Measures of Financial Network Complexity
A Topological Approach

Mark D. Flood, OFR
Joint work with:
Jonathan Simon, U. of Iowa
Mathew Timm, Bradley U.

Financial Risk & Network Theory

Cambridge Centre for Risk Studies Seminar
Cambridge University, September 13-14, 2016

Views expressed in this presentation are those of the speaker(s) and not necessarily of the Office of Financial Research.
Views and opinions expressed are those of the speaker(s) and do not necessarily represent official OFR positions or policy.
Why worry about market complexity?

“I believe the threats to the financial system stem largely from two increasingly dominant market characteristics. The first is the complexity of the markets. The second is the tendency for the markets to move rapidly into a crisis mode with little time or opportunity to intervene.

... 

The challenges in supervising the financial system, and particularly in safeguarding against market crises and systemic risk, are centered in dealing with these two characteristics..”

-- Rick Bookstaber

Testimony before the House Financial Services Committee
October 2, 2007
There are (too) many options

- Defining “complexity” as an emergent phenomenon – Mitchell (2009)
  - Size
  - Entropy
  - Algorithmic information content
  - Logical depth
  - Thermodynamic depth
  - Statistical complexity
  - Fractal dimension
  - Degree of hierarchy

- Catalogs of measures
  - Lloyd (2001): 42 approaches to complexity measurement
  - Bonchev and Buck (2005): 54+ specific formulas
Complexity of financial markets

- Again, many **conceptual** possibilities
  - Arinaminpathy, Kapadia, and May (2012)
  - Caballero and Simsek (2009, 2013)
  - Markose, Giansante, and Shaghaghi (2012)
  - Delpini, Battiston, Riccaboni, et al. (2013)
  - Gai, Haldane, and Kapadia (2011)
  - Bookstaber (2007)
  - Haldane and May (2011)
  - Marsili and Anand (2013)
  - Schwarcz (2009)
  - Sheng (2010)
Requirements for complexity measurement

• Focus on financial markets
  – Transactions
  – Contractual exposures
  – Trading relationships

• Statistics at the network-level
  – Comparison across markets
  – Cardinal (numeric) metrics

• Financially meaningful
  – Sensitive to local interactions
  – Intuitive interpretation re: market institutions/practices
  – Capture multifaceted complexity
Mathematical abstraction

Preparing for topological analysis

Market activity

Counterparty network data

We will work with the undirected, unweighted, market graph, $\Delta$

Counterparty graph model

Image sources: Library of Congress, OFR analysis

Views expressed in this presentation are those of the speaker(s) and not necessarily of the Office of Financial Research.
Some Homology

Homology

• Study of cycles
  • Cycles can represent “holes”
  • Cycles can be boundaries

• Examples from simplicial homology
  • Points = vertices = 0-cells
  • Line segments = edges = 1-cells
  • Polygons = 2-cells,
  • etc...

• Boundaries
  • Vertices (0-cells) have no boundary
  • Edges, non-closed paths (1-cells) have endpoints
  • Polygons (other 2-cells) have perimeters
  • etc...

Image source: Kabaria and Lew (2016), OFR analysis
Some More Homology

Extension to more general graphs

• About paths, cycles, and their boundaries
  • Edge-path as a set (unordered) of “connected” edges
    • \langle ABC \rangle
  • Edge-cycle “ends” where it “starts” (boundary = 0)
    • \langle ABCHG \rangle
  • Edge boundary is difference between end points

• Visual depiction can be misleading

Image source: OFR analysis
Homology

- Graph as a collection of vector spaces
  - Linear combinations over $\mathbb{Z}_2 = \{0,1\}$
  - The spaces have vector bases of vertices, edges, etc.

- Example of $\mathbb{Z}_2$ arithmetic:

$$
\langle ABCGH \rangle + \langle DEFGH \rangle = \\
1\langle ABC \rangle + 1\langle H \rangle + 1\langle G \rangle + 1\langle DEF \rangle + 1\langle G \rangle + 1\langle H \rangle = \\
1\langle ABC \rangle + 1\langle DEF \rangle + 1\langle H \rangle + 1\langle H \rangle + 1\langle G \rangle + 1\langle G \rangle = \\
1\langle ABC \rangle + 1\langle DEF \rangle + 0\langle H \rangle + 0\langle G \rangle = \\
1\langle ABC \rangle + 1\langle DEF \rangle = \\
\langle ABCDEF \rangle
$$

Image source: OFR analysis
Homology

- Betti numbers count independent cycles
- Each vector space:
  - Has a dimension, which is its Betti number, $b_k(\Delta)$
  - $b_k(\Delta) = \text{rank } H_k(\Delta) = \text{rank } k^{th}$ homology group
  - $b_k(\Delta)$ is the number of essential cycles ("holes") that must be filled to eliminate $k^{th}$ homology

