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

Disclaimer

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

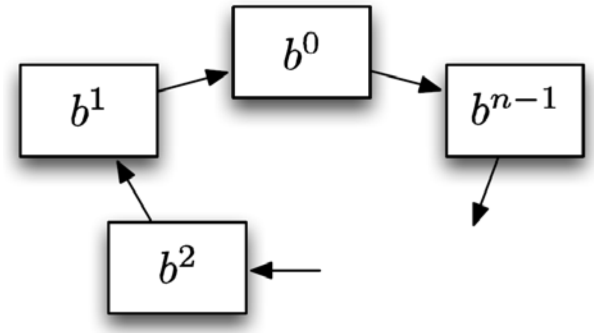
- **Defining "complexity" as an emergent phenomenon – Mitchell (2009)**
  - Size
  - Entropy
  - Algorithmic information content
  - Logical depth
  - Thermodynamic depth
  - Statistical complexity
  - Fractal dimension
  - Degree of hierarchy

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

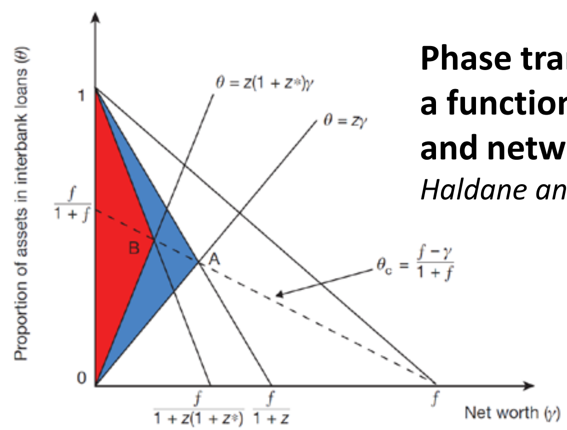
Image source: NASA

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)



Dealer myopia in a "circle" network
Caballero and Simsek (2009, 2013)

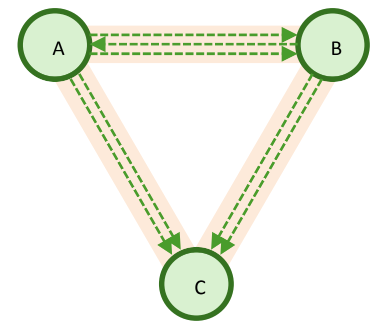


Phase transitions in systemic risk, a function of bank capitalization and network exposure
Haldane and May (2011)

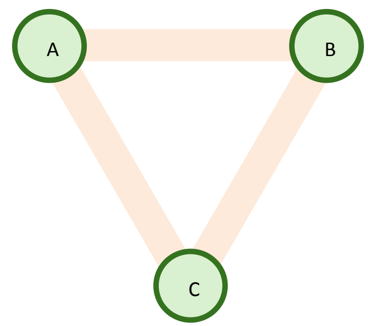
Image sources: Caballero and Simsek (2009), Haldane and May (2011)

Requirements for complexity measurement

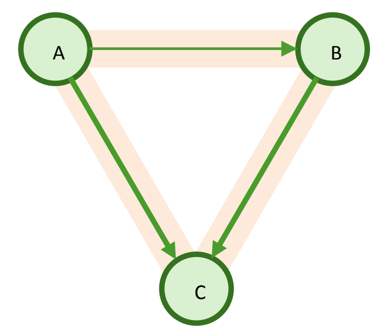
- **Focus on financial markets**
  - Transactions
  - Contractual exposures
  - Trading relationships
- **Statistics at the network-level**
  - Comparison across markets
  - Cardinal (numeric) metrics
- **Financially meaningful**
  - Sensitive to local interactions
  - Intuitive interpretation re: market institutions/practices
  - Capture multifaceted complexity



