haubrich and Lo 2013 and references therein, Roukny et al. 2013, Battiston et al. 2016a,b, Aldasoro and Alves 2018, Berndsen et al. 2018, Constantin et al. 2018, León et al. 2018, Roukny et al. 2018), many of the network methods in that literature are not applicable to the very specific kinds of network topologies that characterize majority-control hierarchies, which are by construction *rooted directed trees* (in the parlance of computer science).

Complex majority-control hierarchies present difficulties for regulators tasked with supervisory oversight (Blair and Kushmeider 2006, National Academy of Sciences 2010, Viñals et al. 2013), particularly the Federal Reserve in its role as consolidated supervisor.[7] Greater complexity (in terms of organizational structure and business activities) of an individual firm makes it harder for a consolidated supervisor to disentangle and understand the firm's structure and increases the likelihood that some parts of the firm's activities and interrelationships go unnoticed. In the case of large multinational organizations, a complexity measure related to oversight would naturally account for the burdens posed by coordinating over multiple national and regulatory environments. Therefore, the identification of metrics that enable comparison across firms that may have very different

---

[6] In this seminal paper, Coase (1937) uses this description to argue that a firm cannot increase in size indefinitely and notes that as a firm grows, the losses due to mistakes will increase.   In a different context (law firm mergers), Briscoe and Tsai (2011) note that following a merger, there is a delay in value creation and coordination due to the time it takes for relationships to develop, a situation that is likely to be the case in the context of financial firms as well. A similar delay in identifying and mitigating systemic importance is likely to occur as a result of the need for supervisory coordination.

[7] Throughout this paper, we will use the terms "supervisor" and "regulator" interchangeably, to refer to an entity responsible for assessing some aspect of risk in an organization. In the context of our discussion, these references relate to external (e.g., governmental) assessment but in principle our proposed metrics may also be useful for risk managers within a firm.

majority-control hierarchies is of critical importance to consolidated supervision in the post-2008 financial crisis environment.[8]

The supervisory challenges we envision are likely what motivated the requirement that SIFIs indicate the number of "material" supervisors and regulators with whom they must interact in the resolution plans that they must file to comply with Title II of the Dodd-Frank Act.[9] In our case, it is not just the *number* of different supervisors and regulators that is relevant but *the way in which those supervisors and regulators have to interact with each other*. Thus our paper provides a natural extension to the existing understanding of the coordination challenges supervisors face, much as network analysis has advanced knowledge regarding interconnectedness by noting that it is not only the number of connections that is important but the way in which those connections arise.

Our paper is most closely related to the work of Carmassi and Herring (2014) who note that institutional complexity "...impeded effective oversight by the authorities ex ante and greatly complicated crisis management and the resolution of institutions ex post." Using Bankscope to obtain data on "majority-owned subsidiaries for which the G-SIB is the ultimate owner with a minimum control path of 50.01%" (i.e., a control definition that is similar to the one used in this paper), Carmassi and Herring (2014) define complexity as the number of subsidiaries but emphasize, "This is a very simplistic indicator of corporate complexity, but it remains the only indicator that can be measured with any degree of accuracy and even that is far from perfect." We agree with both their premise and

---

[8] This point is echoed by Viñals et al. (2013), "Recent evidence from the crisis does not implicate specific bank business models as susceptible to greater risk of failure. Nevertheless, structural measures could be a useful complement to traditional prudential tools under certain conditions. Targeting them to reflect firm-specific risk profiles increases their effectiveness relative to the one-size-fits-all approach envisaged by the recent structural reform proposals, albeit the targeted approach requires firm political commitment and support for supervisors."

[9] Carmassi and Herring (2014) also note in Chapter 4 that the FSB identified the need to cooperate as one of five remaining challenges that regulators faced in order for resolution to proceed in an orderly manner.

characterization of the existing metrics for assessing complexity and indeed, it is the dearth of metrics that motivates our investigation.

This paper proposes network-based metrics to encode organizational complexity. Ours is a novel approach that uses the innate network structure of a majority-control hierarchy. As we explain below, we see this representation, as well as the metrics we construct, as intimately related to the kinds of oversight/regulatory challenges that the Federal Reserve as consolidated supervisor might face. As such we hope that our metrics based on *intra*-firm complexity will be a useful addition to the more-commonly-studied *inter*-firm complexity (i.e., the interconnectedness across firms) metrics.

The metrics we propose also are intended to inform the Federal Reserve's judgment regarding the SIFI designation. The network encoding and associated metrics also admit the use of simulations as a means of assessing changes in complexity should a firm alter its business structure via a change in majority-control hierarchy. Such simulations could prove helpful for understanding the supervisory implications of altering a firm's majority-control hierarchy in the process of winding down a firm (such as in the case of the dismantling of Lehman Brothers), or in arranging a rapid acquisition, (e.g., in the cases of the JP Morgan Chase acquisitions of Bear Stearns and Washington Mutual, the Bank of America acquisition of Merrill Lynch, or the Wells Fargo acquisition of Wachovia). The goal would be to ensure continuity of consolidated supervision in the wake of a crisis, thereby fostering financial stability, a key component of the Federal Reserve's mandate.

Among the questions we address using network-based metrics are:

- Do the SIFI institutions in our sample appear to be more complex?

7

- Do size rankings match complexity rankings that utilize country and SIC (Standard Industrial Classification) code information?

- How has the complexity of the firms in our sample changed over time?

The paper proceeds as follows: Section 2 describes a small dataset that we will use to illustrate the type of insights that can be drawn using our proposed metrics. Section 3 introduces the metrics we employ. Section 4 provides a proof-of-concept by reporting results of our methods on the sample dataset. Section 5 contains additional discussion and Section 6 concludes.

2. Data

To illustrate how our proposed methods might be used in practice, we use an anonymized data set provided to us by Kingland Systems[10] of twenty-nine large financial institutions that include 19 of the original 29 SIFIs and 10 other firms (5 non-SIFI banks and 5 insurance companies). See Appendix A for a complete list of the firms; Table A1 contains some basic descriptive statistics for the anonymized sample. In all subsequent analysis, the data are numbered in random order within group (i.e., SIFIs, non-SIFI banks, and insurance companies) to protect confidentiality. Additionally, because Kingland is not the only company that collects such data, our SIFI sample is restricted to those firms for which Kingland collects data; hence not all SIFIs are available for our analysis. The other 10 firms in our analysis (i.e., the 5 non-SIFI banks and 5 insurance companies) were chosen at random by Kingland based on some rudimentary guidelines (i.e., that the non-SIFI banks be from comparable SIC codes, that there be an equivalent number from the insurance sector

---

[10] Kingland Systems is one of the leading companies that collects entity data, and specifically legal entity identification (see http://www.kingland.com/ for more information). We are grateful to Kingland, and especially to Tony Brownlee, George Suskalo, and Kyle Wiebers for their generosity in providing the data and their patience in answering our various questions.

for comparison and that the two sets be roughly comparable to each other in asset size) and provided to us for analysis. We emphasize that these data are intended to be illustrative of the potential use of our measures and that conclusive evidence will be left for future research on a more representative/complete sample.

For each firm, we obtain underlying data that encodes the control hierarchy. In this context "control" is defined by Kingland Systems as the parent controlling at least 50%+1 of the voting share for the child subsidiary.[11] As described above, this is the intra-institutional system of all parent-child majority-control relationships that stem from the "ultimate parent" (the firm of interest) through the ongoing process of creation, acquisition, and dissolution of subsidiaries by various entities in the institution. We do not know the names of the subsidiaries in the control hierarchies, only the parent-child linkages (as given by anonymized Kingland ID). In addition to the control relationships the dataset also contains the country of origin and four-digit SIC (Standard Industrial Classification) code of each subsidiary; we refer to these characteristics as 'labels' below.[12] We have data for the twenty-nine institutions at two distinct dates, May 26, 2011 and February 25, 2013, spanning either side of the change (due to the Dodd-Frank Act) in the reporting threshold for consolidated entities.

3. Methods

---

[11] We note that in actuality there are numerous definitions of control (e.g., it may depend upon the nature of ownership in terms of the kinds of interest – voting or non-voting), see Carmassi and Herring (2014) for a detailed discussion of alternative sources of control hierarchy data and the limitations inherent in each. While our data has the advantage of being comprehensive in that it is also used by the firms for business and regulatory purposes, its proprietary nature is an obvious limitation. As the goal of our paper is to highlight the way in which metrics can be constructed from network methods to inform supervisory assessments, in particular the assessments of the Federal Reserve tasked with a responsibility as consolidated supervisor, we do not explore alternative definitions.

[12] https://www.osha.gov/pls/imis/sic_manual.html

We draw on techniques from the science of *networks* to analyze the organizational structure of these large market participants (see Newman 2010 for a basic reference). Network analysis has already proven important for its ability to articulate complex interrelationships in financial networks, in particular, Roukney et al. (2018) highlight the importance of the network structure itself (i.e., the topology of the network) in the context of interconnected networks. Our research is related to their idea but we instead consider intrafirm networks; to date there has been little research on this aspect of financial stability.[13] Subsequent to the writing of this paper, Flood et al., 2017 similarly emphasize the importance of the topology of intrafirm networks for financial stability, particularly in times of crisis when rapid decision-making is critical (e.g., in assessing ease of resolution of a failing firm). The relevance of intrafirm networks for financial stability also stems from the microprudential approach to the determination of systemic importance.

3.1. Definitions and Assumptions

Before stating our assumptions, we first need to define some standard network terms. The networks describing the majority-control structures of the firms in our study are characterized by a rooted directed tree structure (see Figure 1). A *tree* is a network without loops. This type of network is composed of *nodes* and *edges*. It is *directed* if the edges (i.e., the links between the nodes) come with a direction. Note (as indicated in Figure 1) a rooted tree has a special node, the "root", also called the *ultimate parent*. In the trees of interest

---

[13] In the *Journal of Financial Stability* special issue on *Network Models, Stress Testing and Other Tools for Financial Stability Monitoring and Macroprudential Policy Design and Implementation,* (April 2018), for example, none of the papers considered intrafirm network structure, suggesting that despite growing literature on the complexity of interconnected financial networks, there is still much work to be done in the area of intrafirm complexity.

(representing majority-control hierarchies) all edges point away from the root. An edge pointing from node A (the "parent") to node B (a "child") encodes the fact that entity A has majority control over entity B (i.e., entity B is a subsidiary of entity A). Nodes that have no directed edges to other nodes (i.e. they do not have majority-control over any entities) are called *leaves.* The maximum number of nodes that a path from the root to any leaf would pass through is referred to as the *depth* of the tree. The number of children for a given node is called the *degree* of that node.[14]
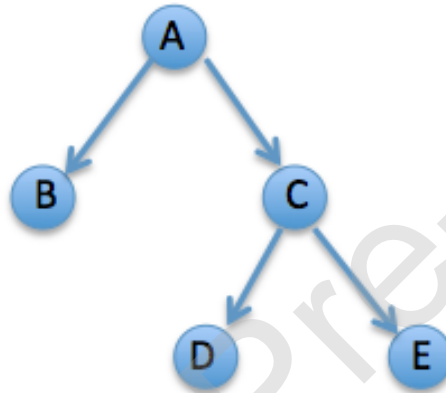

Figure 1. Basic network terminology and structures for *rooted directed trees*. This figure shows a particular example of a rooted directed tree. The root is node A, while B, D, and E are leaves or leaf nodes. Node C is neither the root nor a leaf and is sometimes called an *internal* node. Nodes B and C are children of the node A, which is the parent of these nodes. In addition, this is a regularly branching tree in which each node that has children has exactly two children. This tree has depth 2 (the distance of from the root node A to the farthest children down the tree, in this case either node D or node E) and a total of five nodes. If this were the tree corresponding to a control structure of a financial institution, then the "ultimate parent" would be node A and nodes B and C would be direct subsidiaries of A in which A still held a controlling interest, while D and E would denote subsidiaries of C in which C held a controlling interest.

We refer to the layout of nodes and edges corresponding to the majority control-hierarchy as the *topology* of the network. We make the following assumptions about this topology:

A1. Information: Information flows more easily between nodes with the same label (i.e., the same country or industry). Therefore the number (proportion) of child nodes whose

---

[14] In more general directed networks this is often referred to as the "out-degree" of the node, to distinguish from the "in-degree", that is, the number of immediate parents of a given node. However, in a rooted tree, all nodes except for the root have in-degree one, so throughout this paper we will use "degree" to refer to out-degree.

labels differ from their immediate parent is a proxy for the amount of coordination required to exchange information between those nodes and their immediate parent. The lower the proportion, the greater the ease of accessing needed information.

A2. Firm-level decision-making: Coordination is critical to both short-term and long-term decision-making regarding a firm. It follows from A1 that coordination is easier within the same geographic jurisdiction or industry. In addition, due to consolidated reporting requirements imposed by, among others, the Federal Reserve, supervisory oversight of a multinational firm occurs globally but similarly relies on cross-jurisdictional and cross-industry coordination.

A3. Supervisory evaluation: Coordination is also critical to the supervisory process, particularly for the Federal Reserve as articulated in the Bank Holding Company Supervisory Manual (Board of Governors of the Federal Reserve System, 2017).

