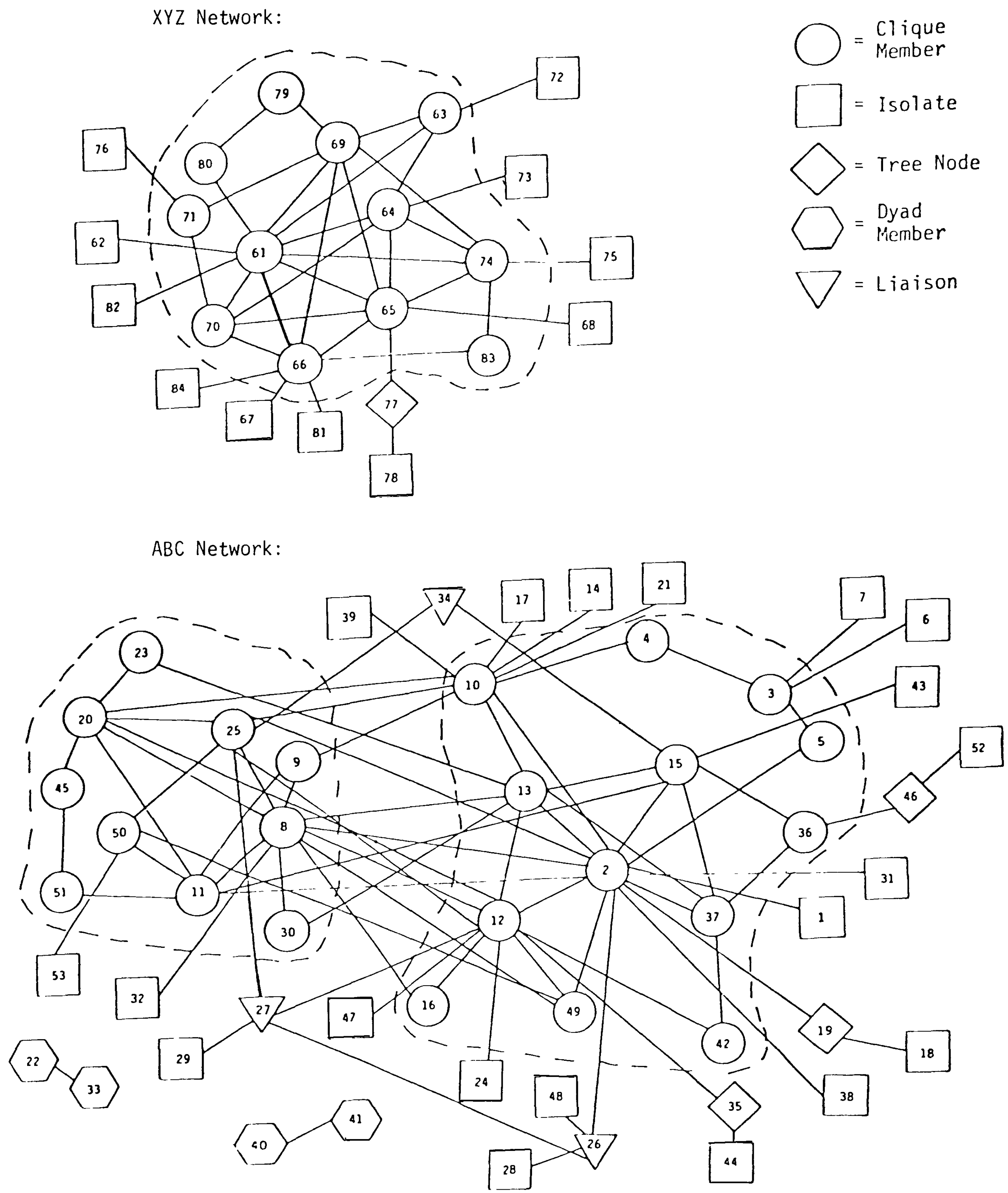
Visualizing Labor Migration Using Quantitative Data

Introduction to Network Analysis

Let's Update Our Files!

- Open RStudio
- Switch to the terminal window
- Type the following:
  cd "/YOUR/GITPATH/HERE"
- Type the following:
  git pull origin/master



Baker, Wayne E. 1984. "The Social Structure of a National Securities Market." *American Journal of Sociology* 89: 775–811.