Transactions



Master agreements

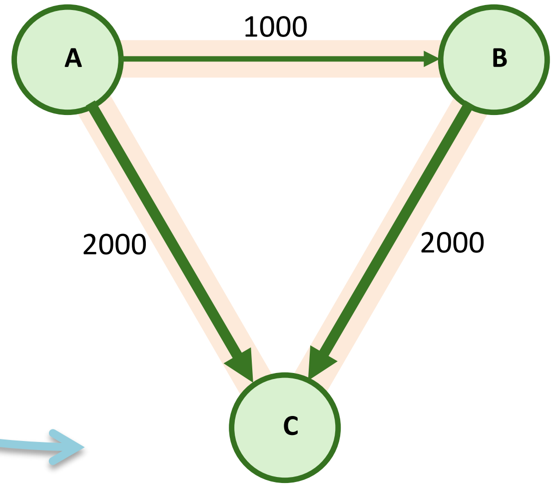


Positions

Image source: OFR analysis

Mathematical abstraction

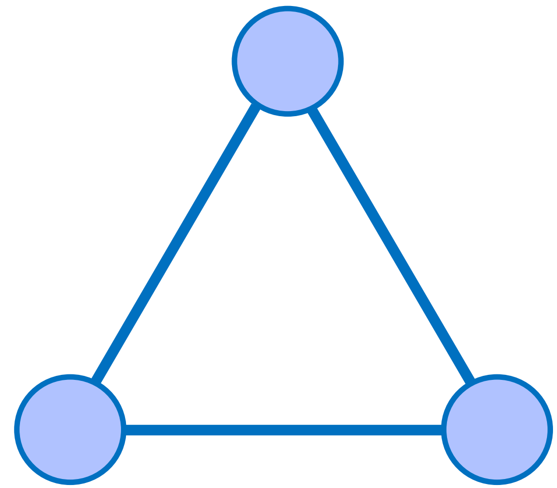
Market activity



Preparing for topological analysis

Measurement

Counterparty network data



We will work with the undirected, unweighted, market graph, Δ

Modeling

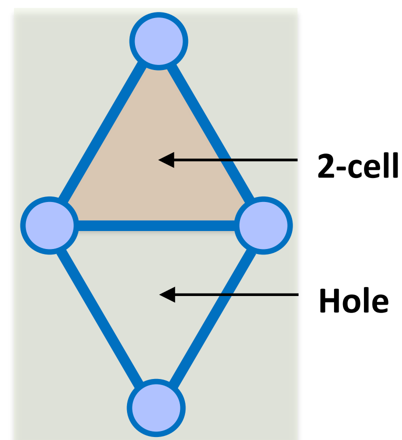
Counterparty graph model

Image sources: Library of Congress, OFR analysis

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…

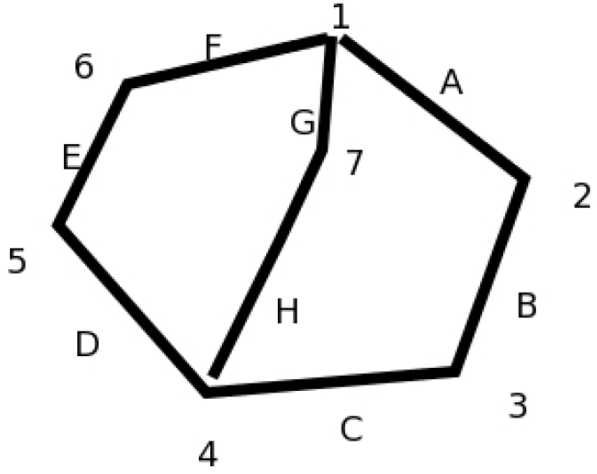


- **Boundaries**
  - Vertices (0-*cells*) have no boundary
  - Edges, non-closed paths (1-*cells*) have endpoints
  - Polygons (other 2-*cells*) have perimeters
  - etc…

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

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**



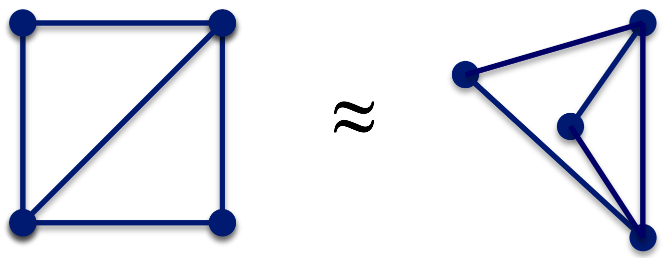


Image source: OFR analysis

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:**

$\langle ABCHG \rangle + \langle DEFGH \rangle =$

$1\langle ABC \rangle + 1\langle H \rangle + 1\langle G \rangle + 1\langle DEF \rangle + 1\langle G \rangle + 1\langle H \rangle =$

$1\langle ABC \rangle + 1\langle DEF \rangle + 1\langle H \rangle + 1\langle H \rangle + 1\langle G \rangle + 1\langle G \rangle =$

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

$1\langle ABC \rangle + 1\langle DEF \rangle =$

$\langle ABCDEF \rangle$

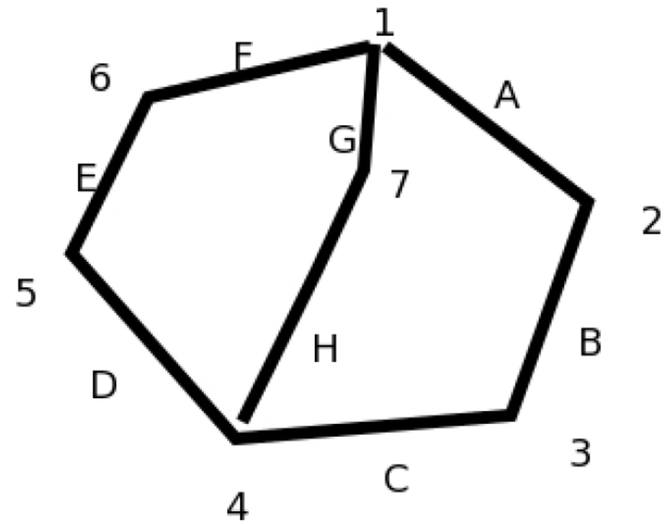


Image source: OFR analysis

Homology

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

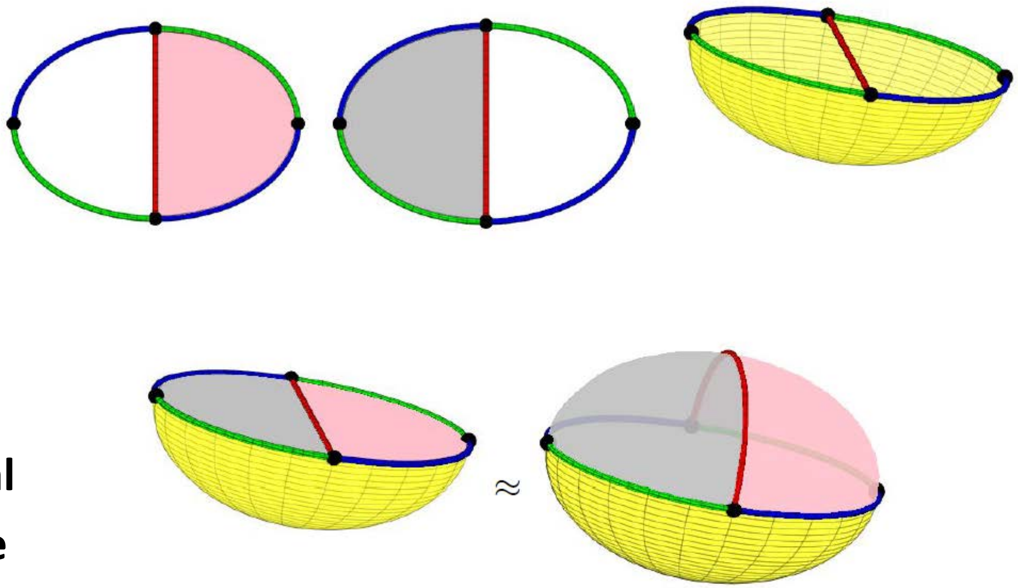
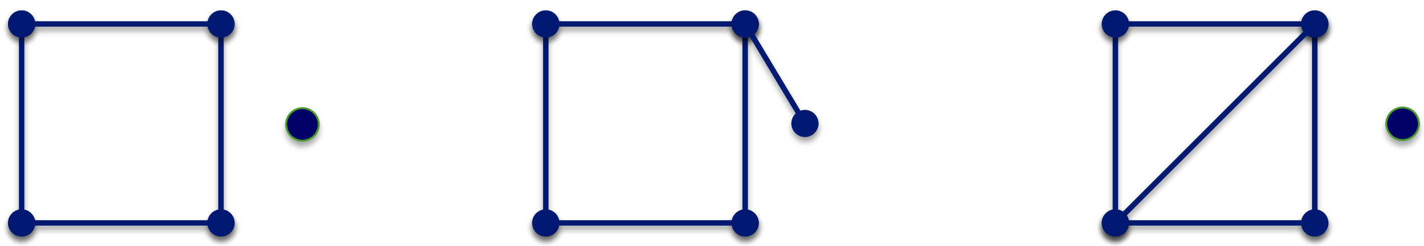