Image source: OFR analysis
Key Theoretical Result

Euler-Poincaré formula

- A new edge **must** connect two components, or create a cycle
  - $v(\Delta) - e(\Delta) = \text{rank } H_0(\Delta) - \text{rank } H_1(\Delta) = b_0(\Delta) - b_1(\Delta)$
  - $b_0(\Delta)$ is the number of connected components
  - $b_1(\Delta)$ is the number of non-redundant cycles

Image source: OFR analysis
Complexity Metrics I

Aggregate size metrics

- $K_v \equiv \nu(\Delta) = \text{number of vertices in the market graph, } \Delta$
- $K_e \equiv e(\Delta) = \text{number of edges, or “deals”}$
- $K_{e/v} \equiv e(\Delta) / \nu(\Delta) = \text{average degree}$

Insensitive to network structure:

$K_v = 6$
$K_e = 5$
$K_{e/v} = 0.8333$

Image source: OFR analysis
Counting simple cycles – closed-form calculations

- $K_{C3} \equiv$ number of triangular cycles in the market graph, $\Delta$
- $K_{C4} \equiv$ number of quadrilateral cycles
- $K_{C5} \equiv$ number of pentagonal cycles
- $K_{C\text{net}} \equiv$ “total” number of nettable cycles = $K_{C3} + K_{C4} + K_{C5}$

Zeroth and first homology on the market graph

- $K_{b1\Delta} \equiv$ rank $H_1(\Delta) = b_1(\Delta) =$ cycle rank of vector space $H_1(\Delta)$
- $K_{b0\Delta} \equiv$ rank $H_0(\Delta) = b_0(\Delta) =$ number connected components in $\Delta$
Netting operates on the directed network of obligations, $\Delta$.

- Oriented cycles in the digraph, $\Delta$, are potentially “nettable”.
- Each oriented (“nettable”) cycle in $\Delta$ has a “shadow” in $\Delta$.
- Netting “kills” a cycle in $\Delta$ by reducing edge weights so one is zero.
- Bookkeeping convention: replace orientations in $\Delta$ by signs on weights.
- Cycles in $\Delta$ are an upper bound for cycles in $\Delta$.
- Rank $b_1(\Delta)$ is lower bound on cycle elimination needed for acyclic graph.

Image source: OFR analysis.
Adding 2-dimensional cells

- Simple cycles (no edge repeats) as netting opportunities
- Topologically “kills” the cycle by filling its hole
- Do this for all simple cycles
- Extends $\Delta$ to a 2-dimensional cell complex
- Call the resulting structure: $\Delta^2_{\text{net}}$

Netting measures on $\Delta^2_{\text{net}}$

- $K_{b_1\Delta^2} \equiv b_1(\Delta^2_{\text{net}}) = \text{rank of vector space } H_1(\Delta^2_{\text{net}})$
- $K_{b_2\Delta^2} \equiv b_2(\Delta^2_{\text{net}}) = \text{netting redundancy} = \text{multiple cycles involving the same deals}$
Netting Example

Order of netting matters

- All nodes start with capital of $50
- Red node ultimately defaults
- Pink nodes are affected through spillovers

Image source: OFR analysis
Every undirected graph, $\Delta$, has a matching line graph, $\Lambda_\Delta$

- Each edge in $\Delta$ corresponds to a vertex in the line graph, $\Lambda_\Delta$
- Thus, $e(\Delta) = v(\Lambda_\Delta)$

Connect nodes in $\Lambda_\Delta$ if the corresponding edges in $\Delta$ share a node
- Thus, a highly central node in $\Delta$ implies a busy line graph, $\Lambda_\Delta$

Image source: OFR analysis
Simple cycles in the line graph, $\Lambda_\Delta$

- $K_{c3\Lambda} \equiv$ number of triangular cycles in $\Lambda_\Delta$
- $K_{c4\Lambda} \equiv$ number of quadrilateral cycles
- $K_{c5\Lambda} \equiv$ number of pentagonal cycles

Exploiting the line graph, $\Lambda_\Delta$

- $K_{e\Lambda} \equiv e(\Lambda_\Delta) =$ edge count in $\Lambda_\Delta =$ deal interactions
- $K_{b1\Lambda} \equiv b_1(\Lambda_\Delta) =$ rank $H_1(\Lambda_\Delta) =$ cycle rank of $\Lambda_\Delta$
Example: Central Counterparties (CCPs)

- $\Lambda_\Delta$ contains no new information ...
- ... but nonetheless reveals facts not obvious in $\Delta$
- Sensitive to the presence of highly central nodes
- CCPs as single point of failure:
Evaluating the Metrics

Does cycle rank detect chains?
- 100K random graphs, except some seeded w/chains
- Each graph has $v(\Delta)=12$ vertices and $e(\Delta)=15$
- Cycle counts/ranks respond to the chains