In the context of a (banking) supervisor, we additionally assume:

a. Risks outside of one's supervisory domain can only be observed indirectly (i.e., through communication and coordination with other regulators). Thus, the risk to financial stability posed by a systemically important firm is related to the ability of its (country or industry) supervisors to coordinate with each other to monitor the firm.

b. A supervisor has access to all information within its own country/industry.[15] In addition, supervisors exert sufficient oversight and control within their domain. In other words, supervisory lapses occur unintentionally as a result of lack of coordination /

---

[15] In some countries, for example the US, there are multiple regulators, so that even within-country coordination may be challenging. For the purposes of our discussion, we ignore this additional layer of complexity at the country level; it is reflected in the discussions of SIC complexity.

communication / information and not due to incompetence or negligence or for any other intentional or malicious reasons.

c. A **consolidated supervisor** (e.g., the Federal Reserve, see Appendix B) is additionally responsible for aggregating information across a number of supervisors of different domains. The consolidated supervisor's ability to do this effectively depends on its ability to obtain this information which in turn is affected by the organizational structure of the firm for which it is responsible, due to the amount of coordination required. A more complex organizational structure renders the task more challenging because it hinders both: (a) communication, including the communication of risk, and (b) the coordination that is necessary for risk mitigation and management, as well as for resolution efforts (Cetorelli and Goldberg 2014, Alessandri et al. 2015, Flood et al. 2017), at both the firm and supervisory level.

The above three assumptions follow from the Bank Holding Company Supervision Manual's (Board of Governors of the Federal Reserve System 2017) emphasis on the importance of Federal Reserve coordination with, and reliance on information from, other primary and functional regulators for the purposes of carrying out its role as consolidated supervisor (see Appendix B for relevant passages).

An implication of Assumptions A1-A3 is that it is not just the *number* of children emanating from a node that is important for consolidated supervision. Rather (Assumption A1) it is the ease by which a consolidated supervisor can access the information necessary in order to make his or her assessment. A subsidiary with children that all fall under the same supervisor or functional regulator is easier to assess than one with children that each

fall under different country or SIC classifications because no coordination with other supervisors is required in the first case while the second requires coordination across a number of different supervisors.[16] The assumptions also imply that less breadth of expertise is required to supervise a subgroup that is totally within one's domain.

The next three assumptions link financial stability to the coordination challenges associated with a firm's organizational structure:

A4. Risk Implications.   In the absence of coordination, opacity ensues, threatening financial stability by increasing the probability that a risk will either occur or go undetected for longer.

A5. Transmission of risk within the firm: Spillover of risk (i.e., from one country to another or one industry to another) within a firm is more likely to occur across heterogeneous linkages (i.e., edges that connect nodes with different labels) than across homogeneous linkages (i.e., edges that connect nodes with the same label) due to the above-mentioned coordination/information assumptions.

A6. Transmission of risk outside the firm: Because complexity is important for the assessment of the risk of an individual organization, it in turn is important for mitigating systemic risk and thus, financial stability.   This assumption follows from the Carmassi and Herring (2014) premise in the Introduction.[17]

---

[16] Cetorelli and Goldberg (2014, footnote 11) make a similar point in the context of the challenges a firm's management faces when a parent and its affiliate are located in different countries. They also mention the monitoring difficulty that organizational complexity might pose in their footnote 14.

[17] This assumption is not really necessary for our analysis but is important to the idea that our metrics may be useful for SIFI designation.   We acknowledge that there is a lack of consensus in the literature regarding whether the failure of a single firm necessarily implies a systemic cascade and emphasize that we are agnostic on this point but mention the Carmassi and Herring (2014) assumption to motivate our focus on a subset of SIFI firms and to highlight the fact that we are not alone in drawing the link to financial stability.

Regulatory considerations create constraints on the arrangement of the nodes and edges in the rooted trees derived from the control hierarchy. For example, the Banking Act of 1933 (more commonly known as the Glass-Steagall Act, after its Senator and Congressman sponsors) required a separation between commercial and investment banks (as well as other restrictions). In a network theoretic framework, this would mean that under Glass-Steagall, commercial and investment banks would need to be in different subtrees in the control hierarchy (e.g., in Figure 1, investment banks would need to be on one side of the root and commercial banks on the other). Similarly, legal and tax incentives might drive patterns of country-incorporation, resulting in a tree structure where nodes associated with a specific business classification are also associated with a specific country. The 1999 passage of the Gramm-Leach-Bliley Act repealed many of the Glass-Steagall restrictions, fostering substantial growth-by-acquisition in the banking sector as banks diversified into new industries and countries (DeYoung et al. 2009). In tree terminology, the repeal would suggest that banking trees are less characterized by country- or SIC-specific subtrees but instead have become more jumbled. As the recent financial crisis unfolded, many viewed the repeal of Glass-Steagall as partially responsible, and calls to reenact it intensified, resulting in the 2010 Dodd-Frank Act. As a result, we might expect to see banking trees moving back toward their pre-1999 subtree separation (moving closer to a "perfect tree" in the terminology of the next section). In addition, the 2012 lowering of the threshold requiring subsidiaries to report on a consolidated basis (as stipulated in the 2010 Dodd-Frank Act) suggests we should see an increase in the number of nodes included in the majority-control hierarchies between 2011 and 2013. Therefore, the patterns of SIC and

15

country codes (as node labels) in the majority-control hierarchy could help to reify these regulatory changes and should be of interest to banking supervisors.

3.2. Statistical Description of Tree Heterogeneity

This section of the paper delineates some metrics to quantify several important characteristics of the institutions' majority-control hierarchy trees that relate to the ease of coordination that are highlighted in Assumptions A1-A6. Despite a high degree of heterogeneity in the firms' structures, these metrics provide a basis for comparison.

A fundamental quantitative descriptor for any network (tree or not) is its *degree distribution*, describing the probability distribution associated with the network's set of degrees (i.e., the function $d(i)$ that records the fraction of nodes with i children). Just as the degree distribution of a tree describing a firm's reporting lines might be used to characterize the spans of control (Urwick 1956) of its management, the degree distribution of a firm's majority-control hierarchy analogously might be used to describe a consolidated supervisor's span of control in assessing the firm's systemic importance.

For this reason, we therefore introduce a new metric of complexity that we see as related to supervisory challenges derived from the need to coordinate oversight efforts across a variety of jurisdictions and agencies. It follows from the assumptions that the simplest supervisory structure a firm can have is a "perfect" supervisory tree, where all nodes of a certain label are clustered (in the language of Flood et al. 2017, this equates to the number of maximally homogeneous subgroups – subtrees that consist of nodes of a single label – being equal to the number of labels).

16

Assumption A3 forms the baseline for our analysis. In assessing the complexity of any given firm, we derive a measure reflecting the distance of the majority-control hierarchy tree from this "ideal" supervisory or *perfect tree* structure.[18]

We illustrate how this can be done using our Kingland data sample. For each firm, we consider a "perfect" supervisory baseline to be one with topology equal to the firm in question, where beginning with the nodes at level 2 each child (subsidiary) has the same label (either country or 2-digit SIC classification) as its immediate parent, -- i.e., we take as given the firm's heterogeneity at level 1, reflecting its decisions regarding the distribution of business or geographic lines to each of the immediate subsidiaries (children) of the ultimate parent. By imposing that all edges beyond level 1 join nodes with similar labels, this process will produce a tree with (country or SIC) distributions that may be quite different from the firm's actual majority-control hierarchy. In our framework, a perfect tree is comprised of *perfect groupings*, that is, country-specific (or SIC-specific) groupings that do not involve other countries (SIC codes). Of course in reality none of the firms we consider will have an exact perfect tree structure (i.e., one where there are no edges connecting two nodes with different labels). To this end we introduce the notion of *perfect tree similarity*.

For each firm, we compare the proximity of their actual control hierarchy to their perfect supervisory baseline via a *perfect tree statistic,* in each case computing the fraction of nodes with the same label (i.e., country or 2-digit SIC code) as their immediate parent. The statistic is therefore bounded between zero and one. Mathematically, let *E* denote the

---

[18] That is, we begin by defining a "perfect tree" as one in which all nodes belong to the same (country or 2-digit SIC) classification. In other contexts, the phrase "perfect tree" refers to a tree with the same number of directed edges emanating from each node. It is important to recognize that our definition differs from that one.

number of firms in our sample and N_i the total number of subsidiaries in firm i, i.e., the count of all nodes that are in firm i's majority-control structure). We index all nodes in the firm with $\{0, 1,...,N_i\}$, where we always assign the root node index 0. Also let $s_{i,j} \in \{1,...,S\}$ be the label (i.e., the country or 2-digit SIC code) associated with subsidiary j of firm i, and $pa(j)$ the parent node of that subsidiary j. Finally let $\delta(\cdot) = 1$ if the expression within the parentheses is true and zero otherwise. Then the perfect tree statistic for firm i is given by:

$$S_i = \frac{1}{N_i} \sum_{j=1}^{N_i} \delta\big(s_{i,j} = s_{i,pa(j)}\big) \tag{1}$$

As noted above, this is just the fraction of non-root nodes in firm i whose parents have the same label as the node itself.

Note that in a perfect tree (i.e., a tree in which the perfect tree statistic is equal to one), removal of a node and all subtrees below it will not change the value of the statistic.[19] In contrast, a value of zero means that the subsidiaries below depth 1 are always different in character from their immediate parent (with respect to a given characteristic – country or SIC code) and therefore would require coordination among supervisors across all countries (industries) in which the firm operates. Thus, to the extent that a firm's tree structure is closer to a perfect supervisory tree, we reason that it is easier for both the firm and the supervisor to track/monitor the activities of its subsidiaries or to achieve resolution should the firm fail. In practice we recognize that the perfect tree characterization is too rigid and the costs of a fully-segmented structure may far outweigh the benefits of globalization and cross-border banking, but we believe it may inform consolidated supervision by providing a regulatory ideal against which real world instances can be compared. We refer to the

---

[19] The supervisory analogue to this network structure might arise when a troubled firm is forced to sell or close one of its subsidiaries. The closer the organizational structure is to a perfect tree, the less likely there will be disruption to the rest of the firm when the subsidiary is pared.

comparison between the perfect supervisory baseline and the firm's actual majority-control hierarchy via their associated perfect tree statistics as the *perfect tree similarity statistic.*

The closer a firm is to its own perfect supervisory tree, the easier it will be for the consolidated supervisor to satisfy the mandate of the Bank Holding Company Act (to aggregate information from supervisors in different jurisdictions or industries evaluating the portion of the firm for which they are responsible) in order to form an assessment of the consolidated organization and without need for additional supervisory review. Firms can be evaluated according to their own internal structure and the level of complexity they exhibit as a result of that structure can be compared, thus informing the allocation of supervisory resources.

3.3. Comparing complexity across heterogeneous firms

Assumptions A1-A6 lead us to the following:

"By comparing the firm's actual organizational structure to a hypothetical perfect tree structure, we can draw inferences about the challenges a firm and/or its supervisors would experience were one of its subsidiaries to experience deterioration."

It additionally follows from assumption A5 that the greater number of perfect groupings in a tree, the less likely that difficulty faced by a subsidiary in one country (industry) will spill over into subsidiaries in other countries (and the easier it will be to 'ring-fence' the risk in a resolution situation).[20]

---

[20] A resolution framework where each national authority has responsibility for the resolution of banks that fall under its jurisdiction is referred to as the "subsidiarization" model in of Carmassi and Herring (2014); in their context, our country and industry perfect tree statistics could inform the geographic subsidiarization and subsidiarization across lines of business that they propose.

The closer a structure is to a perfect tree, the easier it is for a supervisor to evaluate its risk. This is because the supervisor has direct access to information from a higher proportion of subsidiaries in the firm. Similarly, the farther away from a perfect tree, the more coordination is required for information to transmit to the consolidated supervisor. We expect firms to have moved closer to a perfect tree structure between 2011 and 2013.

As noted above, Assumptions A1-A6 imply that organizational structures that are closer to a "perfect tree" will be easier for a supervisor to evaluate. However, simply calculating the proportion of edges in the perfect tree that connects nodes of differing labels will not provide a satisfactory metric because different firms have different business models (and hence the distribution of their country and SIC classifications will differ) and thus some firms more than others will have structures and distributions that lend themselves naturally to be closer to a perfect tree. Such a comparison would give rise to the "apples-to-oranges" and aforementioned one-size-fits-all criticisms that banks often level at regulations, a reference to the idea that by their nature, regulations need to be general and hence don't adequately account for the fact that firms have different business models. Recognizing these criticisms, we therefore need to find a way to determine how unusual a specific firm's organizational structure is, by comparing it to all the other structures that might arise, given the specific distribution of that firm's characteristics. Specifically, we take a firm's tree topology as fixed and bootstrap different labelings of the nodes (country or 2-digit SIC code) according to the firm-specific empirical label transition probabilities. Concretely, for firm i, we first set the label of the root node, $s_{i,0}$, to its observed (actual) value. We then traverse the tree following the edges through each node, sampling the label of each node according to the empirical conditional distribution:

$$p(s_{i,j} = s \,|s_{i,pa(j)} = t) \propto \sum_{j=1}^{N_i} \delta(s_{i,j} = s)\delta(s_{i,pa(j)} = t). \qquad (2)$$

We continue this process recursively for the child nodes of the root node and their children until we reach the leaf nodes. When j is a non-leaf node (i.e., a node with children), we restrict s to be from the subset of labels associated with the non-leaf nodes in the observed tree; similarly, when j is a leaf node (i.e., a node with no children), we restrict s to be from the subset of labels associated with the leaf nodes in the observed tree.[21] When drawing labels for the leaf nodes we add a small correction factor to the empirical probability, $1/|N|$, where $|N|$ is the number of nodes in the tree, to account for the fact that conditional on the simulated parent, drawing from the leaf-only empirical distribution might result in labels that have a 0% conditional probability. This correction factor admits the possibility that another label (i.e., different from the observed one) might be assigned to the leaf; hence a greater variety of tree labelings can be explored.