Image source: OFR analysis

Key Theoretical Result

Euler-Poincaré formula

- **A new edge <u>must</u> connect two components, <u>or</u> 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**



$v(\Delta) = 5$	$v(\Delta) = 5$	$v(\Delta) = 5$
$e(\Delta) = 4$	$e(\Delta) = 5$	$e(\Delta) = 5$
$b_0(\Delta) = 2$	$b_0(\Delta) = 1$	$b_0(\Delta) = 2$
$b_1(\Delta) = 1$	$b_1(\Delta) = 1$	$b_1(\Delta) = 2$

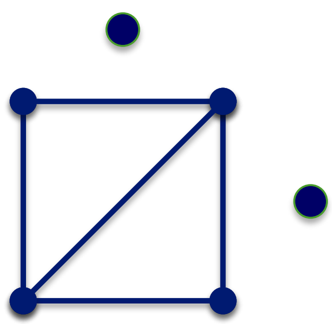
Image source: OFR analysis

Complexity Metrics I

Aggregate size metrics

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

Insensitive to network structure:



$K_v = 6$

$K_e = 5$

$K_{e/v} = 0.8333$

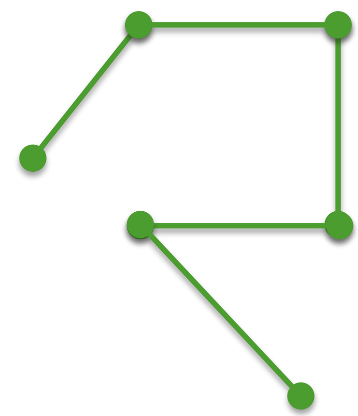


Image source: OFR analysis

OFR

Counting simple cycles – closed-form calculations

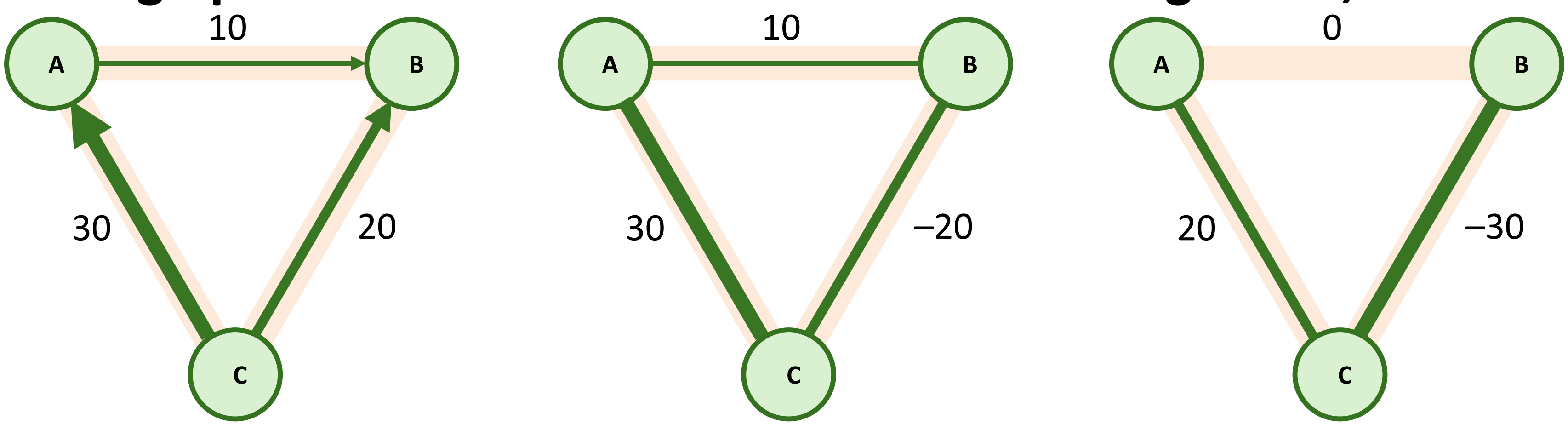
- $K_{C3} \equiv$ number of triangular cycles in the market graph, Δ
- $K_{C4} \equiv$ number of quadrilateral cycles
- $K_{C5} \equiv$ number of pentagonal cycles
- $K_{Cnet} \equiv$ *"total"* number of nettable cycles = $K_{C3} + K_{C4} + K_{C5}$

Zeroth and first homology on the market graph

- $K_{b_1\Delta} \equiv$ rank $H_1(\Delta) = b_1(\Delta)$ = cycle rank of vector space $H_1(\Delta)$
- $K_{b_0\Delta} \equiv$ rank $H_0(\Delta) = b_0(\Delta)$ = number connected components in Δ

Netting

Netting operates on the directed network of obligations, $\vec{\Delta}$



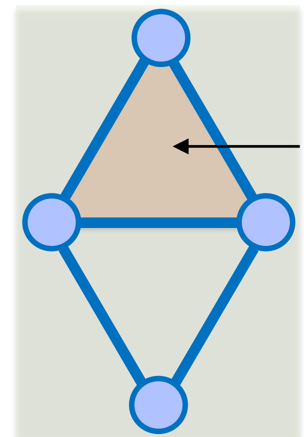
- Oriented cycles in the digraph, $\vec{\Delta}$, are potentially "nettable"
- Each oriented ("nettable") cycle in $\vec{\Delta}$ has a "shadow" in Δ
- Netting "kills" a cycle in $\vec{\Delta}$ by reducing edge weights so one is zero
  - Bookkeeping convention: replace orientations in $\vec{\Delta}$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

Image source: OFR analysis

Complexity Metrics III

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 Δ to a 2-dimensional cell complex
- Call the resulting structure: Δ^2_{net}