ESTIMATES OF CHANGES IN GNP/CAPITA (1955–70) REGRESSED ON CONTROL VARIABLES AND DUMMY VARIABLES REPRESENTING "STRUCTURAL POSITION" (Location in 10-Block Model)

INDEPENDENT VARIABLES[a]	PARTIAL REGRESSION COEFFICIENTS (Standard Errors)			
	Eq. (1)	Eq. (2)	Eq. (3)	Eq. (4)
GNP/capita (1955)	.292 (.080)**	.274 (.081)**	.266 (.082)**	.294 (.082)**
Secondary school enrollment ratio (1955)	20.99 (2.37)**	13.62 (2.95)**	16.98 (2.76)**	11.18 (3.15)**
Block A	...	−468.0 (170.4)**	...	−329.0 (163.1)*
Block B	...	−507.4 (184.7)**	...	−449.6 (182.1)*
Block C′	...	−645.4 (186.0)**	...	−626.2 (184.1)**
Block D	...	−422.7 (139.7)**	...	−386.6 (146.3)**
Block D′	...	−510.0 (180.1)**	...	−490.5 (179.9)**
Block E	...	−542.8 (202.6)**	..	−504.3 (197.9)*
Block E′	...	−582.3 (177.5)**	...	−546.5 (175.9)**
Block F	...	−310.0 (153.4)*	..	−257.3 (158.9)
Block F′	...	−789.5 (204.3)**	..	−758.3 (204.7)**
Diplomats sent (1963–64)	...	...	.142 (.297)	.433 (.309)
Interventions into (1960–67)	...	...	−1.03 (3.54)	−3.44 (3.59)
Interventions by (1960–67)	...	...	−11.47 (4.42)*	−12.68 (4.49)**
Value of exports (1965; millions of U.S. $)	...	...	.036 (.018)*	.010 (.019)
Regression constant	20.33	562.2	51.1	507.6
R^2	68.1%	75.9%	72.7%	79.1%

NOTE.—N = 90 nations.
[a] Dependent variable is change in GNP/capita, 1955–70.
* Significant at .05.
** Significant at .01.

Snyder, David and Edward L. Kick. 1979. "Structural Position in the World System and Economic Growth, 1955–1970: A Multiple-Network Analysis of Transnational Interactions." *American Journal of Sociology* 84: 1096–1126.

Overview

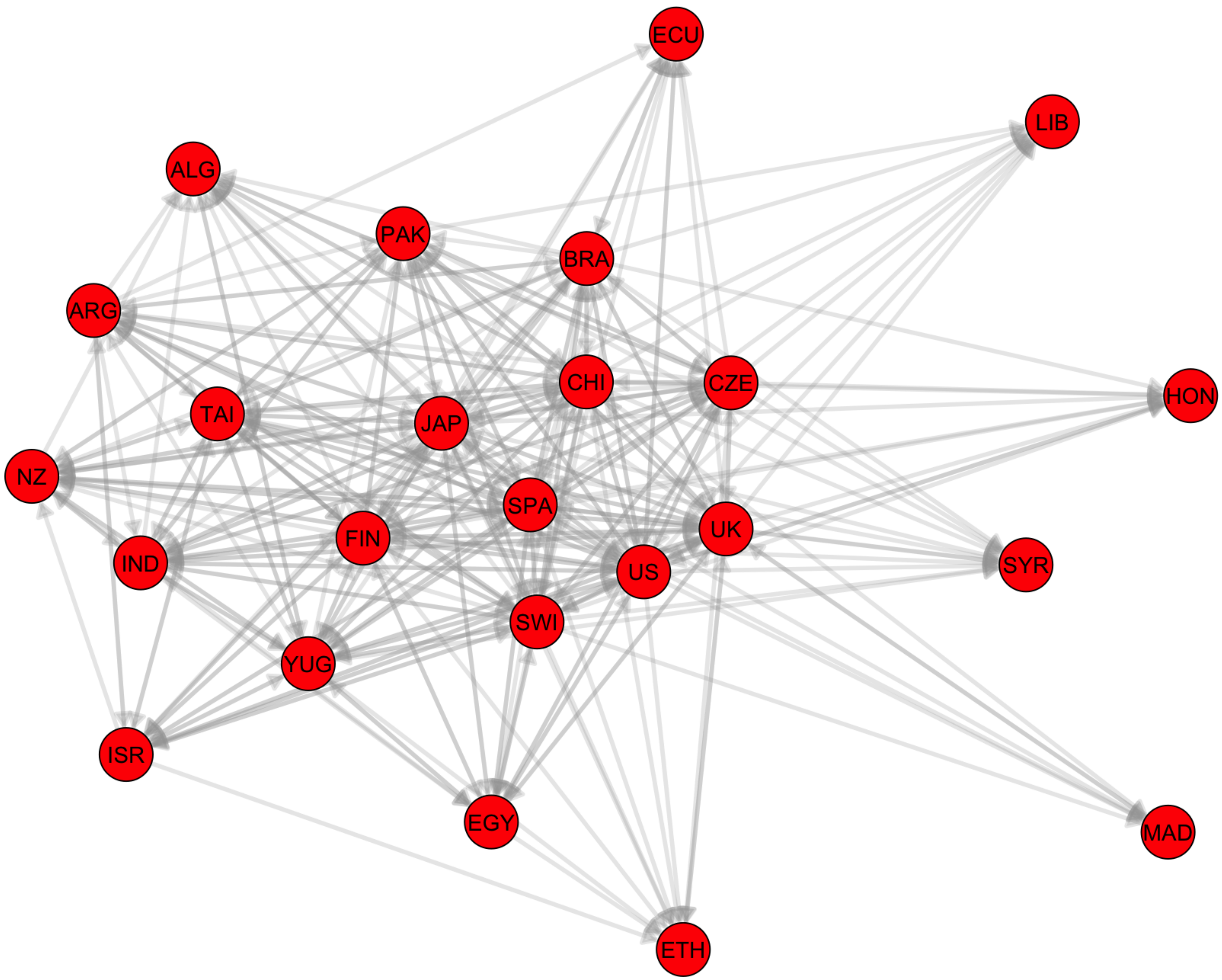
- Anatomy of a network
  - Types of networks/graphs
  - Working with network data
- Analysis
  - Graph properties
  - Centrality
  - Roles and positions