The number of bootstrap replications that we use for each firm varies in order to achieve sufficiently large coverage of the range of possible label combinations. Specifically, for each firm we choose the number of bootstrap replicates according to the following formula: # replicates = (# of unique label-label transitions) x (maximum depth of the firm-specific tree) x (average # of children of the non-leaf nodes in the firm-specific tree), where the average # of children is the mean number of children emanating from all non-leaf nodes.[22] For each replication, the *perfect tree similarity statistic* is computed.

---

[21] This is because in our trees there are many labels that do not appear as leaves and/or never emanate from certain parent labels and others that only appear as leaves and always emanate from the same parent label.

[22] Leaf nodes by definition have zero children, hence including them in the computation of the average would result in a strictly (and sometimes much) lower average. Since we are primarily interested in a larger (rather than smaller) number of replications, we exclude these from the computation of the average. We acknowledge that this formula is somewhat arbitrary. In an earlier version of the paper, the number of replications was fixed

21

The replications therefore allow us to generate a distribution of possible structures for each firm, conditional on its overall topology (organizational structure), and to compute a range of summary statistics from these distributions. Our approach allows a supervisor to compare complexity across firms that have different country and industry profiles, while still holding the tree topology and distribution of countries/industries fixed for each firm. In particular, from the replications we can compute a percentile ranking for each firm, representing the proportion of simulations that produced a perfect tree similarity statistic that was lower (higher) than the one given by the firm's observed structure.

3.4. Two Examples

As explained in Section 2, the majority-control hierarchy imposes a tree structure on the data, driven by consolidated supervision considerations. That is, for the purposes of consolidation, as we navigate through a chain of subsidiaries, we should never loop back to any of the intermediate (or parent) entities. For illustrative purposes, it is useful to select a firm with a small number of nodes to demonstrate the information in the tree layouts further. [23] Figure 2 considers SIFI S11; in 2011, this firm had 43 nodes, corresponding to 14 countries, four 1-digit SIC codes, and seven 2-digit SIC codes, with a tree depth of four. In figures such as this, the largest circle represents the ultimate parent, with the size of other circles decreasing with growing distance from this parent: the smaller the circle, the farther down the tree it is. Figure 2a shows the layout of S11 by depth. Note that S11 has most of

at 1,000. Using a formula to determine the number of replicates recognizes the fact that the number of possible tree combinations for a firm with a small number of nodes (e.g., S11) is much lower than the number for a firm with a large number of nodes (e.g., S19). This formula was chosen to strike a balance between having a large number of replicates for even the smallest firms and the computational burden that would result from simulating a proportionate number of replicates for the largest firms.

[23] All layouts were done using the freely available Gephi software package, available at https://gephi.org/ The layouts are normalized to be consistent across figures – that is, we have created layouts in which nodes are in the same positions from figure to figure.

its subsidiaries at depth 3, with 28 entities distributed among four depth 2 subsidiaries. In

addition, all but one of the children of the root (ultimate parent) is a leaf (does not control

any additional subsidiaries). Thus, most of the control hierarchy in this

tree emanates from one child of the ultimate parent; in the absence of that one node, the tree

would have depth 1. It should be evident, therefore, that severing the link between
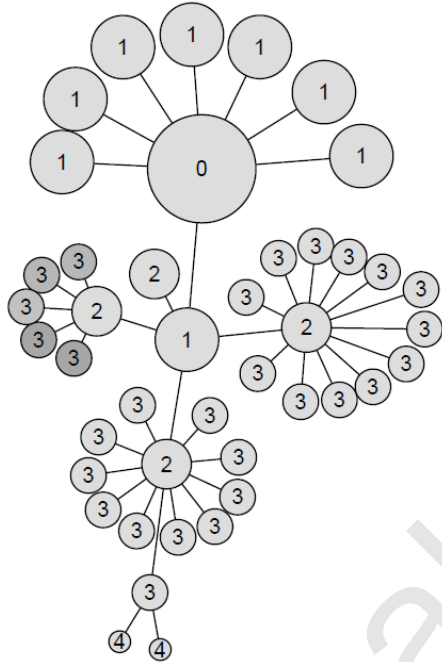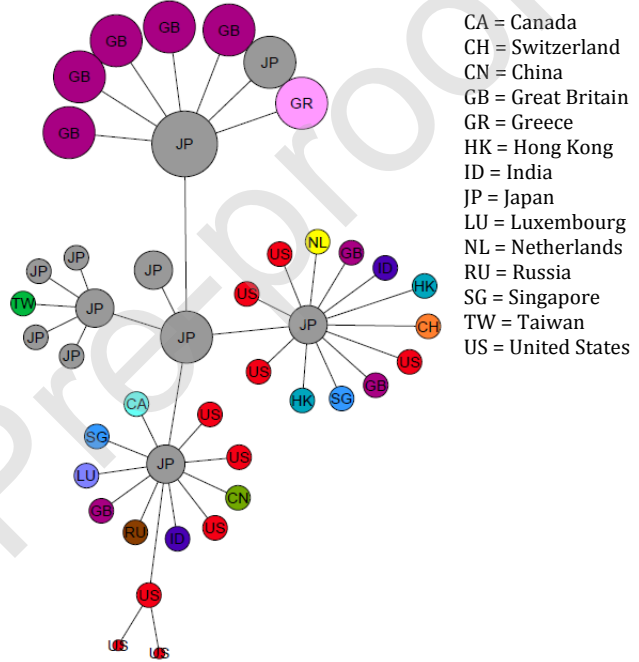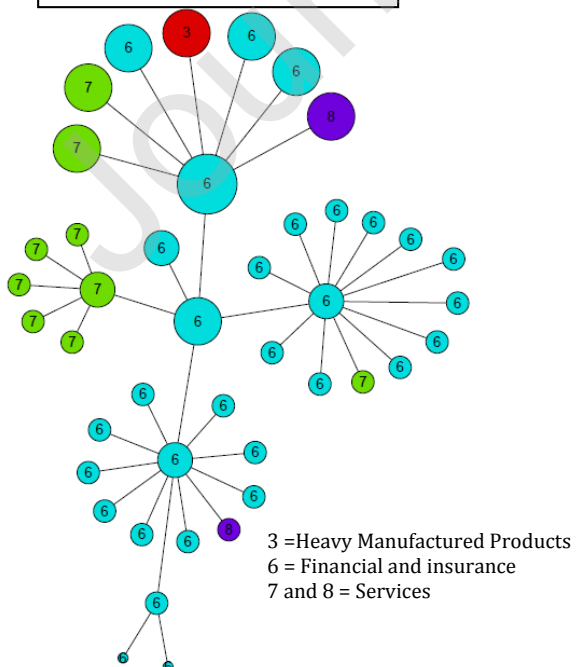
Figure 2a: by depth (n=4)

Figure 2b: by country (n=14)

CA = Canada
CH = Switzerland
CN = China
GB = Great Britain
GR = Greece
HK = Hong Kong
ID = India
JP = Japan
LU = Luxembourg
NL = Netherlands
RU = Russia
SG = Singapore
TW = Taiwan
US = United States

Figure 2c: by 1-digit SIC classification (n=4)

Figure 2d: by 2-digit SIC classification (n=7)

3 = Heavy Manufactured Products
6 = Financial and insurance
7 and 8 = Services
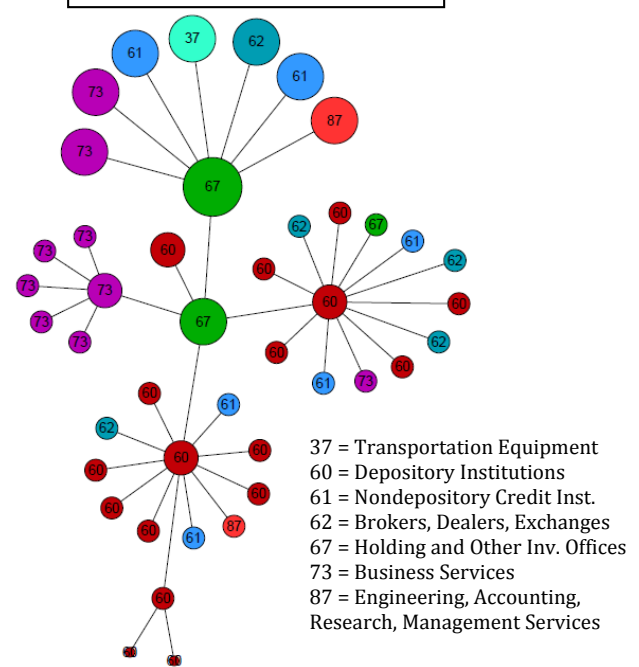
37 = Transportation Equipment
60 = Depository Institutions
61 = Nondepository Credit Inst.
62 = Brokers, Dealers, Exchanges
67 = Holding and Other Inv. Offices
73 = Business Services
87 = Engineering, Accounting,
Research, Management Services

Figures 2a-d. The control hierarchy of SIFI S11, color-coded by distance from the ultimate

the root and that particular node would dramatically change the tree configuration, whereas severing the link with any other nodes at depth one would hardly change the configuration at all. Put another way, that node is an essential component of the complexity of the firm's majority-control hierarchy. Hence it may be a node of particular focus for a supervisor charged with evaluating the complexity of the SIFI.

Figure 2b considers S11's control hierarchy when labeled by country. Despite the parent company being incorporated in Japan, most (five of eight) of its immediate children are incorporated in Great Britain, with only two incorporated in Japan and one in Greece. Yet among the 35 other children in the tree, all except the two US subsidiaries at depth 4 have an immediate parent that is also incorporated in Japan, suggesting that it will be relatively easy for a Japanese supervisor to obtain information at all levels of control of the firm since there are only two entities beyond the immediate reach of the supervisor.[24]

Figure 2c illustrates S11's control hierarchy labeled by 1-digit SIC code. At the 1-digit level, this firm is fairly homogeneous, with 31 of the children of the ultimate parent operating in the same industry. In addition to financial services, this firm has control over one subsidiary in SIC area 3 (roughly construction and equipment), and 11 in services (areas 7 and 8); for the most part the services are concentrated in one subtree (to the left of the diagram). With most of the tree falling into the same SIC classification, it is evident that a financial services supervisor would be able to assess most of the firm's activities without having to rely on coordination with supervisors from other industries.

---

[24] The tree structure with most subsidiaries being controlled by a Japanese subsidiary may reflect a keiretsu arrangement (Berglof and Perotti, 1994). We thank a referee for bringing this to our attention. Cetorelli and Goldberg (2014) similarly note that using their metrics and definition of complexity, Japanese banking organizations appear to be less complex.

In addition, since the subsidiaries that fall under SIC classification 7 are concentrated in one subtree, the diagram highlights a single link that might warrant additional scrutiny or that might be severed should either the firm or its supervisor wish to reduce the range of the firm's business activities. An additional implication is that should this link (between the subtree under SIC classification 7 and its parent node with SIC classification 6) be severed, either due to deacquisition or the subtree falling below the majority-control threshold, there would be little change to the burden of the consolidated supervisor. Figure 2d analogously illustrates S11's control hierarchy labeled by 2-digit SIC code.