Cycle counts, $C_n(\Delta)$, and cycle ranks, $b_1(\Delta)$, on unseeded and seeded graphs

<table>
<thead>
<tr>
<th></th>
<th>$C_3(\Delta)$</th>
<th>$C_4(\Delta)$</th>
<th>$C_5(\Delta)$</th>
<th>$b_1(\Delta)$</th>
<th>$e(\Lambda_{\Delta})$</th>
<th>$C_3(\Lambda_{\Delta})$</th>
<th>$C_4(\Lambda_{\Delta})$</th>
<th>$C_5(\Lambda_{\Delta})$</th>
<th>$b_1(\Lambda_{\Delta})$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unbiased</td>
<td>2.19</td>
<td>2.81</td>
<td>3.20</td>
<td>4.58</td>
<td>32.32</td>
<td>21.90</td>
<td>36.65</td>
<td>77.71</td>
<td>18.37</td>
</tr>
<tr>
<td>two 4-chains</td>
<td>2.44</td>
<td>3.46</td>
<td>4.09</td>
<td>4.86</td>
<td>33.84</td>
<td>24.42</td>
<td>43.79</td>
<td>99.66</td>
<td>19.89</td>
</tr>
<tr>
<td>three 4-chains</td>
<td>2.58</td>
<td>3.81</td>
<td>4.64</td>
<td>5.04</td>
<td>34.67</td>
<td>25.83</td>
<td>48.11</td>
<td>113.43</td>
<td>20.71</td>
</tr>
<tr>
<td>three 3-chains</td>
<td>2.59</td>
<td>3.64</td>
<td>4.16</td>
<td>4.89</td>
<td>34.43</td>
<td>25.92</td>
<td>49.12</td>
<td>115.96</td>
<td>20.48</td>
</tr>
<tr>
<td>four 3-chains</td>
<td>2.72</td>
<td>3.94</td>
<td>4.54</td>
<td>5.02</td>
<td>35.14</td>
<td>27.23</td>
<td>53.38</td>
<td>129.92</td>
<td>21.18</td>
</tr>
</tbody>
</table>

Image source: OFR analysis
Evaluating the Metrics

Weighted graph stratification and persistent homology

- Exposure concentration – random (blue) vs concentration bias (red)
- Distribution of edge weights – 3 beta distributions
  - Unimodal ($\alpha=2.0$, $\beta=2.0$)
  - Skewed ($\alpha=2.0$, $\beta=5.0$)
  - Bimodal ($\alpha=0.5$, $\beta=0.5$)

Persistant homology of triangle counts in the line graph, $\Lambda_{\Delta}$, for random graphs with 3 different beta distributions of edge weights

Results

- Weight stratification exhibits noticeable distinctions
- Cycle counts strongly detect edge distribution differences

Image source: OFR analysis
Cycles need not depend on size

- Random regular graphs
- $K_v = 20 = \text{number of nodes}$
- $K_{e/v} = 10 = \text{average degree}$
- $K_{b0}$ is constant
- $K_{b1}$ is constant
- $K_{e\Lambda}$ is constant

$\Rightarrow$ Conditionally independent variation

Distribution of $K_{C3}$

Image source: OFR analysis
Evaluating the Metrics

Correlations between edges and triangles, $G(n,p)$ graphs

$$\text{Correlation}(\text{edg}, \text{tri}) = \frac{3 \binom{n}{3}(p^3 - p^4)}{\sqrt{\binom{n}{2}(p - p^2)} \times \left( \binom{n}{3}(p^3 - p^6) + 12 \binom{n}{4}(p^5 - p^6) \right)}$$

$G(20,p)$ graphs

- Correlation for 1000 graphs
- Expected value (from formula)

Image source: OFR analysis
Core-Periphery Theorem

“Simple” core-periphery graph, $\Delta^*_{cp}$

- Core has $k$ dealer nodes, completely connected
- Periphery has $(n - k)$ client nodes, each with one dealer

Interior optimum core size

**Theorem 1.** In a simple core-periphery graph, $\Delta^*_{cp}$, with $n$ total vertices of which $k$ are in the core, the number of edges in the line graph, $\Lambda_{\Delta^*_{cp}}$, and its cycle rank, are each minimized when:

- Each core party has the same number of clients (or whole numbers of clients within one of each other), and
- The optimum size $k$ of the core is $k \approx \sqrt{n}/3$. 

*Image source: OFR analysis*
Network complexity attributable to node \( v \)

Systemic importance designation depends on “interconnectedness”
- Does/should it depend on characteristics of the network?

Counterfactual thought experiment
- What is the network without \( v \)?