Anatomy of a Network

- Vertices/Nodes
  - Represent one or types of entities
    - One- vs. two-mode data
- Edges
  - Represent one or more relationships between entities
  - Edges come in different types
    - Unweighted vs. weighted
    - Directed vs. undirected
    - Unsigned vs. unsigned

One-Mode Data

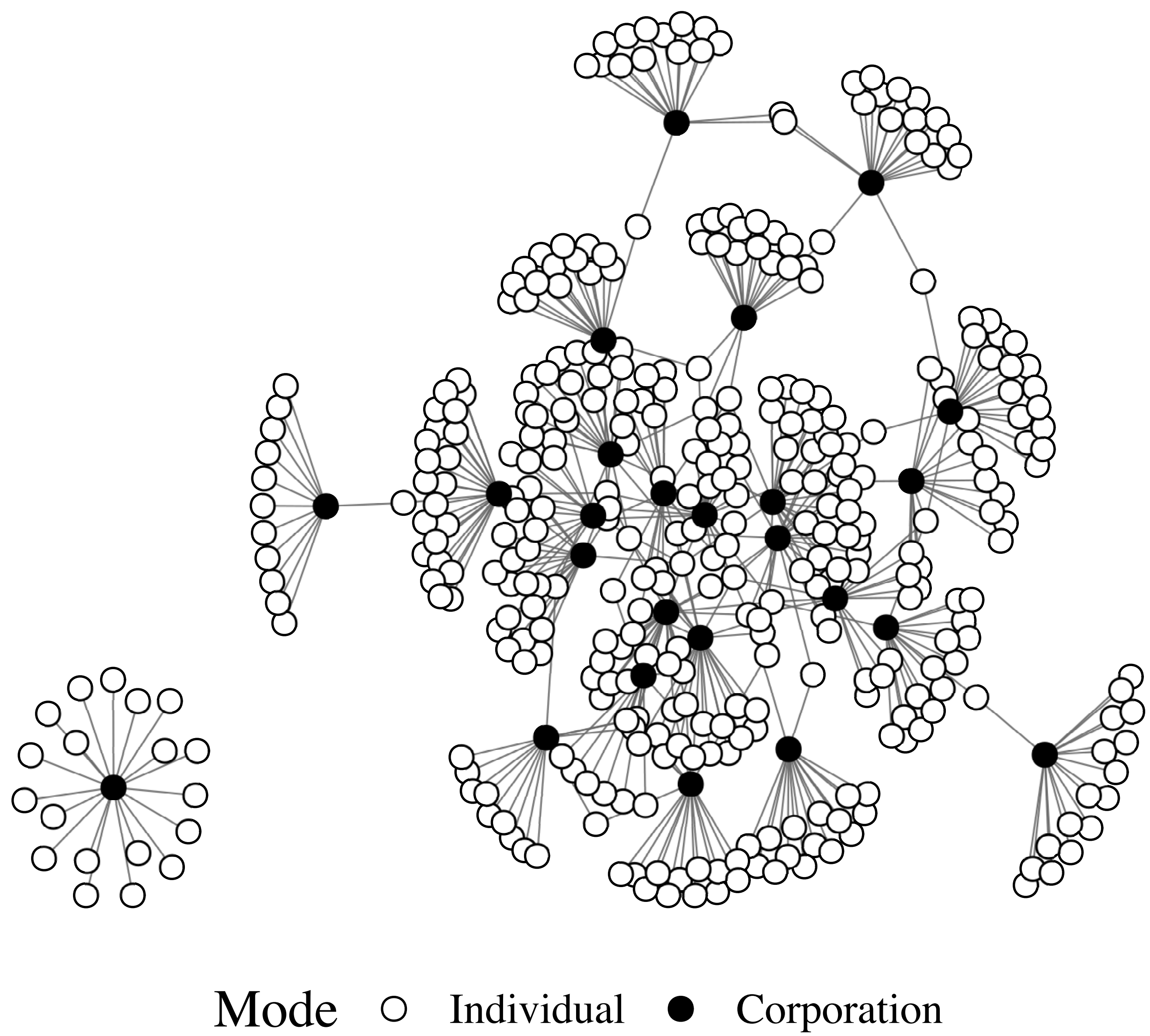
Imports of basic manufactured goods in 1984