For comparison, consider Figure 3 in which we see the same kinds of snapshots (at the same date) for a much larger firm (S16): this firm has 1778 nodes, corresponding to 32 countries and 100 4-digit SIC codes, with a tree depth of 5. We provide three representations color-coded according to distance from the root, country of origin, and 1-digit SIC code, respectively. Figure 3a shows that the majority of nodes are at depth 2. Specifically, there are 299 nodes at depth 1, 1186 nodes at depth 2, 188 nodes at depth 3, 24 nodes at depth 4 and 80 nodes at depth 5. Figure 3b illustrates that in addition to the parent company being located in the US (the largest circle in the center), the majority of S16's subsidiaries are also located in the US. In addition, some of those US subsidiaries themselves have quite elaborate majority-control hierarchies, judging from the large clusters at the top of the figure, as well as the ones on the far left, far right and bottom of the figure, that are almost exclusively comprised of US subsidiaries. There are also a large number of subsidiaries located in the UK (green), Germany (yellow), and Spain (light blue), with the UK and Spanish subsidiaries having a number of children that are located in the same country as their immediate parent. Figure 3c shows that while many

25

Figure 3a: by depth (n=5)

Figure 3b: by country (n=32)

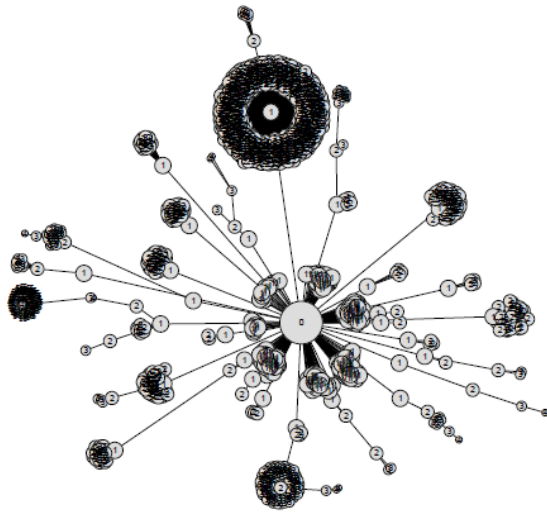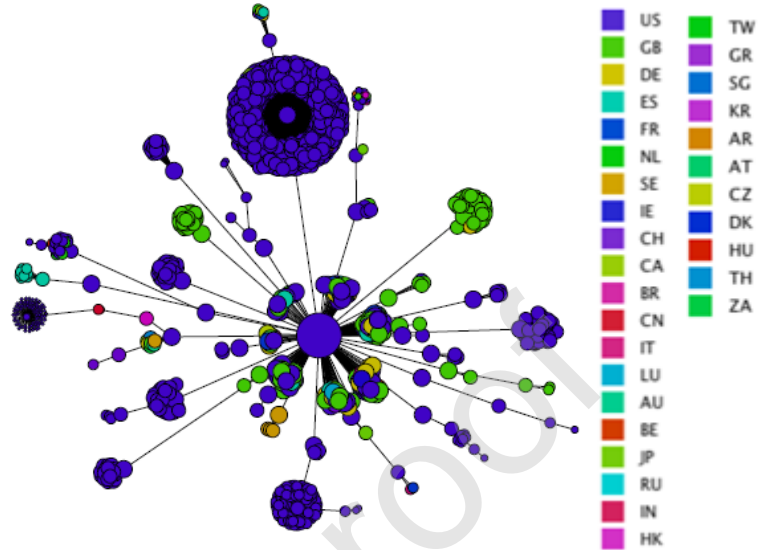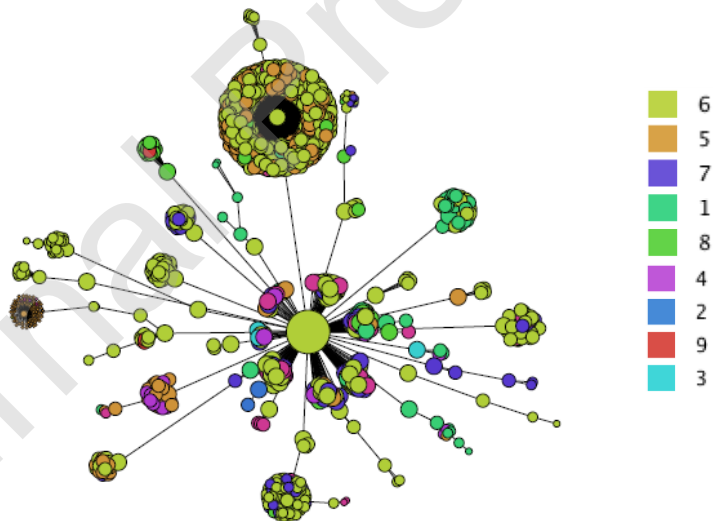Figure 3c: by 1-digit SIC classification (n=9)

Figures 3a-c. The control hierarchy of SIFI S16, color-coded by distance from the ultimate parent, by depth, country, and 1-digit SIC classification. A consistent layout is used for all three representations, for comparability. Node size is proportional to distance to the ultimate parent, with larger nodes corresponding to closer distance.

26

of S16's subsidiaries are in the same 1-digit SIC classification as the parent company (classification 6 – Finance, Insurance, and Real Estate), the company also has a fairly diversified range of subsidiaries with the second largest 1-digit classification in the wholesale and retail trade sectors (SIC code 5).

Taken together, Figures 2 and 3 illustrate that this view of complexity is multidimensional, that is, a firm that has a complex SIC classification structure may not have a complex country structure and vice versa.[25] Identification of these complexity dimensions provides a way to evaluate and compare sometimes very disparate firms and may inform allocation of supervisory resources.   In addition, network information can be used to develop metrics with which to compare complexity across firms.

4. Results

In this section we present several results describing the tree topology of the firms in our sample; discussion of the implications of these results for our three main questions of interest is reserved for the next section.   The main results are given in the following subsections, corresponding to different metrics:

1) Degree statistics

2) Markov statistics (parent-child transitions)

3) Perfect tree similarity statistics

We then use these metrics to illustrate the manner in which supervisors might draw inferences about the firms and explore whether there are differences between the coordination efforts required for different types of institutions.   Finally, we consider

---

[25] The observation that the risks firms face may be multidimensional forms the basis for much of the work on multilayer networks. See, for example, Battiston et al. (2016a,b) and Berndsen et al. (2018)

whether our metrics provide additional information beyond the size delineation that has traditionally been used to classify institutions that warrant greater regulatory oversight due to their systemic importance. Note that while in this paper we focus our discussion exclusively on the overall tree topology of the firms in our sample, a similar analysis can be readily conducted on any subtrees of interest, for example to inform ring-fencing decisions that supervisors might need to make once a risk has been identified.

4.1. Overall degree statistics.

One way to characterize the organizational structure of the firms is via the hierarchy distribution, that is, the proportion of nodes at each level of the tree hierarchy. To aid our understanding of how tree hierarchies might be used in the context of large financial institutions, it is useful to first consider how different organizational structures correspond to different hierarchy distributions. For example, an institution with a very flat (i.e., "entrepreneurial") management structure would have a large proportion of nodes at level one and relatively few branches extending from those nodes. In contrast, an institution that concentrates its decision-making among only a few senior managers who are then held accountable for large portions of the firm would have a larger proportion of nodes at lower levels of the tree. Such a diffuse tree might also be found among organizations that have experienced significant growth by acquisition, such as many financial institutions in the decade preceding the 2008 financial crisis, where the tree of an acquired complex organization may have been grafted to the tree of the acquiring parent somewhere below the highest level, creating a very hierarchical structure of great

depth (a "bureaucratic" structure). Firms also might be arranged along geographical ("divisional") or industry ("functional") lines (Armour and Teece 1978).[26]

The hierarchy distribution for our sample of firms is summarized in Figure 4, for 2011 and 2013 separately. Note that there is substantial variation across the firms. For example, while in 2011 more than one-third of the firms have more than half of their nodes at the first level of the tree hierarchy, others have relatively few nodes branching from the ultimate parent and instead have a large concentration of nodes farther down the tree (e.g., S13 and B1). None of the firms with the deepest trees (i.e., more than six levels) have node concentration at the first level of the tree hierarchy, indicating a flatter or more diffuse organizational structure. Note firm I4 has a tree structure that spreads out at each level in 2011. Across all firms in our sample, roughly one-third of the nodes are in each of the first two levels, another 22% in the third level, and only 10% at deeper levels in the tree hierarchy.

In contrast, by 2013, the tree hierarchies of the firms in our sample deepened substantially. For example, only half as many (five) firms now have more than half their nodes at the first level, while 14 have less than 10% of their nodes at the first level. In addition, 11 firms now have more than seven levels while just two years earlier, none did. Across all firms in the sample, by 2013 roughly 25% of the nodes were at deeper than the third level. Thus, from the perspective of consolidated supervision, the challenges associated with assessing these firms increased dramatically, with many entities in the organization being much farther removed from the ultimate (root) parent.

---

[26] An earlier version of our paper (Lumsdaine, et al., 2016) discussed the similarity of the firms' degree distributions to power law distributions (distributions that have the form x^{-r} for some $r>1$. Power law distributions suggest certain patterns of growth in the network and also have implications for its robustness and stability.
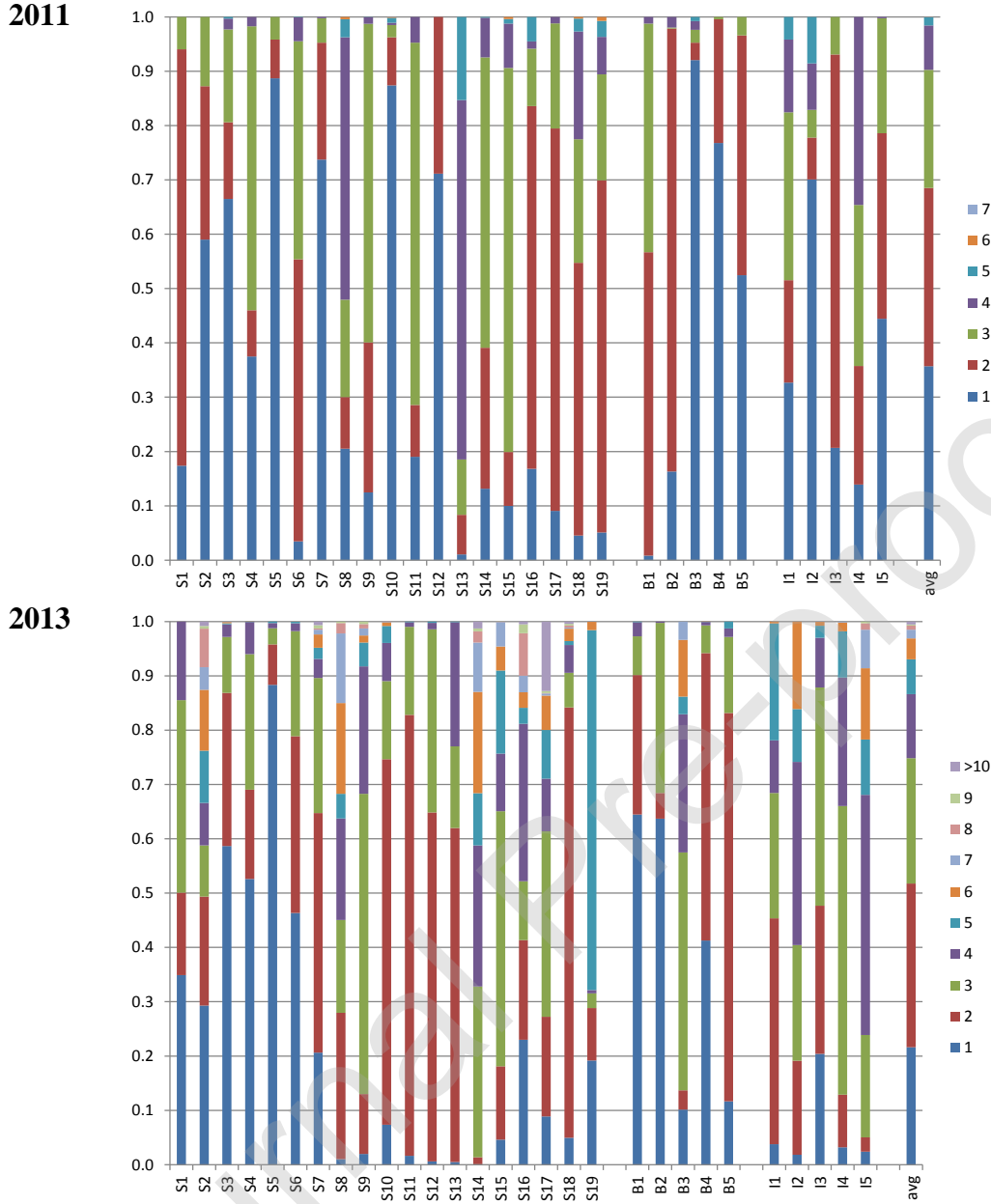
2011

2013

Figure 4. For each firm, we plot the fraction of nodes at each level in its control hierarchy (i.e., the distance from the root), for 2011 and 2013. For ease of comparability between 2011 and 2013, for firms with more than nine levels in 2013, the fraction of nodes beyond level 9 are aggregated into the ">10" distance. The fraction of nodes at each level for the entire sample of firms we consider is shown in the right-most column ("avg").

4.2 Aggregation Across a Sample of Firms: Markov Statistics (parent-child similarities)

As noted above in the discussion of both degree and hierarchy distributions, from

a supervisory perspective it is not just the *number* (or proportion) of child nodes that

30

emanate from a parent node that matters but also the *similarity* between the parent and child.[27] In this section we therefore ask the following: Given that a node has a particular label *s*, what is the conditional ('birth') probability *P(s/s)* that the node below it also has label s, not just for a given firm but across a sample of firms? Birth probabilities are one simple measure of a type of "homophily" in a network, or the predilection for a node (in this case a parent) to be connected to another node of the same kind.[28]   In the context of the rooted directed tree networks we consider, the Markov similarity statistic *P(s/s)* for label *s* is computed as:

$$P(s|s) = \left(\sum_{i=1}^{E} \sum_{j=1}^{N_i} \delta\big(s_{i,pa(j)} = s\big) \right)^{-1} \sum_{i=1}^{E} \sum_{j=1}^{N_i} \delta\big(s_{i,j} = s\big) \delta\big(s_{i,pa(j)} = s\big). \qquad (3)$$

Note the resemblance between the second sum and the formula for the perfect tree similarity statistic. In the case of the Markov similarity statistic, however, the sum is computed per unique value of the label *s* over all firms in the sample. A higher Markov similarity statistic indicates greater within-label (country or industry) linkage (as opposed to across-label linkage) and hence may suggest a greater likelihood of spillover to other firms with the same label. From a financial stability perspective, if we were able to observe the entire network of firms, such a metric provides a measure of the degree of coordination burden involved in a particular dimension (e.g., with respect to that country) across the entire sample and hence may be useful in identifying potential contagion effects should a firm with a particular label (e.g., a specific country) start to exhibit signs

---

[27] Although for expositional purposes much of our discussion has focused on the ease of a supervisor to "look below" in examining entities that are lower down the tree, our focus on the similarity between parent and child nodes stems from the view that consolidated reporting and risk management will be easier when, for example, a child has the same legal, accounting, tax, or supervisory framework as its parent.
[28] Flood et al. (2017) describe edges connecting nodes of the same kind as "homogeneous edges" in their paper linking intrafirm network structure to the ease of supervisory coordination.

of stress. Such a statistic may help alert supervisors of high Markov-similarity countries or industries to the possibility of *a systemic event in that country or industry.* In contrast, a low Markov similarity statistic may indicate greater cross-label linkage and hence may suggest a greater likelihood of cross-country or cross-industry contagion, pointing to the possibility of a global disruption to financial stability. Discussion of the role of homophily in interconnected networks and its implications for financial stability and systemic risk are contained in Berndsen et al. (2018) and references therein.