Netting measures on Δ^2_{net}

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

Image source: OFR analysis

Order of netting matters

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

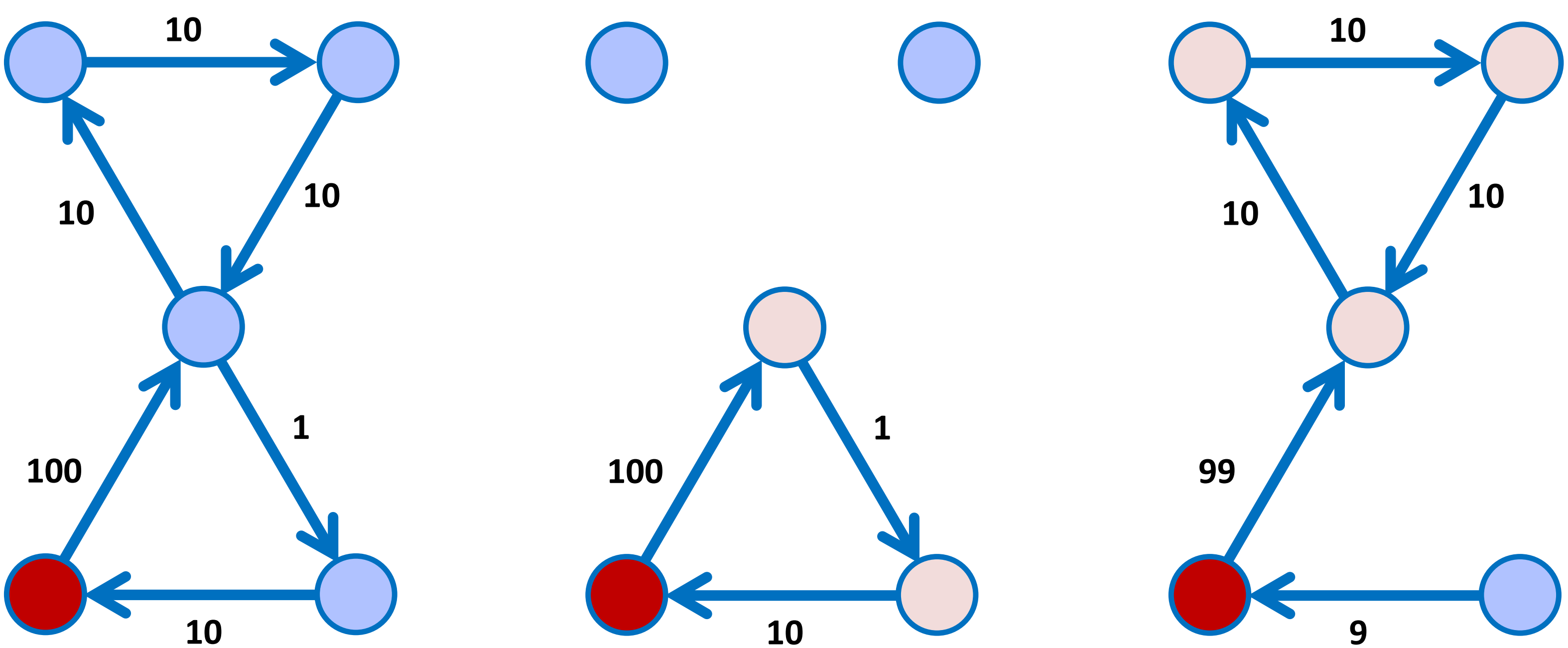
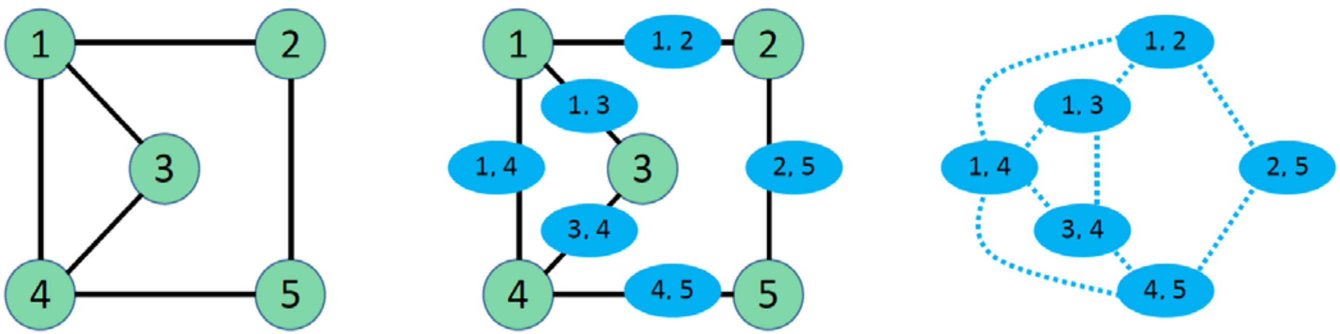


Image source: OFR analysis

Every undirected graph, Δ, has a matching line graph, Λ_Δ

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



Original graph, Δ

Line graph, Λ_Δ

- Connect nodes in Λ_Δ if the corresponding edges in Δ share a node
- Thus, a highly central node in Δ implies a busy line graph, Λ_Δ

Image source: OFR analysis

Complexity Metrics IV

OFR

Simple cycles in the line graph, Λ_Δ

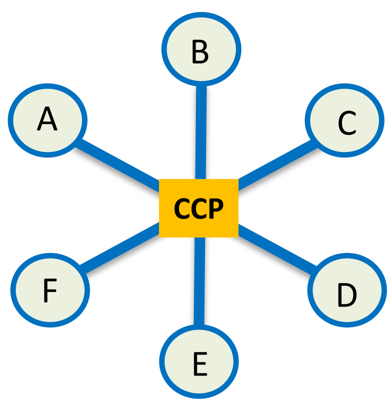
- $K_{C3\Lambda}$ ≡ number of triangular cycles in Λ_Δ
- $K_{C4\Lambda}$ ≡ number of quadrilateral cycles
- $K_{C5\Lambda}$ ≡ number of pentagonal cycles

Exploiting the line graph, Λ_Δ

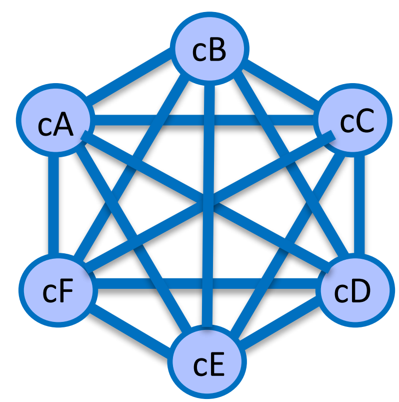
- $K_{e\Lambda} \equiv e(\Lambda_\Delta)$ = edge count in Λ_Δ = deal interactions
- $K_{b_1\Lambda} \equiv b_1(\Lambda_\Delta) = \text{rank } H_1(\Lambda_\Delta)$ = cycle rank of Λ_Δ

Example: Central Counterparties (CCPs)

- Λ_Δ contains no new information …
- … but nonetheless reveals facts not obvious in Δ
- Sensitive to the presence of highly central nodes
- CCPs as single point of failure:



CCP market graph



CCP line graph

Image source: OFR analysis

Evaluating the Metrics

Does cycle rank detect chains?