Two-Mode Data

Graph of board interlocks among the twenty-five largest railroad corporations in the United States in 1905

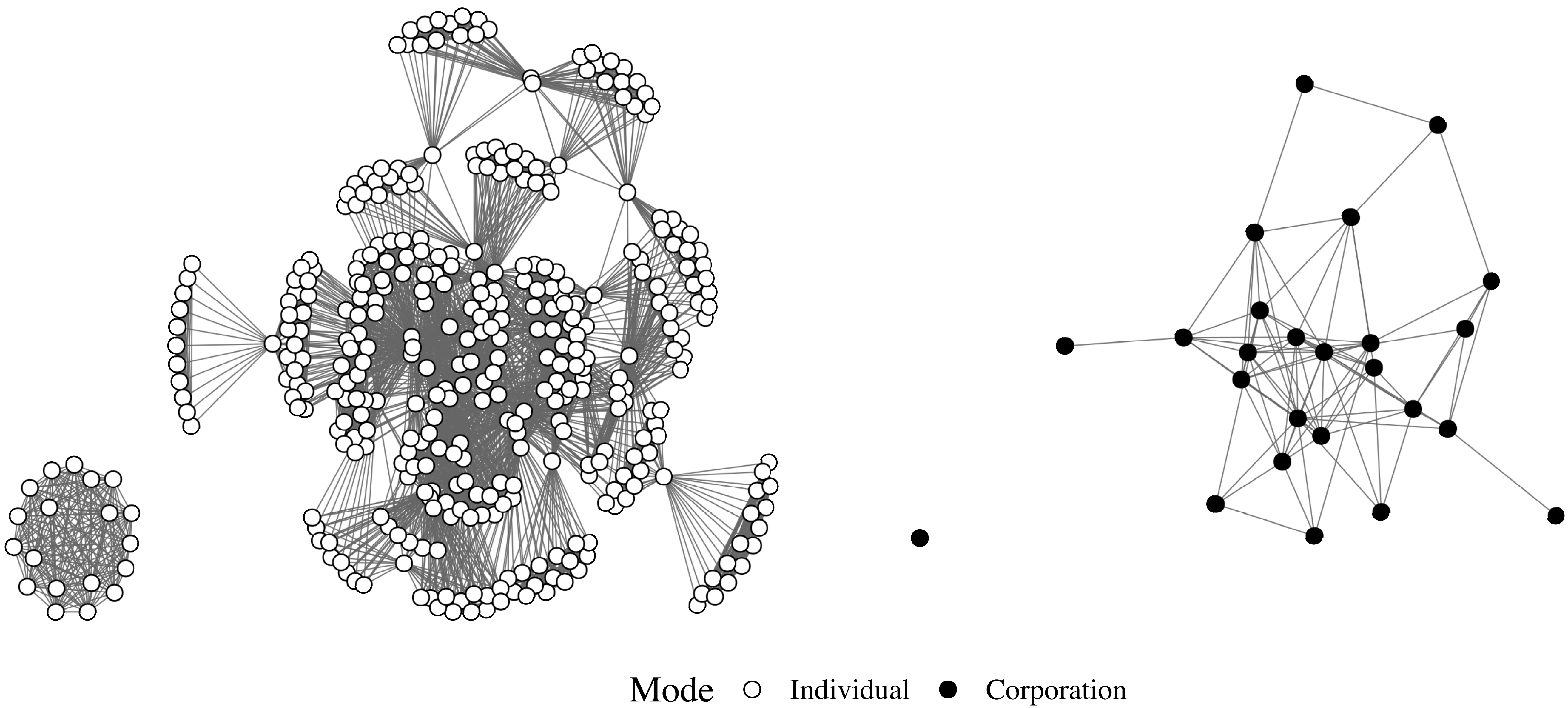
(A) Individual-to-Corporation



Projecting Two-Mode Data

Graph of board interlocks among the twenty-five largest railroad corporations in the United States in 1905

(B) Individual-to-Individual **(C) Corporation-to-Corporation**



Anatomy of a Network

- Vertices/Nodes
  - Represent one or types of entities
    - One- vs. two-mode data
- Edges
  - Represent one or more relationships between entities
  - Edges come in different types
    - Unweighted vs. weighted
    - Directed vs. undirected
    - Unsigned vs. unsigned

Overview

- Anatomy of a network
  - Types of networks/graphs
  - Working with network data
- Analysis
  - Graph properties
  - Centrality
  - Roles and positions

Matrix Representation of an Unweighted and Undirected Graph

	Acciaiuoli	Albizzi	Barbadori	Bischeri	Castellani	Ginori	Guadagni	Lamberteschi	Medici	Pazzi	Peruzzi	Pucci	Ridolfi	Salviati	Strozzi	Tornabuoni
Acciaiuoli	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
Albizzi	0	0	0	0	0	1	1	0	1	0	0	0	0	0	0	0
Barbadori	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0
Bischeri	0	0	0	0	0	0	1	0	0	0	1	0	0	0	1	0
Castellani	0	0	1	0	0	0	0	0	0	0	1	0	0	0	1	0
Ginori	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Guadagni	0	1	0	1	0	0	0	1	0	0	0	0	0	0	0	1
Lamberteschi	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
Medici	1	1	1	0	0	0	0	0	0	0	0	0	1	1	0	1
Pazzi	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0
Peruzzi	0	0	0	1	1	0	0	0	0	0	0	0	0	0	1	0
Pucci	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Ridolfi	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	1