To illustrate, in Table 1 we compute the Markov similarity statistics for each country, using all firms in our sample, for both 2011 and 2013. The probabilities vary dramatically for different countries. In 2011, Canada has the highest probability, with P(Canada | Canada) = 0.97; in contrast Switzerland has the lowest, with P(Switzerland | Switzerland) = 0.11. Part of the reason for this variation is differences in country frequency; that is, even under random assignment a country that has more nodes in the network has a greater likelihood of being paired with its own country than does a country that has fewer nodes in the network. By comparing the probabilities in 2011 to those in 2013, however, we see that for most countries, the Markov similarity statistic increases, with many countries above 0.9. This suggests that across the network of 29 firms there has been a shift to consolidate subsidiaries from a given country under parents from the same country, consistent with our prediction based on Assumptions A1-A6.

4.3 Complexity and Changing Structure – Perfect Tree Similarity Statistics

In this section we document the fluidity of majority-control hierarchies by comparing the perfect tree statistics in 2011 and 2013. As is the case for many complex

Table 1: Within-country "birth" probabilities P(A|A) aggregated over all firms in our sample, 2011 and 2013.

Country "A"	2011		2013	
	P(A|A)	Rank	P(A|A)	Rank
Canada	0.97	1	0.974	9
United States	0.94	2	0.958	11
Brazil	0.93	3	0.999	2
Malaysia	0.92	4	0.777	23
Poland	0.86	5	0.936	14
Russia	0.86	6	0.919	17
Great Britain	0.86	7	0.892	21
Norway	0.83	8	0.917	18
Ireland	0.77	9	0.430	31
Hong Kong	0.74	10	0.626	28
Spain	0.67	11	0.986	6
Jersey	0.67	12	0.200	35
Portugal	0.63	13	0.985	7
Czech Republic	0.62	14	0.902	20
Sweden	0.61	15	0.867	22
China	0.56	16	1.000	1
Austria	0.51	17	0.914	19
Germany	0.50	18	0.688	26
Trinidad & Tobago	0.50	19	0.375	32
France	0.42	20	0.926	15
Belgium	0.35	21	0.968	10
Italy	0.32	22	0.976	8
Singapore	0.29	23	0.748	25
Netherlands	0.28	24	0.648	27
Japan	0.25	25	0.942	12
Luxembourg	0.20	26	0.590	29
South Africa	0.19	27	0.926	16
Denmark	0.13	28	0.750	24
Bermuda	0.11	29	0.209	34
Switzerland	0.11	30	0.456	30
Mexico			0.997	3
India			0.995	4
Kenya			0.345	33
Argentina			0.987	5
Australia			0.940	13

Notes to Table 3. This table summarizes the "in" transition statistics with respect to country labels. That is, for any country A consider all the subsidiaries incorporated in A. Here we list the proportion of the children of such subsidiaries that are also incorporated in A (P(A|A)) for any country where in 2011 the in-country probability is neither zero or one, along with some additional countries from 2013.

systems, the structure of the majority-control hierarchy is a response to a variety of endogenous and exogenous forces. A significant component of the latter comes from regulatory frameworks, which can include pressures that come from tax and corporate legal structures as well as supervisory restrictions on activities. For example, in December 2012, the lowering of the threshold for consolidated supervision as a result of the Dodd-Frank Act of 2010 went into effect.   Because our two data snapshots straddle this event, we can use our sample to illustrate how our metrics might be used to study or evaluate changes in policy.

4.3.1 Country Perfect Tree Similarity Statistics

Table 2 summarizes the country-level perfect tree statistics and bootstrapping simulations for each firm and both dates, grouped by firm type (SIFIs, non-SIFI banks, and insurance companies). For each firm and each date, four statistics are reported: (1) the firm's "actual" perfect tree statistic (described in Section 3.4), (2) the mean of the same statistic computed over all replications of that firm's simulated data, (3) the standard deviation over all replications, (4) the empirical quantile corresponding to where the observed statistic would lie in the empirical distribution generated from the replications (that is, the percentage of replications that resulted in a statistic that was less than the firm's observed statistic).

It is important to note that while perfect tree statistics of different values can be compared *within* a single firm (i.e., a lower value means more subsidiaries are different from their immediate parent than a higher value), care needs to be taken when comparing these statistics *across* firms because some empirical label distributions are more easily

34

Table 2. Country Perfect Tree Similarity Statistics

	May 26, 2011				February 25, 2013			
	Perfect	Mean	Stdev	Quantile	Perfect	Mean	Stdev	Quantile
SIFIs								
S1	0.776	0.677	0.349	27.63%	0.850	0.523	0.271	84.98%
S2	0.351	0.362	0.330	58.43%	0.609	0.571	0.253	47.51%
S3	0.526	0.375	0.209	75.75%	0.727	0.527	0.212	84.97%
S4	0.800	0.615	0.364	47.97%	0.987	0.950	0.089	100%
S5	0.529	0.275	0.181	97.92%	0.964	0.869	0.205	100%
S6	0.651	0.397	0.387	65.96%	0.939	0.784	0.220	100%
S7	0.906	0.799	0.237	99.47%	0.834	0.574	0.232	92.58%
S8	0.869	0.655	0.308	58.19%	0.953	0.732	0.204	98.10%
S9	0.895	0.831	0.272	33.57%	0.953	0.875	0.198	43.97%
S10	0.261	0.258	0.127	53.79%	0.925	0.712	0.217	100%
S11	0.279	0.094	0.051	100%	0.939	0.795	0.226	100%
S12	0.264	0.096	0.081	93.39%	0.942	0.777	0.244	100%
S13	0.857	0.693	0.343	35.08%	0.951	0.794	0.218	100%
S14	0.864	0.746	0.341	29.33%	0.908	0.532	0.239	100%
S15	0.895	0.639	0.243	100%	0.924	0.684	0.239	99.53%
S16	0.860	0.665	0.218	100%	0.826	0.539	0.243	99.71%
S17	0.575	0.349	0.200	90.20%	0.856	0.445	0.154	100%
S18	0.963	0.896	0.173	67.42%	0.973	0.899	0.201	99.94%
S19	0.933	0.854	0.164	100%	0.935	0.726	0.300	100%
Mean	*0.687*	*0.541*			*0.894*	*0.700*		
Non-SIFIs								
B1	0.962	0.903	0.129	99.61%	0.959	0.881	0.161	100%
B2	0.941	0.830	0.255	37.04%	0.991	0.956	0.090	100%
B3	0.520	0.159	0.136	100%	0.510	0.157	0.135	97.16%
B4	0.669	0.374	0.241	100%	0.990	0.964	0.107	99.98%
B5	0.429	0.183	0.091	100%	0.783	0.428	0.241	99.96%
Mean	*0.704*	*0.490*			*0.847*	*0.677*		
Insurance Companies								
I1	0.665	0.536	0.367	41.18%	0.862	0.569	0.231	91.43%
I2	0.297	0.181	0.171	64.52%	0.752	0.384	0.190	99.74%
I3	0.576	0.281	0.270	79.46%	0.744	0.373	0.167	100%
I4	0.857	0.718	0.338	34.77%	0.884	0.602	0.230	90.91%
I5	0.781	0.626	0.254	60.35%	0.909	0.558	0.205	99.91%
Mean	*0.635*	*0.468*			*0.830*	*0.497*		

Notes to Table 4: Summary statistics describing the regularity of a firm's country control hierarchy for two periods in time, May 26, 2011 (left panel) and February 25, 2013 (right panel). The columns labeled "Perfect" refer to the perfect tree-similarity statistic equal to the number of nodes that have the same country designation as their immediate parent divided by the total number of nodes in the firm's tree. In a perfect tree, deletion of a node and its children will not change the value of the statistic. From a supervisory perspective, this means that supervisory oversight follows firm organizational lines. A number closer to one signifies greater proximity to perfect. The columns "Mean", "Stdev" and "Quantile" are summary statistics computed over all replications of the simulated data, where each firm's tree structure is taken as given and the node identifiers are randomly drawn according to the conditional probability distribution of the firm's observed tree. See section 3.2 for details.

shuffled to achieve proximity to a perfect tree than others. For this reason, it is useful also to consider the moments and/or quantiles generated using the perfect tree simulations. For example, a comparison of S9 and S11 in 2011 reveals that although S9 has a perfect tree statistic that is close to one, the arrangement of its subsidiaries is actually far from perfect (falling in the lowest third of the simulated distribution). In contrast, S11 has a much lower perfect tree statistic (0.279) but is in the 100% quantile compared to other control hierarchies that could result from its given country distribution, meaning that the simulations produced no better way to arrange the subsidiaries for this firm (i.e., in a manner that would result in less coordination burden among supervisors) than the observed arrangement.

In 2011, nine of the 29 firms had country structures that were significantly different from their corresponding perfect tree, given their topology (that is, a test of the null hypothesis that the firm's perfect tree similarity statistic is equal to one is rejected). In addition, 25 of 29 firms had country distributions that differed from a structure where all subsidiaries differ from their immediate parent (that is, a test of the null hypothesis that the firm's perfect tree similarity statistic is equal to zero is also rejected), suggesting most firms have majority-control structure that follows geographical lines. By 2013, only one firm had a country structure that was significantly different from its corresponding perfect tree and for all firms, the null hypothesis that the firm's perfect tree similarity statistic is equal to zero was rejected, corroborating the results of the Markov statistics of a shift toward geographical consolidation.

4.3.2 SIC Perfect Tree Similarity Statistics

Table 3 summarizes the 2-digit SIC-level perfect tree simulations for each firm and both dates, again grouped by firm type.[29] Similar to the country tree statistics, while in 11 of the 29 firms we fail to reject the null hypothesis that the firms have 2-digit SIC structures where all subsidiaries differ from their immediate parent in 2011 (perfect tree statistic is equal to zero), by 2013 the hypothesis is rejected for *all* of the firms' structures (at the 95% level of confidence). The results document a shift toward a perfect tree configuration; all except one firm's perfect tree similarity statistic (S19) increased and the number of firms with SIC perfect tree-similarity significantly less than one fell from 20 to 10. Nonetheless, in many cases the change in 2-digit SIC structure did not actually render the firms less complex, as the associated quantile declined for 14 of 29 firms.

An interesting example from this table is S18. From Table A1, we know that this firm experienced a modest (13%) reduction in nodes between 2011 and 2013, despite adding five additional countries to its majority-control hierarchy. Over the same period, it reduced its number of SIC categories by more than 50% and doubled its degree depth.

As noted above in Section 4.2, the increased degree depth is an indication of a shift toward a more bureaucratic organizational structure. Commensurate with this change, the firm's country perfect tree statistic increased both in level and quantile (see Table 4), indicating that the new control hierarchy was closer to a segmented structure, where children mimic the country of their immediate parent. This shift to a more

---

[29] We use 2-digit SIC classifications for computation of the perfect tree similarity statistic rather than the more granular 4-digit since the former will be farther away from the lower bound of zero and hence is more informative for illustrative purposes. In practice, of course, these statistics may be computed at whichever level of granularity is deemed most appropriate, i.e., the statistics also could be computed using 1-, 3-, or 4-digit granularity if desired, recognizing that higher levels of granularity involve increased computational time for the associated simulations required for inference.