- **100K random graphs, except some *seeded* w/chains**
- **Each graph has *v*(Δ)=12 vertices and *e*(Δ)=15**
- **Cycle counts/ranks respond to the chains**

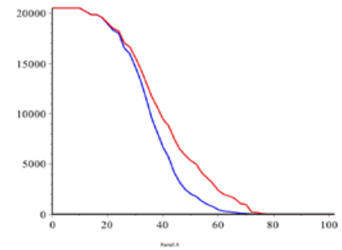
Cycle counts, $C_n(\Delta)$, and cycle ranks, $b_1(\Delta)$, on unseeded and seeded graphs

	$\mathbf{C}_3(\Delta)$	$\mathbf{C}_4(\Delta)$	$\mathbf{C}_5(\Delta)$	$b_1(\Delta)$	$e(\Lambda_\Delta)$	$\mathbf{C}_3(\Lambda_\Delta)$	$\mathbf{C}_4(\Lambda_\Delta)$	$\mathbf{C}_5(\Lambda_\Delta)$	$b_1(\Lambda_\Delta)$
Unbiased	2.19	2.81	3.20	4.58	32.32	21.90	36.65	77.71	18.37
two 4-chains	2.44	3.46	4.09	4.86	33.84	24.42	43.79	99.66	19.89
three 4-chains	2.58	3.81	4.64	5.04	34.67	25.83	48.11	113.43	20.71
three 3-chains	2.59	3.64	4.16	4.89	34.43	25.92	49.12	115.96	20.48
four 3-chains	2.72	3.94	4.54	5.02	35.14	27.23	53.38	129.92	21.18

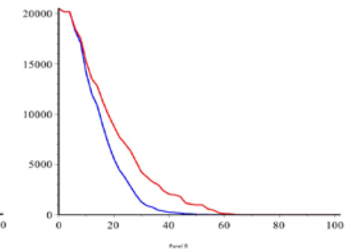
Image source: OFR analysis

Weighted graph stratification and persistent homology

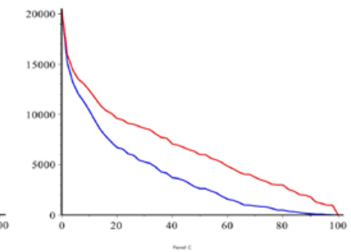
- Exposure concentration – random (blue) vs concentration bias (red)
- Distribution of edge weights – 3 beta distributions
  - Unimodal (α=2.0, β=2.0)
  - Skewed (α=2.0, β=5.0)
  - Bimodal (α=0.5, β=0.5)



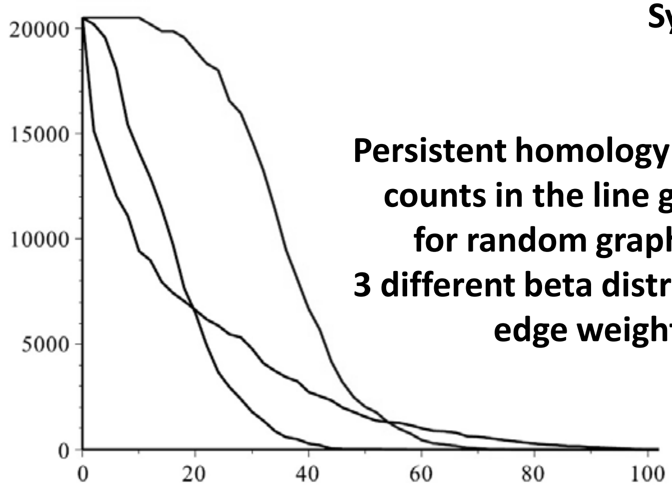
Symmetric unimodal



Skewed right



Symmetric bimodal



Persistent homology of triangle counts in the line graph, Λ_{Δ}, 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 = number of nodes
- $K_{e/v}$ = 10 = average degree
- $K_{b_0\Delta}$ is constant
- $K_{b_1\Delta}$ is constant
- $K_{e\Lambda}$ is constant

➔ Conditionally independent variation

Distribution of K_{C3}

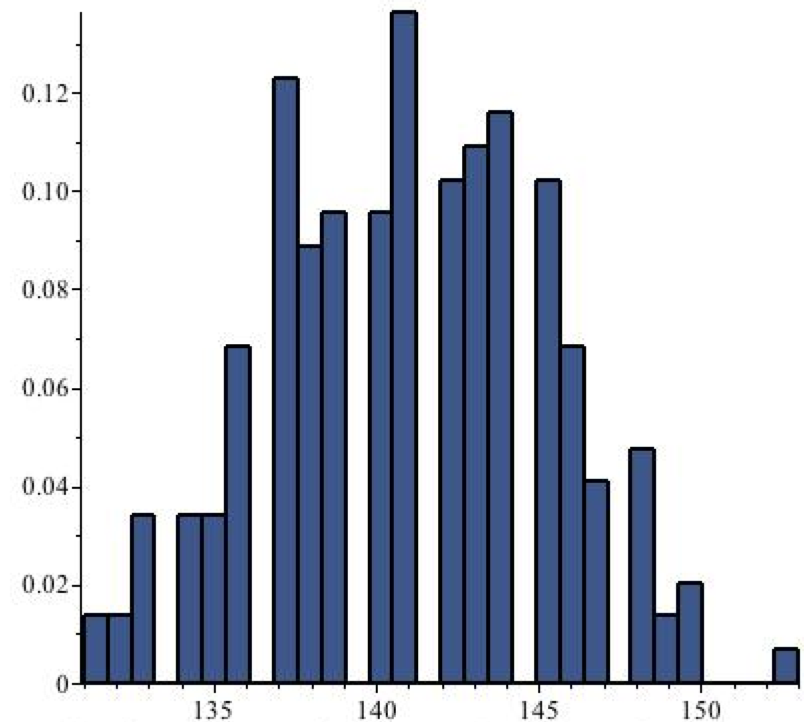
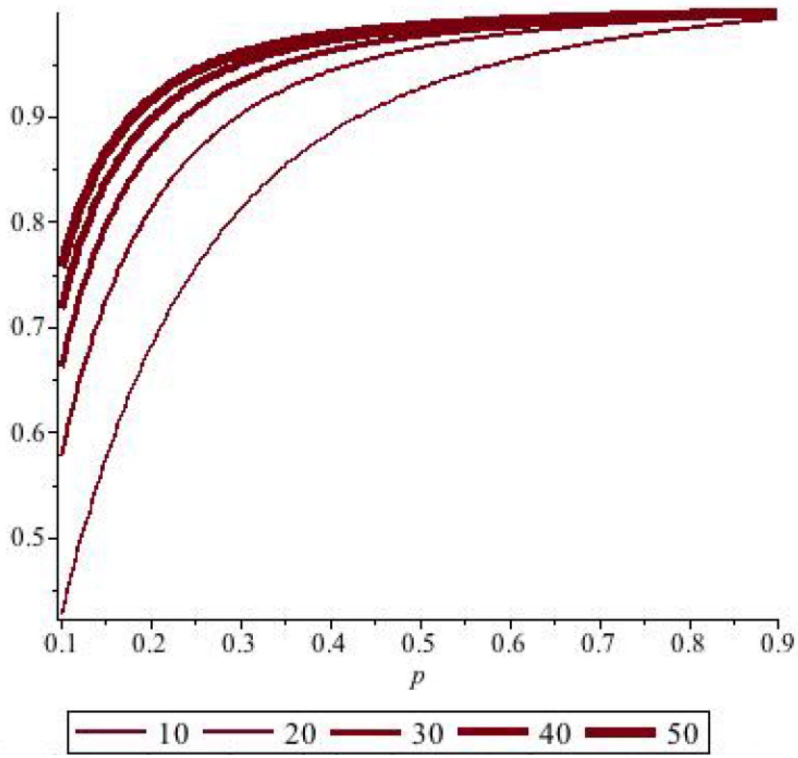