Salviati	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0
Strozzi	0	0	0	1	1	0	0	0	0	0	1	0	1	0	0	0
Tornabuoni	0	0	0	0	0	0	1	0	1	0	0	0	1	0	0	0

Matrix Representation of an Unweighted and Directed Graph

	ALG	ARG	BRA	CHI	CZE	ECU	EGY	ETH	FIN	HON	IND	ISR	JAP	LIB	MAD	NZ	PAK	SPA	SWI	SYR	TAI	UK	US	YUG
ALG	0	0	0	1	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1
ARG	1	0	1	1	0	1	0	0	1	0	1	1	1	0	0	0	1	1	1	0	1	0	1	0
BRA	1	1	0	1	1	1	1	0	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1
CHI	1	1	1	0	1	0	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1
CZE	1	1	1	1	0	1	1	1	1	1	1	0	1	1	0	1	1	1	1	1	1	1	1	1
ECU	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
EGY	0	0	0	0	1	0	0	1	1	0	0	0	1	0	0	0	0	1	1	0	0	1	1	1
ETH	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0
FIN	1	1	1	1	1	1	1	1	0	1	1	1	1	0	0	1	1	1	1	1	1	1	1	1
HON	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
IND	1	0	0	1	1	0	1	0	1	0	0	0	1	0	0	1	1	1	1	0	1	1	1	1
ISR	0	1	0	0	0	0	0	1	1	0	0	0	1	0	0	1	0	1	1	0	1	1	1	1
JAP	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1
LIB	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MAD	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0
NZ	1	0	0	1	0	0	1	0	0	0	1	0	1	0	0	0	1	1	0	0	1	1	1	1
PAK	0	0	0	1	1	0	0	0	1	0	1	0	1	1	0	1	0	1	1	1	1	1	1	0
SPA	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1
SWI	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1
SYR	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
TAI	0	0	1	1	0	0	0	0	1	0	1	1	1	0	0	1	1	1	1	1	0	1	1	1
UK	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1
US	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1
YUG	1	1	0	1	1	0	1	1	1	0	1	1	1	0	0	1	1	1	1	1	1	1	1	0