Table 3. SIC Perfect Tree Similarity Statistics

	May 26, 2011				February 25, 2013			
	Perfect	Mean	Stdev	Quantile	Perfect	Mean	Stdev	Quantile
SIFIs								
S1	0.305	0.069	0.071	99.94%	0.631	0.376	0.158	99.25%
S2	0.103	0.063	0.036	81.66%	0.466	0.649	0.197	18.33%
S3	0.092	0.053	0.041	90.03%	0.575	0.358	0.169	99.99%
S4	0.561	0.704	0.301	28.95%	0.955	0.917	0.133	32.60%
S5	0.089	0.064	0.066	79.74%	0.940	0.885	0.174	54.21%
S6	0.435	0.300	0.314	64.14%	0.830	0.731	0.191	51.65%
S7	0.022	0.090	0.028	6.26%	0.601	0.384	0.195	79.31%
S8	0.473	0.467	0.272	45.21%	0.754	0.663	0.180	61.85%
S9	0.359	0.360	0.194	42.19%	0.841	0.666	0.261	56.91%
S10	0.032	0.044	0.026	42.63%	0.767	0.591	0.257	59.14%
S11	0.465	0.283	0.107	94.64%	0.938	0.869	0.134	90.72%
S12	0.170	0.273	0.170	36.56%	0.820	0.749	0.229	42.85%
S13	0.856	0.565	0.313	94.99%	0.879	0.794	0.182	50.48%
S14	0.078	0.115	0.101	50.96%	0.706	0.410	0.268	76.23%
S15	0.683	0.489	0.217	79.96%	0.784	0.583	0.171	93.88%
S16	0.417	0.262	0.169	69.79%	0.491	0.517	0.157	37.28%
S17	0.291	0.160	0.118	85.78%	0.471	0.211	0.056	100%
S18	0.666	0.501	0.203	73.02%	0.844	0.721	0.252	39.30%
S19	0.836	0.676	0.227	75.03%	0.755	0.432	0.265	92.38%
Mean	*0.365*	*0.291*			*0.739*	*0.606*		
Non-SIFIs								
B1	0.754	0.651	0.261	52.07%	0.871	0.754	0.240	53.46%
B2	0.621	0.360	0.303	66.37%	0.968	0.930	0.122	64.92%
B3	0.276	0.117	0.077	96.83%	0.671	0.447	0.259	67.00%
B4	0.103	0.037	0.033	90.63%	0.957	0.933	0.155	16.40%
B5	0.302	0.213	0.110	71.40%	0.659	0.552	0.284	44.54%
Mean	*0.411*	*0.276*			*0.825*	*0.723*		
Insurance Companies								
I1	0.375	0.433	0.299	45.65%	0.546	0.366	0.167	85.66%
I2	0.212	0.112	0.042	98.46%	0.488	0.272	0.081	99.33%
I3	0.125	0.069	0.047	84.64%	0.596	0.432	0.210	63.10%
I4	0.478	0.415	0.246	54.83%	0.497	0.401	0.126	71.40%
I5	0.119	0.055	0.018	99.78%	0.564	0.371	0.112	96.94%
Mean	*0.262*	*0.217*			*0.538*	*0.368*		

Notes to Table 5: Summary statistics describing the regularity of a firm's SIC control hierarchy for two periods in time, May 26, 2011 (left panel) and February 25, 2013 (right panel). The columns labeled "Perfect" refer to the perfect tree-similarity statistic equal to the number of nodes that have the same SIC designation as their immediate parent divided by the total number of nodes in the firm's tree. In a perfect tree, deletion of a node and its children will not change the value of the statistic. From a supervisory perspective, this means that supervisory oversight follows firm organizational lines. A number closer to one signifies greater proximity to perfect. The columns "Mean", "Stdev" and "Quantile" are summary statistics computed over all replications of the simulated data, where each firm's tree structure is taken as given and the node identifiers are randomly drawn according to the conditional probability distribution of the firm's observed tree. See section 3.2 for details.

divisional arrangement was not without cost, however; from Table 3 despite more parent-child SIC alignment (the perfect tree statistic increased from 0.666 to 0.844), the resultant firm was less functionally-arranged as it dropped from the 73rd percentile to the 39th percentile in its proximity to a perfect tree.

4.3.3 Additional Thoughts

As noted in Section 3, complexity in our framework is a multidimensional concept. It is evident from a comparison of Tables 2 and 3 that while in many cases inference regarding a firm's complexity may be similar in the two dimensions we consider (i.e., country and SIC), in other cases the two dimensions may conflict, for example with a firm having a fairly complex country-label topology and a not-so-complex SIC-label topology, or vice versa.[30] In principle, such conflict can be addressed by redefining (and expanding) the label space to accommodate all possible label pairs (i.e., all country-SIC combinations). This framework also can be used to consider additional types of labels beyond just country and SIC classifications (e.g., by activity), were such information available. More generally one can define labels via *n*-tuples, where *n* represents the number of properties that describe each node. In practice, however, such a definitional expansion soon becomes computationally prohibitive.

---

[30] A similar point is made by Carmassi and Herring (2015) in discussion of their subsidiarization proposal (discussed above) where they note, "Subsidiarization does not provide a complete solution to the problem of corporate complexity, because subsidiaries can be organized on the basis of their location or on the basis of the kind of business conducted, but these approaches often conflict."

5. Discussion

In this section we provide some thoughts to illustrate how results such as ours might be used in policy discussion and the implications of our findings for both the determination of systemic importance and financial stability. We emphasize that due to our limited sample, the remarks are offered as a proof-of-concept rather than intended to be definitive.

5.1. Are the SIFIs in our sample really more complex from a supervisory perspective?

In Table A1, we see that in 2011 SIFIs had tree structures with more than three times as many nodes, higher degree depth, greater geographical reach and more than double the amount of 4-digit SIC variation in their subsidiaries than both non-SIFI banks and insurance companies. In addition, data from Bloomberg® on total consolidated assets of all firms in our sample reveal that on average the SIFIs had $1.82tr in total consolidated assets as compared to $0.72tr for the non-SIFI banks and $0.61tr for the insurance companies in our dataset. So do our complexity measures tell a regulator anything more about the appropriateness of a systemic importance designation than can be gleaned from simply looking at the size of an institution?

Note that in asking this question, we do not mean to downplay the relevance of the other metrics (e.g., number of nodes, tree depth, etc.) or to suggest that our perfect tree similarity statistic is the only important property (or even the most important property) of a majority-control hierarchy network. To reiterate, we consider the term "complexity" from a supervisory perspective regarding the topology of the firm, namely: (1) how difficult is it for a consolidated supervisor to access the information it needs to

40

assess the firm, (2) how likely are they to identify a problem with the firm if it exists, and (3) how easily can a problem be mitigated/remedied once identified? Recall our assumption in this context is that the closer a firm's control hierarchy is to a perfect tree, the easier will be its access to needed information.

From this supervisory perspective, one might at first glance consider the SIFI sample to be on average more complex than either the non-SIFI bank or insurance company samples, as a result of the SIFI sample's much more elaborate control hierarchies based on the dimensions given in the data (# of nodes, countries, SIC groups, and degree depth). Yet for the most part the SIFI sample also had a greater proportion of child nodes that were from the same country or SIC classification as their immediate parent than did the insurance companies. Importantly, given their elaborate control hierarchy, the firms in the SIFI sample were in many cases closer to a perfect tree than most of the simulated trees, judging from their quantile position. Thus the SIFIs in our sample may not necessarily present greater supervisory challenges, despite their larger size and more elaborate structures, assuming sufficient oversight in each country and industry. Put another way, the ability of a firm to quickly reduce its exposure with respect to a specific country or industry appears to be similar between the SIFI and non-SIFI bank samples. In contrast, the sample of insurance companies in 2011 appears to be more complex, with more extensive cross-country and cross-industry reporting structures, that might prove harder to untangle in a crisis. For this reason, we believe our proposed intrafirm metrics may be a useful complement in the supervisory context to the more general well-established network metrics that are used to evaluate interconnected networks. We emphasize that at this stage, our findings are merely suggestive and not conclusive, given

the subsample of data that we have for analysis. Of course it remains to be seen whether our findings also appear in a larger, representative sample.

5.2. *Is size a sufficient statistic?*

In the aftermath of the 2008 financial crisis, calls to end "too big to fail," the so-called practice of bailing out the largest, most systemically important financial institutions, have intensified. Most often, the concept of too big to fail implies a firm whose size is larger than a specified threshold. Yet size is but one of the criteria mentioned in the SIFI definition.[31]

Consistent with the idea that size and systemic importance are synonymous, regulators typically delineate a size threshold to identify firms that pose considerable risk to the global financial system. Despite the ease of implementation, a size-based threshold is in many ways unsatisfactory, precisely because it does not take into account the level of complexity of a firm's business activities. To quantify this point more generally, we computed the Pearson rank correlation between total consolidated assets and both country- and SIC- perfect tree similarity statistics.[32] The rank correlation was -0.32 and -0.36, respectively. In contrast, the rank correlation between size and the number of nodes was 0.58, reflecting the fact that firms with more total consolidated assets generally have more subsidiaries. Taken together, these numbers highlight the fact that the link between asset size, number of subsidiaries, and the number of countries and

---

[31] There are a variety of definitions of size that arise in the banking and finance literature. The most common in recent banking regulations (e.g., Basel II, Basel III, Dodd-Frank) is specified in terms of total consolidated assets. Other definitions might include market capitalization, number of distinct entities, number of employees. Generally speaking, however, in the "too big to fail" context, size is usually considered in financial (e.g., dollar) terms, rather than in terms of features of organizational structure.
[32] We report results using rank-based statistics in order to preserve confidentiality of the firms. The results are qualitatively similar using ln(assets).

industries in which a firm operates is not exact; that is, the perfect tree statistics are measuring something different than just the number of nodes. This point is also illustrated in Figure 5, comparing the ranks of the firms' size to the ranks of their country (left chart) and SIC (right chart) perfect tree similarity statistics. In particular, if size and the perfect-tree statistics measured the same thing, we would expect the points in this figure to lie along the 45-degree line. Instead there is a slight negative relationship, particularly among the SIFI firms; the larger firms are less complex from a supervisory standpoint, with a larger proportion of trees being self-contained (where the child node is in the same country as its immediate parent).

To see how such information might be used in practice, we offer two examples:

(1) Firms S4 and S5 have very similar asset-size, yet in 2013 S4 has three times as many subsidiaries (nodes) and activities in 50% more industries than S5. In contrast, it is only active in half as many countries. Hence the supervisory challenges associated with these firms may be very different, S4 requiring coordination among many more industry regulators and S5 requiring substantial coordination among different country supervisors. Yet both their country and SIC perfect tree-similarity statistics are similar and close to one. And while their country perfect tree similarity statistics would place them both at the 100% quantile (i.e., closer to a perfect tree than all simulated firms with identical tree structure), S4's SIC statistic, while higher in value than S5's, is below the 33rd percentile (by comparison, S5's is just below the 54th percentile). Taken together, these statistics would alert regulators to the SIC dimension of S4's business. Note from Table A1 that the absolute number of 4-digit SIC activities of S4 is not particularly unusual: there are 12 firms that have
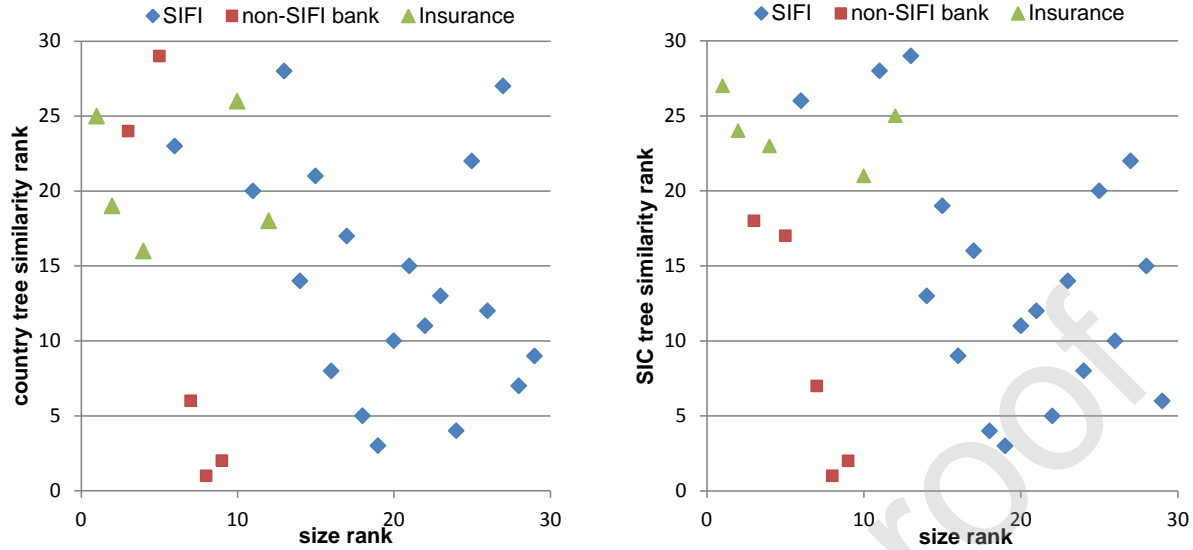
43

Figure 5. Scatter plot of firms' size rank versus country and SIC tree similarity ranks. These charts show the distribution of size rank (in terms of 2012 total consolidated assets) versus 2013 country (top chart) and SIC (bottom chart) tree similarity ranks, for all 29 firms in our sample, shown separately for SIFIs, non-SIFI banks, and insurance companies. *A higher rank indicates greater complexity and/or larger size.* While it is evident that the non-SIFI banks and insurance companies are smaller in size than the SIFIs, there is also evidence that the insurance companies have country and SIC perfect tree similarity statistics that are farther from a perfect tree (i.e., they are more intermingled and hence potentially more difficult to supervise/regulate). In contrast, the non-SIFI are similar to their SIFI counterparts in that some have tree structures that are close to perfect and others far away. Data on size as of December 31, 2012 are obtained from Bloomberg®.

subsidiaries in a greater number of industries. Yet there is only one firm (S2) that has a

2-digit SIC perfect tree similarity statistic with a lower quantile score.