Image source: OFR analysis

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

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



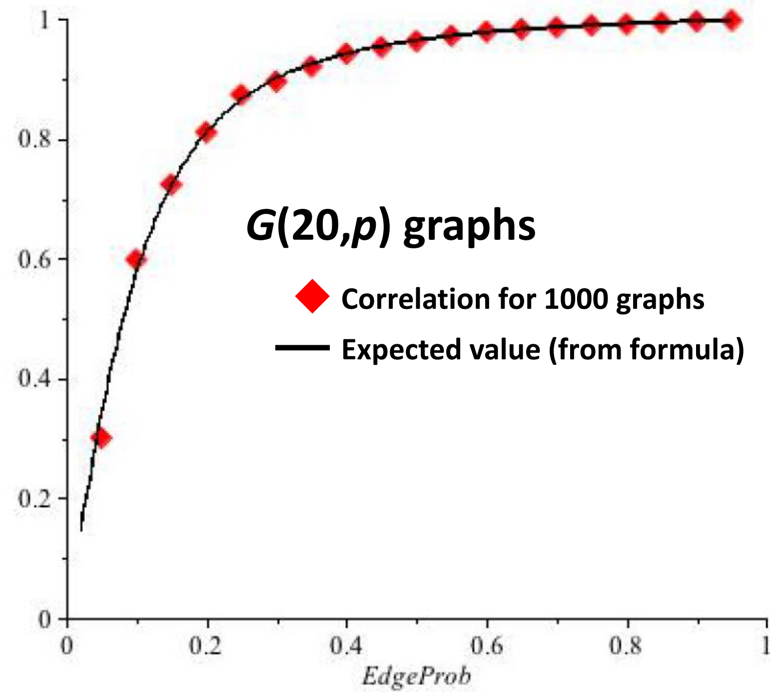
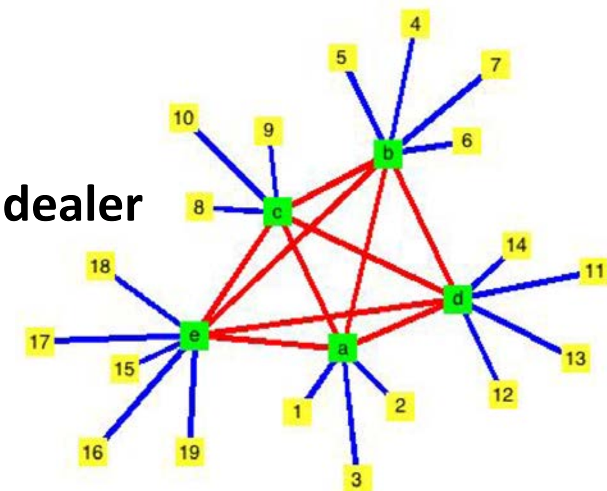


Image source: OFR analysis

Core-Periphery Theorem

"Simple" core-periphery graph , Δ^*_{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, Δ^*_{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?**

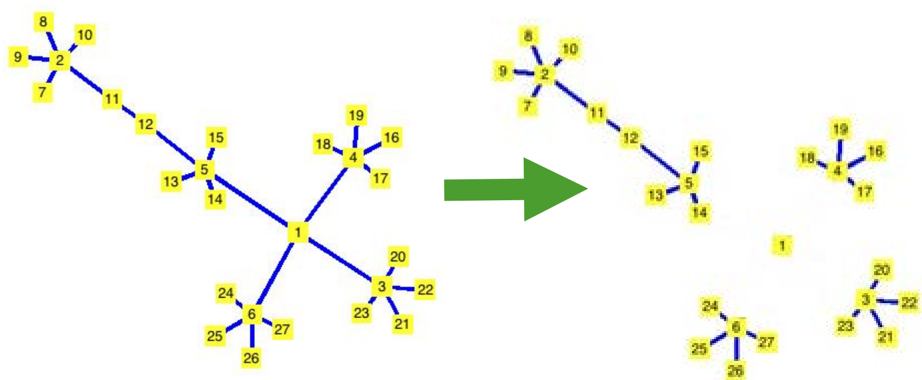
Counterfactual thought experiment

- **What is the network without *v*?**

Depends on:

- **Which nodes would rebalance?**
- **How they would rebalance?**
- **Which complexity measure to use?**