Matrix Representation of a Weighted and Directed Graph

	ARG	DZA	ECU	GBR	ISR	USA	JPN	CHE	ESP	HND
ARG	0	0	2183	10198	36951	133210	3061	1359	83803	139
DZA	0	0	17	40555	46734	11717	179	3959	23269	3
ECU	2035	0	0	9796	2018	302705	0	1157	176212	334
GBR	2383	185	743	0	21489	833858	13718	23360	86985	75
ISR	1432	1	182	7729	0	125896	562	1100	912	65
USA	10338	27	10824	165086	70135	0	50202	14729	17204	4946
JPN	5348	3	427	59546	74	500527	0	3591	2545	101
CHE	1820	62	669	5241	966	51605	306	0	43159	45
ESP	137084	1598	3030	60888	75	119977	1523	88369	0	383
HND	0	0	159	1	5	288779	127	152	2824	0

Edgelist Representation of a Weighted and Directed Graph

From	To	Count
ARG	ECU	2183
ARG	GBR	10198
ARG	ISR	36951
ARG	USA	133210
ARG	JPN	3061
ARG	CHE	1359
ARG	ESP	83803
HND	ECU	159
HND	GBR	1
HND	ISR	5
HND	USA	288779
HND	JPN	127
HND	CHE	152
HND	ESP	2824

Let’s find out how to do this all in R!

Overview

- Anatomy of a network
  - Types of networks/graphs
  - Working with network data
- Analysis
  - Graph properties
  - Centrality
  - Roles and positions

Describing Graphs

- Number of components
  - Number of subgraphs in which there is (a) a path between every pair of nodes in the subgraph and (b) no paths to nodes in other subgraphs
    - One component = connected
    - Two or more components = disconnected
- Diameter
  - Length of the largest shortest path in a connected graph
- Density
  - The share of possible edges that are actually observed

Let’s find out how to do this all in R!

Overview

- Anatomy of a network
  - Types of networks/graphs
  - Working with network data
- Analysis
  - Graph properties
  - Centrality
  - Roles and positions

Two Measures of Node Centrality

- Degree centrality
  - Number of ties
    - In-degree vs. out-degree in directed graphs
- Eigenvector centrality
  - Weighted degree centrality in which ties to nodes with more connections count more
    - Gets at distinction between bigness vs. centrality
  - Can be interpreted as a measure of coreness

Let’s find out how to do this all in R!

Overview

- Anatomy of a network
  - Types of networks/graphs
  - Working with network data
- Analysis
  - Graph properties
  - Centrality
  - Roles and positions