(2) Another comparison of interest is between S6 and S12. These two firms are very

similar in their perfect tree similarity statistics (both country and SIC), as well as the

quantiles to which these statistics correspond. Both firms' SIC quantiles are near the

median of the simulated distribution for their corresponding tree structures and hence

might warrant additional supervision. Yet S6 is nearly 50% bigger in asset size and

has more than four times the number of subsidiaries (nodes) than S12. A size-only

threshold would potentially miss the complexity of S12.

Finally, Cohen and Lou (2012) use a Herfindahl-type measure to assess how complicated a firm is, computed as the sum of squared percentages of each business type (e.g., 0.4^2 $+0.3^2 + 0.3^2$ if there are three types that comprise 40%, 30%, and 30% of the firm's overall business) but do not take into account the topology of the firm's organizational structure. The closer their measure is to one, the less complicated a firm is. To compare our work to this important paper, using our 2011 data we analogously compute indices that measure the country and SIC concentration of our firms. The correlation between these indices and our perfect tree statistics is 0.83 (country) and 0.72 (SIC). Yet when we consider the rank correlation between the indices and the perfect tree statistic quantiles, there is a negative relationship, providing further evidence that our complexity metric is detecting something different than business concentration.[33]

5.3. Has complexity of the firms in our sample changed over time?

A comparison of the left and right blocks of columns in Table 2 and 3 (corresponding to results as of May 26, 2011, and February 25, 2013, respectively), reveals a number of interesting observations:

- In 2011, four of the 29 firms had country structures that were not significantly different from a random tree structure (a perfect tree similarity statistic of zero). By 2013, all of the firms' structures differed.

- Despite their nonrandom structure, in the 21 months between May 2011 and February 2013, the firms in our sample substantially reduced their level of geographical complexity. In the earlier sample, nine firms (6 SIFIs, 2 non-SIFI banks, and 1

---

[33] The values of the indices and scatter plots comparing the ranks discussed in this paragraph are available on request.

45

insurance company) had country structures that were statistically significantly different from a perfect tree (a perfect tree similarity statistic of one), while by 2013, only one (B3) remained. Carmassi and Herring (2015) also note that complexity (defined as the number of subsidiaries of a firm) has declined since 2011.

- The reduction in geographical complexity was partially achieved through a change in structure and reduction in geographical intraconnectedness for each firm. For 24 of the 29 firms, the country perfect-tree similarity statistic moved closer to one. In addition, 20 firms experienced quantile increases as a result of their change in structure between the two years, indicating closer proximity to a perfect tree structure. In addition, over all 29 firms the range of statistics narrowed (from [0.261 to 0.963] in 2011 to [0.510 to 0.991] in 2013).

- By 2013, most firms were in the very upper tail of their respective country perfect tree distributions, with nearly half of the statistics above *all* of the simulated values corresponding to their firm's structure (quantile = 100%).

- In contrast, while most firms' SIC structure moved more closely to a perfect tree in 2013 (i.e., the number of firms with SIC perfect tree-similarity statistic significantly less than one fell by 50%, from 20 to 10), the associated quantiles declined in almost half the firms, indicating that the change in SIC structure did not commensurately reduce the firms' complexity along this dimension.

- As noted above, supervisors might also use measures such as these to inform their choices of which firms merit additional scrutiny. For example, for firm S2, despite an increase in its county and SIC perfect tree similarity statistics between 2011 and 2013, the associated quantiles declined from 58.42% to 47.51% for the country

perfect tree similarity statistic and from 81.66% to 18.33% for the SIC perfect tree similarity statistic, indicating that the firm's geographical and industry complexity may not have declined commensurately with the change in structure.

Taken together, these results indicate some reduction in country and SIC complexity for most of the firms in our sample. While the firms in our SIFI sample have made the most progress, there have also been large changes in the non-bank SIFIs and insurance company samples. There is little evidence that the increase in size as a result of the post-2008 financial crisis consolidation of the banking sector has led to greater jurisdictional complexity; for most firms, coordination with respect to both supervisory oversight and possible wind-down or paring of assets would be easier now, given their organizational structure, than in 2011.

6. Conclusions

The 2008 financial crisis highlighted the risks that large, multinational, complex, interconnected banks pose. Since then, debate concerning which firms warrant a SIFI designation has led to a large amount of research into the web of interconnections inherent in financial transaction networks and the associated risk to the global financial system. There has been comparatively less emphasis on the risks that a firms' internal complexity might pose, particularly the coordination challenges for supervisors tasked with evaluating a firm's riskiness on a consolidated basis. As noted by Carmassi and Herring (2014):

"[Also,] opacity of organizational structures impedes regulatory oversight. If regulators do not have a clear understanding of how lines of business map into

47

legal entities and how the legal entities interact with each other, they cannot perform effective prudential supervision, nor can they implement an orderly resolution"

"Information tends to flow more freely within a regulatory organization than between organizations and more freely among regulatory institutions within a country than across national borders."

In this paper, we propose using a firm's control hierarchy as a proxy for such supervisory challenges. By defining complexity as a function of the firm's tree topology we demonstrate that complexity and size are not synonymous and thus warrant distinct mention in the FSB's SIFI definition. We additionally propose using a perfect tree similarity statistic to quantify the ease of supervision in a number of dimensions: (1) a consolidated supervisor's need to rely on coordination or information from supervisors from other countries, (2) the implications of severing the link to a subsidiary, and (3) the ability to assess how to wind down a firm through paring of subsidiaries with minimal risk.[34] By comparing data from 2011 and 2013, we find that on average the ease of supervision along these dimensions improved. While our data sample prevents definitive conclusions regarding this improvement, one possible reason is the 2012 lowering of the threshold under which subsidiaries must submit to consolidated supervision. Although purely speculative at this stage, the pattern that we have observed in our data, namely of dramatic increases in the number of nodes and decreases in the level of complexity as defined by our metrics, is consistent with what we would expect from such a change in

---

[34] Although we have emphasized the potential value of our metrics to the supervisory/regulatory community, it is also worth noting that our metrics additionally might be useful as measures of operational risk. The Risk Management Association on their website (http://www.rmahq.org/operational-risk/?gmssopc=1, accessed 8/12/16) notes, "The definition of operational risk is the risk of loss resulting from inadequate or failed internal processes, people, and systems, or from external events, but is better viewed as the risk arising from the execution of an institution's business functions." We thank several people at various conferences and seminars where an early version of this paper was presented for mentioning this connection to operational risk.

threshold. No doubt other explanations are possible. In addition, our results suggest that contrary to conventional wisdom, some of the SIFI-designated institutions may not pose greater supervisory challenge since their control hierarchy more closely resembles a perfect tree than some of the other firms in our sample. Specifically, we find little difference between the SIFI and non-SIFI banks in our sample. In contrast, the insurance companies in our sample are more complex according to the perfect tree similarity statistics, despite being smaller in size, having fewer subsidiaries, and being less geographically or industry-diverse than the banks. We emphasize, however, that our sample is highly selected and that generalizability of these findings remains a topic for further research.

In closing we stress the importance of further development of metrics aimed at measuring complexity as it relates to the challenges of supervisory oversight and coordination. As a key aspect of the FSB's definition of systemic importance, viable measures of complexity are critical to ensuring financial stability. In this paper, we make the argument that the firms' topology may provide insights into the challenges that external supervisors are known to face. In our view, the metrics proposed in this paper represent just the beginning of the discussion on this topic. We emphasize that, distinct from size and interconnectedness, much work on complexity remains to be done.

References.

Adamic L, Brunetti C, Harris JH, Kirilenko A (2012) Trading Networks, unpublished working paper.

Adams, RB (2010) Governance of Banking Institutions. Anderson R, Baker HK, eds. *Corporate Governance* (John Wiley & Sons, Inc., New York) 451-467.

Aldasoro I, Alves I (2018) Multiplex interbank networks and systemic importance: An application to European data. *Journal of Financial Stability* 35 (April), 17-37

Alessandri P, Masciantonio S, Zaghini A (2015) Tracking banks' systemic importance before and after the crisis. *Banca d'Italia occasional paper 259.*

Arinaminpathy N, Kapadia S, May RM (2012) Size and complexity in model financial systems. *Proceedings of the National Academy of Sciences* 109(45), 18338-18343.

Armour H, Teece D (1978) Organizational structure and economic performance: a test of the multidivisional hypothesis. *The Bell Journal of Economics* 9(1), 106-122.

Basel Committee on Bank Supervision (2014). *The G-SIB assessment methodology: score calculation* (Bank for International Settlements, Basel).

Battiston S, Farmer JD, Flache A, Garlaschelli D, Haldane AG, Heesterbeek H, Hommes C, Jaeger C, May R, Scheffer M (2016a) Complexity theory and financial regulation. *Science* 351 (6275), 818-819. doi: 10.1126/science.aad0299

Battiston S , Martinez-Jaramillo S (2018) Financial networks and stress testing: Challenges and new research avenues for systemic risk analysis and financial stability implications. *Journal of Financial Stability* 35 (April) 6-16.

Battiston S, Puliga M, Kaushik R, Tasca P, Caldarelli G (2012) DebtRank: Too central to fail? Financial Networks, the FED and Systemic Risk. *Nature* 2:541, doi:10.1038/srep00541

Battiston S, Caldarelli G, May RM, Roukny T, Stiglitz JE (2016b) The price of complexity in financial networks. *Proceedings of the National Academy of Sciences* 113(36), 10031-10036.

Berglof E, Perotti E (1994) T he governance structure of the Japanese financial keiretsu. *Journal of Financial Economics* 36, 259-84.

Bernanke BS (2013) Monitoring the Financial System, speech delivered at the 49th Annual Conference on Bank Structure and Competition, Chicago, IL, May 10.

Berndsen RJ, León C, Renneboog L (2018) Financial stability in networks of financial institutions and market infrastructures. *Journal of Financial Stability* 35 (April), 120-135.

Billio M, Getmansky M, Lo AW, Pelizzon L (2012) Econometric measures of connectedness and systemic risk in the finance and insurance sectors. *Journal of Financial Economics* 104: 535-559, doi:10.1016/j.jfineco.2011.12.010.

Bisias D, Flood M, Lo AW, Valavanis S (2012) A survey of systemic risk analytics. *Annual Review of Financial Economics* 4: 255-96.

Blair CE, Kushmeider RM (2006) Challenges to the Dual Banking System: The Funding of Bank Supervision. *FDIC Bank Review* 18(1): 1-22.

Board of Governors of the Federal Reserve System (2017). *BHC Supervision Manual*. (Board of Governors of the Federal Reserve System, Washington, DC), September. https://www.federalreserve.gov/publications/files/bhc.pdf

Briscoe F, Tsai W (2011) Overcoming relational inertia: how organizational members respond to acquisition events in a law firm. *Administrative Science Quarterly* 56(3): 408-440.

Caballero RJ, Simsek A (2013) Fire sales in a model of complexity. *The Journal of Finance* 68(6), 2549-2587.

Caccioli F, Farmer JD, Foti N, Rockmore DN (2015) Overlapping portfolios, contagion, and financial stability. *Journal of Economic Dynamics and Control* 51:50–63.

Carmassi J, Herring RJ (2014) Corporate Structures, Transparency and Resolvability of Global Systemically Important Banks. Unpublished working paper.

Carmassi J, Herring RJ (2015) The Corporate Complexity of Global Systemically Important Banks. Unpublished working paper.

Cetorelli N, Goldberg LS (2014) Measures of Global Bank Complexity. *Economic Policy Review* 20(2).

Cetorelli N, McAndrews J, Traina J (2014) Evolution in Bank Complexity. *FRBNY Economic Policy Review* 20(2).

Coase RH (1937) The nature of the firm. *Economica* 4(16): 386-405.

Cohen L, Lou D (2012) Complicated firms. *Journal of Financial Economics* 104: 383-400.

Cohen-Cole E, Patacchini E, Zenou Y (2010) Systemic risk and network formation in the interbank market. CAREFIN, Università Bocconi working paper #25/10.

Constantin A, Peltonen TA, Sarlin P (2018) Network linkages to predict bank distress. *Journal of Financial Stability* 35 (April), 226-241.

De Young R, Evanoff DD, Molyneux P (2009) Mergers and acquisitions of financial institutions: A review of the post-2000 literature. *Journal of Financial Services Research* 36(2), 87-110.

Elsinger H, Lehar A, Summer M (2006) Risk Assessment for Banking Systems. *Management Science* 52(9): 1301-1314. http://dx.doi.org/10.287/mnsc.1060.0531.

Financial Stability Board (2011) Policy measures to address systemically important financial institutions, available at www.financialstabilityboard.org/wp-content/uploads/r_111104bb.pdf?page_moved=1, accessed May 11, 2011.

Financial Stability Board (2019) 2019 list of global systemically important banks (G-

SIBs), available at http://www.fsb.org/2019/11/2019-list-of-global-systemically-important-banks-g-sibs/, accessed January 30, 2020.

Flood MD, Kenett DY, Lumsdaine RL, Simon JK (2017) The Complexity of Bank Holding Companies: A Topological Approach. NBER working paper #23755.

Haldane AG, May RM (2011) Systemic risk in banking ecosystems. *Nature* 469:351-355.

Haubrich JG, Lo AW (2013), eds. Quantifying systemic risk (The University of Chicago Press: Chicago, IL).

Hautsch N, Schaumburg J, Schienle M (2012) Financial network systemic risk contributions. CRC 649 working paper 2012-053.