Effect of isolating node #1



$K_{b_1\Lambda}(\Delta)$ $K_{b_1\Lambda}(\Delta')$

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

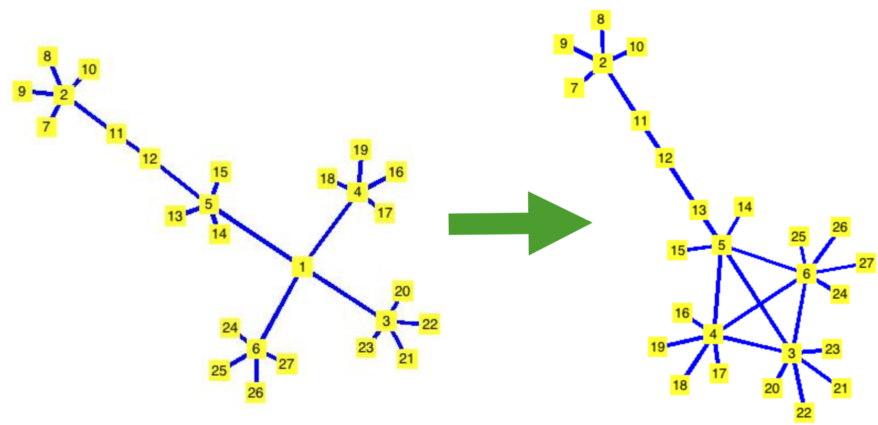
Image source: OFR analysis

Litigation Complex Theorem

Litigation complex, Λ_v

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

Effect of removing node #1



Theorem 1. *If Γ is a graph, v a vertex of Γ, and Λ_v the associated litigation complex, then*

(1) $b_0(\Lambda_v) = b_0(\Gamma)$, *and*

(2) $b_1(\Lambda_v) = b_1(\Gamma) + \frac{1}{2}(d_v - 1)(d_v - 2) = b_1(\Gamma) + \binom{d_v - 1}{2}$, *where d_v is the degree of v in Γ.*