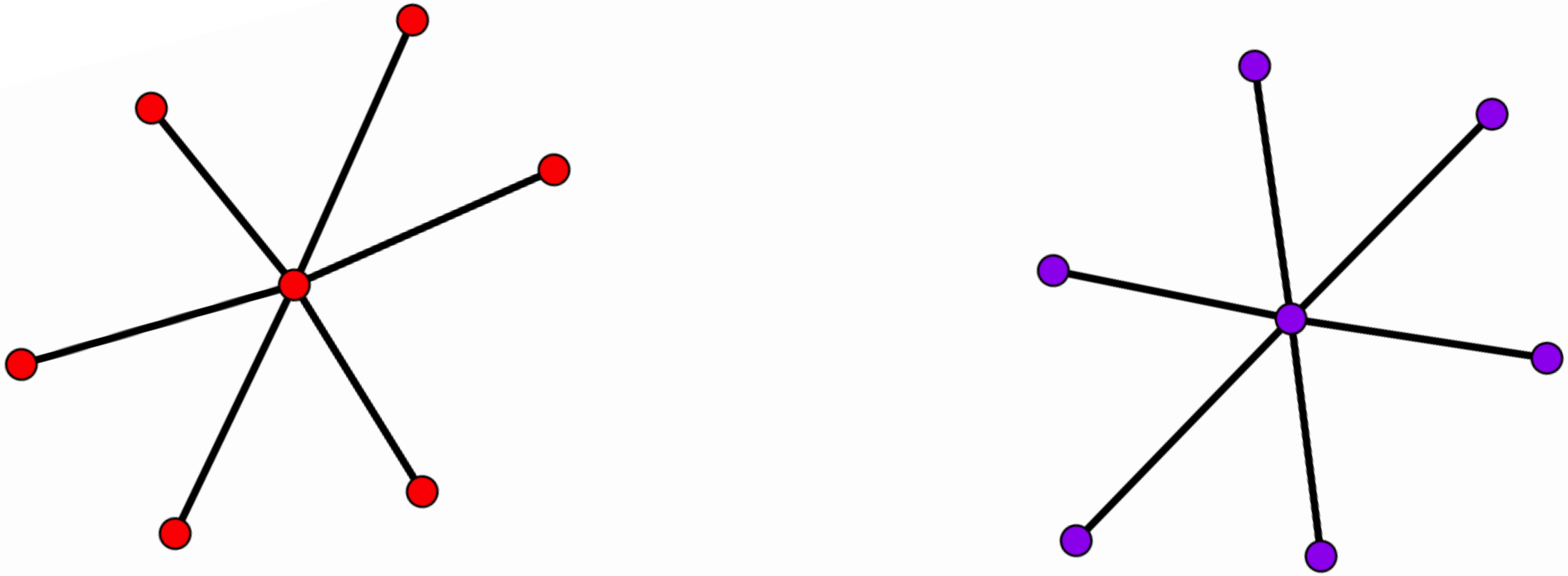
Cohesion vs. Equivalence

- Cohesive subgroups refer to sets of nodes that are closely connected to one another
  - Cliques
    - Everyone is connected to everyone else
- Two actors are structurally equivalent if they share the same ties to other actors in the network
  - Actors can be structurally equivalent without being tied to one another
  - We can use measures of structural equivalence to construct a blockmodel