Depends on:
- Which nodes would rebalance?
- How they would rebalance?
- Which complexity measure to use?

\[
K_{\text{marg}}(\text{node 1}, K_{b_1\Lambda}) = K_{b_1\Lambda}(\Delta) - K_{b_1\Lambda}(\Delta')
\]

Image source: OFR analysis
Litigation complex, $\Lambda_v$

- A “local” line graph to capture the complexity of node removal
- Delete node $v$
- Completely connect its former neighbors

**Theorem 1.** If $\Gamma$ is a graph, $v$ a vertex of $\Gamma$, and $\Lambda_v$ the associated litigation complex, then

\[ (1) \quad b_0(\Lambda_v) = b_0(\Gamma), \quad \text{and} \]
\[ (2) \quad b_1(\Lambda_v) = b_1(\Gamma) + \frac{1}{2}(d_v - 1)(d_v - 2) = b_1(\Gamma) + \binom{d_v - 1}{2}, \quad \text{where} \ d_v \ \text{is the degree of} \ v \ \text{in} \ \Gamma. \]
DTCC data on credit default swaps (CDSs)

- Transaction-level data since 2003 on U.S. trades
- Weekly position data since 2010

London Whale: JPMC 2012

- U.S. Senate report, 2013
- JPMorgan Chief Investment Office (CIO) takes a large speculative bet
  - Net notional of $157 billion
  - Jamie Dimon indicates, the strategy was (FT, 2012):
    
    "flawed, complex, poorly reviewed, poorly executed, and poorly monitored"

  - Total loss of $6.2 billion
Markit CDS indexes: London Whale core trades

- CDX.NA.IG = North American, Investment Grade
- CDX.NA.IG.9 made the headlines, but over 100 other CDS types involved

CDX.NA.IG.[8-15]
All positions, one week in 2011

Image source: DTCC data, OFR analysis
Interdealer versus customer trades

Distribution of Node Degrees
Example for one week in 2012

- **Number of nodes**
- **Average interdealer degree**

**Core-periphery topology**

**Interdealer versus customer trades**

Image source: DTCC data, OFR analysis
Cycle rank complexity, $K_{b1\Delta}(G)$
CDX.NA.IG.[8-23], 2010-2014

Image source: DTCC data, OFR analysis
Summary Statistics – London Whale

All index contracts: Markit CDX.NA.IG.[8-23]

- Jan 2010 – Dec 2014
- Correlations across complexity measures are very high

<table>
<thead>
<tr>
<th></th>
<th>K_v</th>
<th>K_e</th>
<th>K_b1D</th>
<th>K_c3</th>
<th>K_Lb1D</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Mean</strong></td>
<td>332.8</td>
<td>2567.5</td>
<td>2235.7</td>
<td>304236.9</td>
<td>430500.9</td>
</tr>
<tr>
<td><strong>Std Dev</strong></td>
<td>31.3</td>
<td>236.1</td>
<td>215.0</td>
<td>80217.1</td>
<td>80006.5</td>
</tr>
<tr>
<td><strong>K_v</strong></td>
<td>1</td>
<td>0.712</td>
<td>0.637</td>
<td>0.714</td>
<td>0.831</td>
</tr>
<tr>
<td><strong>K_e</strong></td>
<td>0.712</td>
<td>1</td>
<td>0.995</td>
<td>0.957</td>
<td>0.950</td>
</tr>
<tr>
<td><strong>K_b1D</strong></td>
<td>0.637</td>
<td>0.995</td>
<td>1</td>
<td>0.947</td>
<td>0.923</td>
</tr>
<tr>
<td><strong>K_c3</strong></td>
<td>0.714</td>
<td>0.957</td>
<td>0.947</td>
<td>1</td>
<td>0.977</td>
</tr>
<tr>
<td><strong>K_Lb1D</strong></td>
<td>0.831</td>
<td>0.950</td>
<td>0.923</td>
<td>0.977</td>
<td>1</td>
</tr>
</tbody>
</table>

*Image source: DTCC data, OFR analysis*
Linear dependence in complexity measures

Core-periphery graphs are special
• Core approaches complete graph
• Core dominates the complexity measures
• New edges in the core tend strongly to create cycles
• New edges in the periphery almost never create cycles

Example comparison
• Two graphs of identical size, large complexity difference
  • $K_v(G) = 90$ and $K_e(G) = 180$
• Core-periphery graph, $G_{CP}$: 15 core nodes, each with 5 clients
  • Triangle count, $K_{C3}(G_{CP}) = 455$
• Random graph, $G_R$: edges distributed arbitrarily
  • Triangle count, $K_{C3}(G_R) \approx 10$
Directions for future research

• Correlations
  • Market stress events
  • Systemic risk measures
  • Market liquidity measures

• Modeling extensions
  • Directed graphs
  • Weighted graphs
  • Stratification by collateral type
  • Persistent homology

• Market extensions
  • Transaction networks
  • Other markets

Image sources: Library of Congress, OFR analysis
Thanks!