Image source: OFR analysis

Practical Example – London Whale

OFR

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

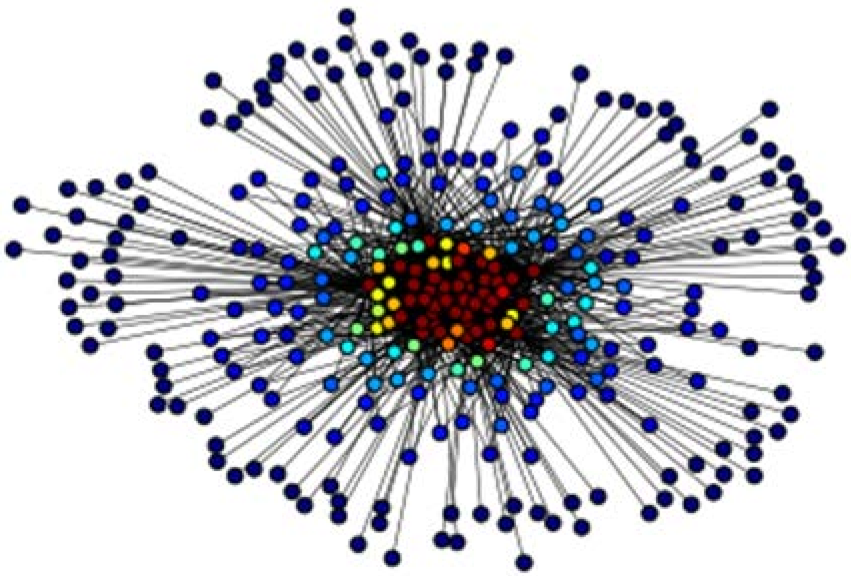


Image source: DTCC data, OFR analysis

Interdealer versus customer trades

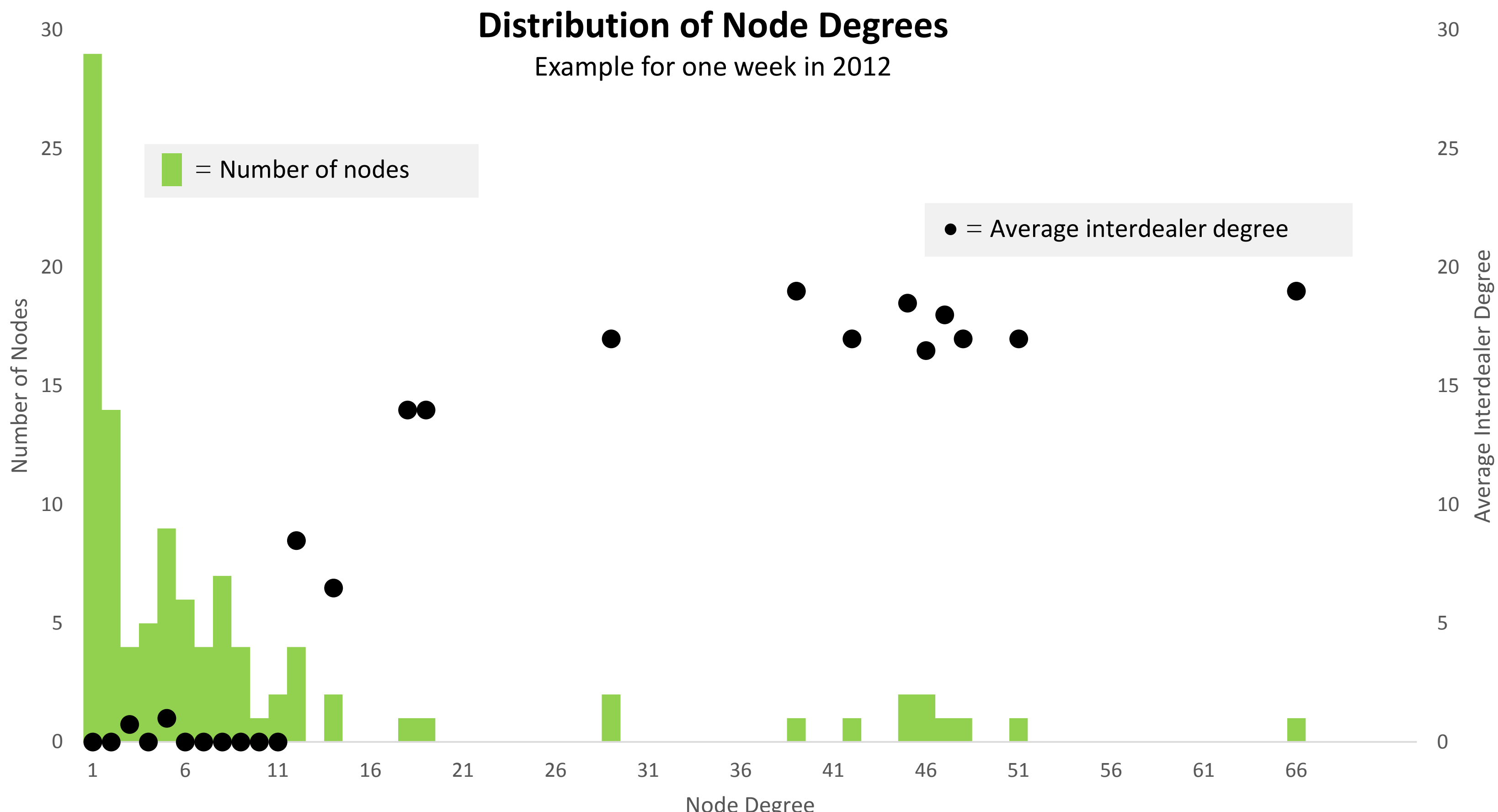


Image source: DTCC data, OFR analysis

Practical Example – London Whale

Cycle rank complexity, $K_{b1\Delta}(G)$

CDX.NA.IG.[8-23], 2010-2014

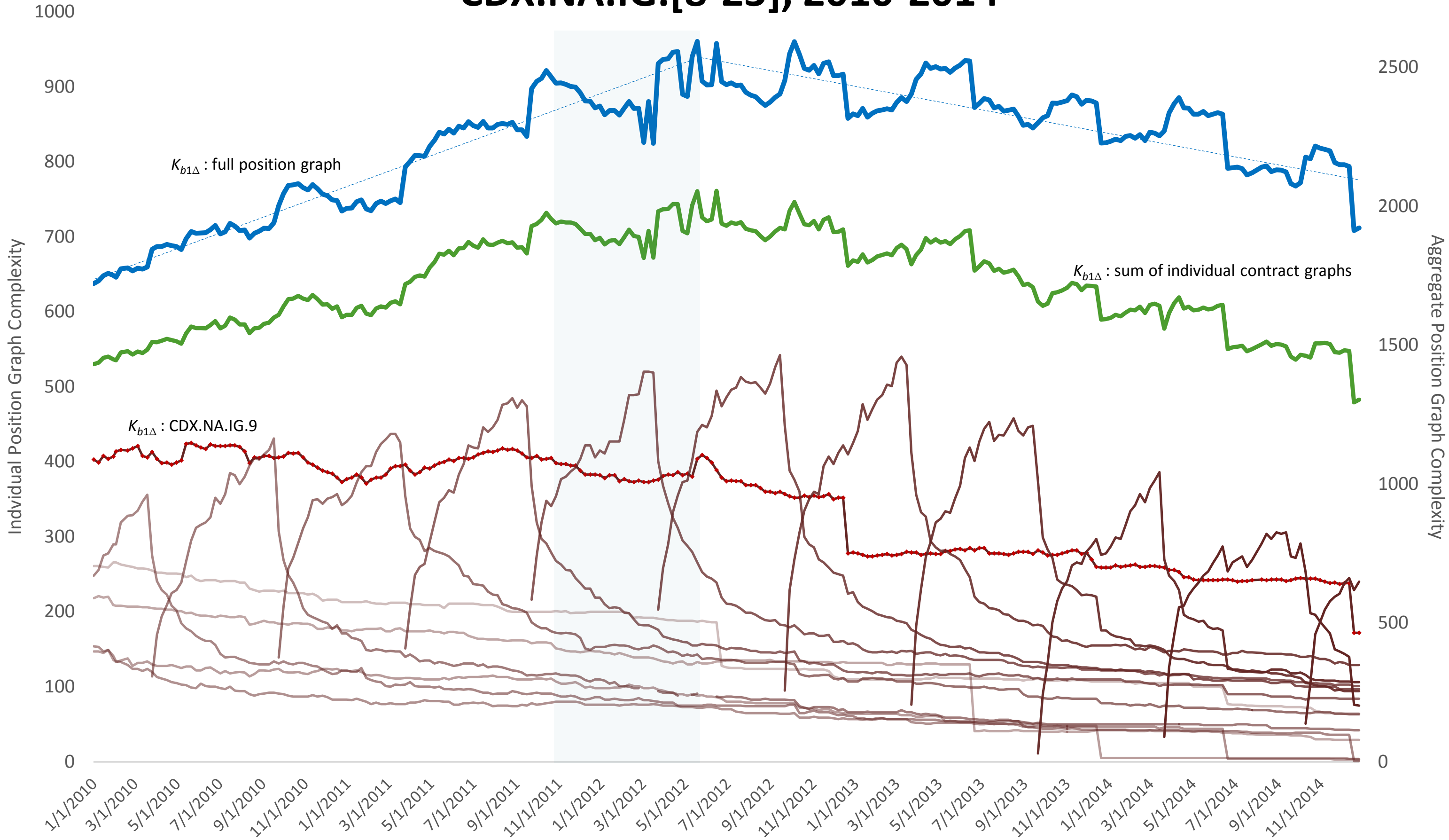


Image source: DTCC data, OFR analysis

OFR

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

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

	K_v	K_e	K_b_1D	K_c3	K_Lb_1D
Mean	332.8	2567.5	2235.7	304236.9	430500.9
Std Dev	31.3	236.1	215.0	80217.1	80006.5
K_v	1	0.712	0.637	0.714	0.831
K_e	0.712	1	0.995	0.957	0.950
K_b_1D	0.637	0.995	1	0.947	0.923
K_c3	0.714	0.957	0.947	1	0.977
K_Lb_1D	0.831	0.950	0.923	0.977	1

Image source: DTCC data, OFR analysis

Linear dependence in complexity measures

Core-periphery graphs are special

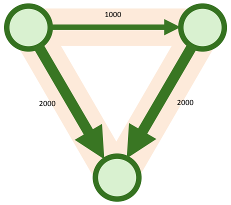
- Core approaches complete graph
- Core dominates the complexity measures
- New edges in the core tend strongly to create cycles
- New edges in the periphery almost never create cycles

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



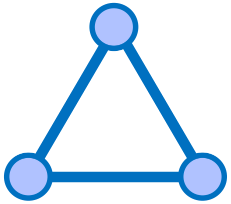


Image sources: Library of Congress, OFR analysis

Gratitude

Thanks!