---- (2013) Forecasting systemic impact in financial networks, SFB649 discussion paper 2013-008.

Kapadia S, Drehmann M, Elliott J, Sterne G (2013) Liquidity risk, cash flow constraints, and systemic feedbacks. Haubrich JG, Lo AW, eds. Quantifying systemic risk (The University of Chicago Press: Chicago, IL), 29-61.

León C, Machado C, Sarmiento M (2018) Identifying central bank liquidity super-spreaders in interbank funds networks. *Journal of Financial Stability* 35 (April 2018), 75-92.

Lumsdaine RL, Rockmore DN, Foti N, Leibon G, Farmer JD (2016) The Intrafirm Complexity of Systemically Important Financial Institutions, available at https://ssrn.com/abstract=2604166

National Academy of Sciences (2010) *Technical Capabilities Necessary for Regulation of Systemic Financial Risk: Summary of a Workshop* (National Academy of Sciences, Washington, DC).

Newman M (2010) *Networks: An Introduction* (Princeton University Press, Princeton, NJ).

Roukny T, Battiston S, Stiglitz JE (2018) Interconnectedness as a source of uncertainty in systemic risk. *Journal of Financial Stability* 35 (April 2018), 93-106.

Roukny T, Bersini H, Pirotte H, Caldarelli G, Battiston S (2013) Default cascades in complex networks: Topology and systemic risk. *Scientific Reports* 3, 2759. doi: 10.1038/srep02759

Squartini T, van Lelyveld I, Garlashelli D (2013) Early-warning signals of topological collapse in interbank networks. Unpublished working paper.

Urwick LF (1956) The manager's span of control. *Harvard Business Review* (May-June): 39-47.

Viñals J, Pazarbasioglu C, Surti J, Narain A, Erbenova M, Chow J (2013) Creating a safer financial system: will the Volcker, Vickers, and Liikanen structural measures help? *IMF Staff Discussion Note* SDN/13/14.

Vitali S, Glattfelder JB, Battiston S (2011) The network of global corporate control. *PLoS ONE* 6(10): e25995. doi:10.1371/journal.pone.0025995

Appendix A: List of Financial Institutions and basic descriptive statistics
This is the list of financial institutions analyzed in this paper, broken out into banks and insurance companies and grouped by country of incorporation.[35] Those that are among the twenty-nine systemically important financial institutions (SIFIs), as determined by the FSB and IMF, are preceded by an asterisk. The country of incorporation is included in parentheses. For completeness, the SIFIs that are not included in our dataset (as a result of not being managed by Kingland Systems, the data provider) are also listed.

Banks

*Bank of America (US)
*Citigroup (US)
*Goldman Sachs (US)
*JP Morgan Chase (US)
*Morgan Stanley (US)
  Royal Bank of Canada (CA)
*Credit Suisse AG (CH)
*UBS AG (CH)
*Deutsche Bank AG (DE)
*Banco Santander SA (ES)
  BBVA (ES)
*BNP Paribas SA (FR)
*Société Générale SA (FR)
*Barclays PLC (GB)
*HSBC Holdings PLC (GB)
*Royal Bank of Scotland PLC (GB)
  Standard Chartered (GB)
  Banca Intesa (IT)
UniCredit (IT)
*Mitsubishi UFJ FG (JP)
*Mizuho FG (JP)

Banks (continued)

Nomura (JP)
*Sumitomo Mitsui FG (JP)
*ING Groep NV (NL)

Insurance Companies

Allianz (DE)
Aviva (GB)
Axa (FR)
Swiss Re (CH)
Zurich (CH)

SIFIs not included in the dataset:
[*Bank of China (CN)]
[*Banque Populaire (FR)]
[*Commerzbank (DE)]
[*Crédit Agricole (FR)]
[*Dexia (BE)]
[*Lloyds (GB)]
[*Nordea (SW)]
[*Wells Fargo (US)]

In Table A1 on the next page, we present some basic descriptive statistics about the (anonymized) firms in our sample, giving for each firm the total number of nodes (subsidiaries) in its tree, the number of distinct countries and 4-digit SIC codes represented by the nodes, and the number of levels (i.e., depth) of each tree. The table illustrates the variation in tree structure across the firms and across time. In particular, in 2011 the number of nodes in a tree ranges from 43 to 16,443; the number of distinct countries and 4-digit SIC codes represented by the nodes ranges from 10 to 89 and 13 to 281, respectively. In addition, the tree depth varies from 2 to 7. By 2013, there is less variation across firms. The number of nodes in a tree ranges from 330 to 12,752 while the number of distinct countries and 4-digit SIC codes ranges from 23 to 86 and 27 to 164, respectively. The decline in number of countries and 4-digit SIC codes is offset by the increase in tree depth for all but two firms, likely a result of post-crisis acquisitions. For some firms the increase in tree depth is substantial: for example, firm S17 increased from a tree depth of 4 in 2011 to 20 in 2013.

[35] This is the list that was in effect at the time the initial data collection took place (2011). Since then, the Financial Stability Board has issued a revised list every November. The SIFI status of all of the banks we analyze has not changed since the initial list. See Financial Stability Board (2019) for the most recent list.

Table A1. Descriptive Statistics

	May 26, 2011				February 25, 2013			
	#Nodes	#Countries	#SIC	Depth	#Nodes	#Countries	#SIC	Depth
SIFIs								
S1	1007	34	72	3	1519	57	60	4
S2	887	40	133	3	1585	47	59	10
S3	2568	55	210	5	4001	70	125	7
S4	1897	37	72	4	12752	33	86	6
S5	1034	42	122	3	4272	56	57	5
S6	3221	87	210	5	7289	73	96	7
S7	5850	58	198	4	5477	72	164	11
S8	6483	68	157	4	9564	76	147	9
S9	5502	48	194	5	8455	47	127	9
S10	1815	35	222	7	4012	48	107	8
S11	43	14	16	4	1468	23	34	5
S12	53	18	13	2	1520	29	45	5
S13	935	32	46	5	2224	32	46	5
S14	9815	76	281	5	3243	56	152	13
S15	9084	89	240	6	10211	86	127	9
S16	1778	32	100	5	2545	50	86	11
S17	2334	49	250	4	1117	38	104	20
S18	11487	47	279	6	10077	52	134	12
S19	16443	58	172	6	11231	61	114	7
Mean	*4328*	*48*	*157*	*4.5*	*5398*	*53*	*98*	*8.6*
Non-SIFI Banks								
B1	2678	19	72	4	2378	30	65	5
B2	1998	20	110	4	9079	31	63	4
B3	127	37	26	5	681	50	49	7
B4	475	32	147	3	7006	29	53	6
B5	205	28	34	3	387	29	42	5
Mean	*1097*	*27*	*78*	*3.8*	*3906*	*34*	*54*	*5.4*
Insurance Companies								
I1	793	40	48	5	1373	39	67	6
I2	118	25	27	5	330	30	27	6
I3	1564	74	154	3	2738	81	131	7
I4	1752	54	98	4	2544	48	86	7
I5	379	10	47	4	1254	33	67	9
Mean	*921*	*41*	*75*	*4.2*	*1648*	*46*	*76*	*7.0*

Notes to Table A1: Basic descriptive information and statistics on the control hierarchies for the twenty-nine institutions in our sample at two points in time: May 26, 2011 and February 25, 2013. *#Nodes* is the total number of nodes in the hierarchy; *#Countries* is the number of different countries that occur in the hierarchy; *#SIC* is the number of different four-digit SIC codes that occur in the hierarchy; *Depth* is the number of levels in the hierarchy tree.

Appendix B: The Federal Reserve's Responsibility for Consolidated Supervision

The following are relevant passages from the Bank Holding Company Supervision Manual (Board of Governors of the Federal Reserve System 2017) regarding the Federal Reserve's responsibility as consolidated supervisor (all page section numbers quoted from this source).

1. "The BHC Act provides for all BHCs, including financial holding companies formed under the Gramm-Leach-Bliley Act (GLBA), to be supervised on a consolidated basis by the Federal Reserve. Consolidated supervision of a BHC encompasses the parent company and its subsidiaries, and allows the Federal Reserve to understand the organization's structure, activities, resources, and risks, as well as to address financial, managerial, operational, or other deficiencies before they pose a danger to the BHC's subsidiary depository institutions." (page 61)

Consolidated supervision is important not just for the US but for other countries; thus any metrics we develop that can aid consolidated supervision will be relevant beyond just the US and Federal Reserve supervision.

2. "The Federal Reserve's consolidated supervision program has served as the benchmark for many of the current and evolving international standards for the consolidated supervision of financial groups. Key concepts that have been part of the Federal Reserve's approach to consolidated supervision for many years are reflected in the Basel Committee on Banking Supervision's Minimum Standards for Internationally Active Banks (1992), capital accords (1988 and 2006), and Core Principles for Effective Banking Supervision (1997 and 2006), and are now used by the International Monetary Fund and the World Bank in connection with their assessments of countries' bank supervisory regimes." (page 62)

The next passage highlights the Federal Reserve's assumed link between microprudential evaluation of an individual firm and financial stability.

3. "The consolidated supervision framework has two primary objectives: (1) enhancing resiliency of a firm to lower the probability of its failure or inability to serve as a financial intermediary, (2) reducing the impact on the financial system and the broader economy in the event of a firm's failure or material weakness" (section 2124.05). "The Federal Reserve's approach to consolidated supervision <is>…designed to further the objectives of fostering financial stability and deterring or managing the potential for possible financial crises." (Page 57).

The next passages highlight the need for coordination and reliance on information from other regulators (including state regulators) to avoid duplication of effort and links this to business model.

4. Throughout the Bank Holding Company Supervision Manual, and stated explicitly e.g., on page 61: "The guidance reiterates the importance of coordination with, and reliance on, the work of other relevant primary supervisors and functional regulators". "Effective consolidated supervision requires strong, cooperative relationships between the Federal Reserve and relevant primary supervisors and functional regulators." (page 62). "Information sharing among domestic and foreign supervisors, consistent with applicable law and the jurisdiction of each supervisor, is essential to ensure that a banking organization's global activities are supervised on a consolidated basis." (page 63). "For example, more independent Federal Reserve work typically will be required to assess consolidated asset quality or earnings for large complex BHCs with significant nonbank activities that are not functionally regulated. However, where all material holding company assets are concentrated in a single depository institution subsidiary, a minimal level of incremental Federal Reserve efforts typically will be required to assess consolidated asset quality and earnings." (page 84). "The Federal Reserve's conduct of consolidated supervision is central to and dependent on the coordination with, and reliance on, the work of other relevant primary supervisors and functional regulators." (page 374) "Many large complex institutions have interstate operations that expand with the continuation of mergers and acquisitions. In this environment, close cooperation with the other federal and state banking agencies is critical." (page 563).

The next quotes note the challenges for large multinational BHCs and link these challenges to macroprudential supervision.

5. "Financial trouble in one part of an organization can spread rapidly to other parts of the organization; moreover, large BHCs increasingly operate and manage their businesses on an integrated basis across corporate boundaries. Risks that cross legal entities or that are managed on a consolidated basis cannot be monitored properly through supervision directed at any one of the legal entity subsidiaries within the overall organization." (page 61) "There are, however, areas of focus for the Federal Reserve that are unique to a holding company's international operations. For example, some host-country legal and regulatory structures and supervisory approaches are fundamentally different from those in the United States. As a result, the banking organization often must devote additional resources to maintain expertise in local regulatory requirements. In some instances, privacy concerns have led to limits on the information a BHC's foreign office may share with its parent company, thereby limiting the parent company's ability to exercise consolidated risk management on a global basis. Additionally, while considerable progress has been made to strengthen supervisory crossborder cooperation and information sharing, the Federal Reserve and other U.S. supervisors have, at times, faced challenges in accessing information on a bank's or BHC's foreign operations or in carrying out examinations of cross-border or foreign activities." (page 89) "The Federal Reserve aims to reduce systemic risk…Supervision carried out under this framework will support a variety of macroprudential supervisory approaches beyond those already discussed including…(f) Enhancing international coordination with foreign counterparts, including national supervisors and international bodies such as the Basel

Committee on Bank Supervision, the Financial Stability Board, and the Senior Supervisors Group. These activities focus on enhancing oversight of internationally active financial firms and markets and on minimizing the opportunities for firms to take advantage of weaker or inconsistent regulations." (page 597)

The next quote delineates the relevance of the majority-control threshold used in this paper.

6. "section 3(a)(3) serves to prevent an existing bank holding company from increasing, without prior Board approval, its ownership in an existing subsidiary bank unless the BHC already owns 50 percent of the shares of the bank (section 3(a)(5)(B)). A bank holding company which owns more than 50 percent of a bank's shares may buy and sell those shares freely without Board approval, provided the ownership remains above 50 percent. If a bank holding company owns less than 50 percent of a bank's shares, prior Board approval is required before each additional acquisition of shares until the bank holding company's ownership of the bank reaches more than 50 percent." (page 1382)

The final quote provides the regulatory definition of "subsidiary" in the context of consolidated supervision.

7. "A banking and finance subsidiary generally is defined as any company engaged in banking or finance in which the parent institution directly or indirectly holds more than 50 percent of the outstanding voting stock, or which is otherwise controlled or capable of being controlled by the parent organization." (page 1509)