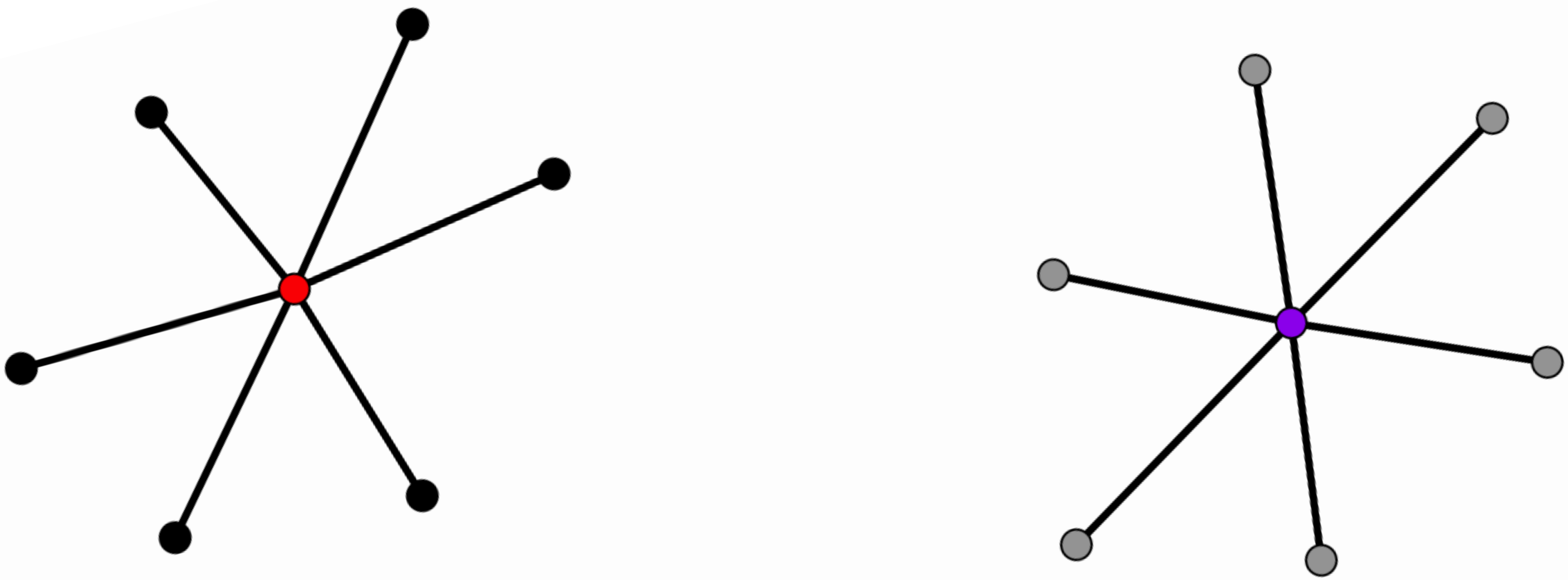
Two Interpretations of a Blockmodel

- Role theory
  - Positions are occupied by sets of structurally equivalent actors
  - Roles are defined by the relationship between positions
  - The set of all roles comprises a role set
- Atheoretical data reduction
  - Structurally equivalent actors are exchangeable
  - When there is perfect equivalence, we can reduce the graph with no loss of information

Illustrating Cohesive Subgroups



Illustrating Structural Equivalence



Blockmodels

- A blockmodel is simplified representation of a graph depicting ties between generalized positions occupied by (structurally) equivalent nodes
  - Extends to networks comprised of multiple relations
- Steps for constructing a blockmodel

1. Measure the degree of equivalence between each pair of nodes
   - Correlation, Euclidean distance
2. Assign nodes to groups using some form of clustering algorithm
   - CONCOR, HCA, or community detection
3. Aggregate ties up to the group level

Constructing a Blockmodel

1. Measure the degree of equivalence between each pair of nodes
    - Correlation vs. distance
2. Partition vertices
    - CONCOR vs. HCA v. community detection
3. Permute the rows and columns of the original matrix
4. Collapse the data to produce a new matrix
    - Dichotomize densities to produce an image matrix
5. Graph the new data

Constructing a Blockmodel

1. Measure the degree of equivalence between each pair of nodes
   - Correlation vs. distance
2. Partition vertices
   - CONCOR vs. HCA v. community detection
3. Permute the rows and columns of the original matrix
4. Collapse the data to produce a new matrix
   - Dichotomize densities to produce an image matrix
5. Graph the new data

ESTIMATES OF CHANGES IN GNP/CAPITA (1955–70) REGRESSED ON CONTROL VARIABLES AND DUMMY VARIABLES REPRESENTING "STRUCTURAL POSITION" (Location in 10-Block Model)

INDEPENDENT VARIABLES[a]	PARTIAL REGRESSION COEFFICIENTS (Standard Errors) Eq. (1)	Eq. (2)	Eq. (3)	Eq. (4)
GNP/capita (1955)	.292 (.080)**	.274 (.081)**	.266 (.082)**	.294 (.082)**
Secondary school enrollment ratio (1955)	20.99 (2.37)**	13.62 (2.95)**	16.98 (2.76)**	11.18 (3.15)**
Block A	...	−468.0 (170.4)**	...	−329.0 (163.1)*
Block B	...	−507.4 (184.7)**	...	−449.6 (182.1)*
Block C′	...	−645.4 (186.0)**	...	−626.2 (184.1)**
Block D	...	−422.7 (139.7)**	...	−386.6 (146.3)**
Block D′	...	−510.0 (180.1)**	...	−490.5 (179.9)**
Block E	...	−542.8 (202.6)**	..	−504.3 (197.9)*
Block E′	...	−582.3 (177.5)**	...	−546.5 (175.9)**
Block F	...	−310.0 (153.4)*	..	−257.3 (158.9)
Block F′	...	−789.5 (204.3)**	..	−758.3 (204.7)**
Diplomats sent (1963–64)	...	...	.142 (.297)	.433 (.309)
Interventions into (1960–67)	...	...	−1.03 (3.54)	−3.44 (3.59)
Interventions by (1960–67)	...	...	−11.47 (4.42)*	−12.68 (4.49)**
Value of exports (1965; millions of U.S. $)	...	...	.036 (.018)*	.010 (.019)
Regression constant	20.33	562.2	51.1	507.6
R^2	68.1%	75.9%	72.7%	79.1%

NOTE.—N = 90 nations.
[a] Dependent variable is change in GNP/capita, 1955–70.
* Significant at .05.
** Significant at .01.

Snyder, David and Edward L. Kick. 1979. "Structural Position in the World System and Economic Growth, 1955–1970: A Multiple-Network Analysis of Transnational Interactions." *American Journal of Sociology* 84: 1096–1126.

Results from CONCOR

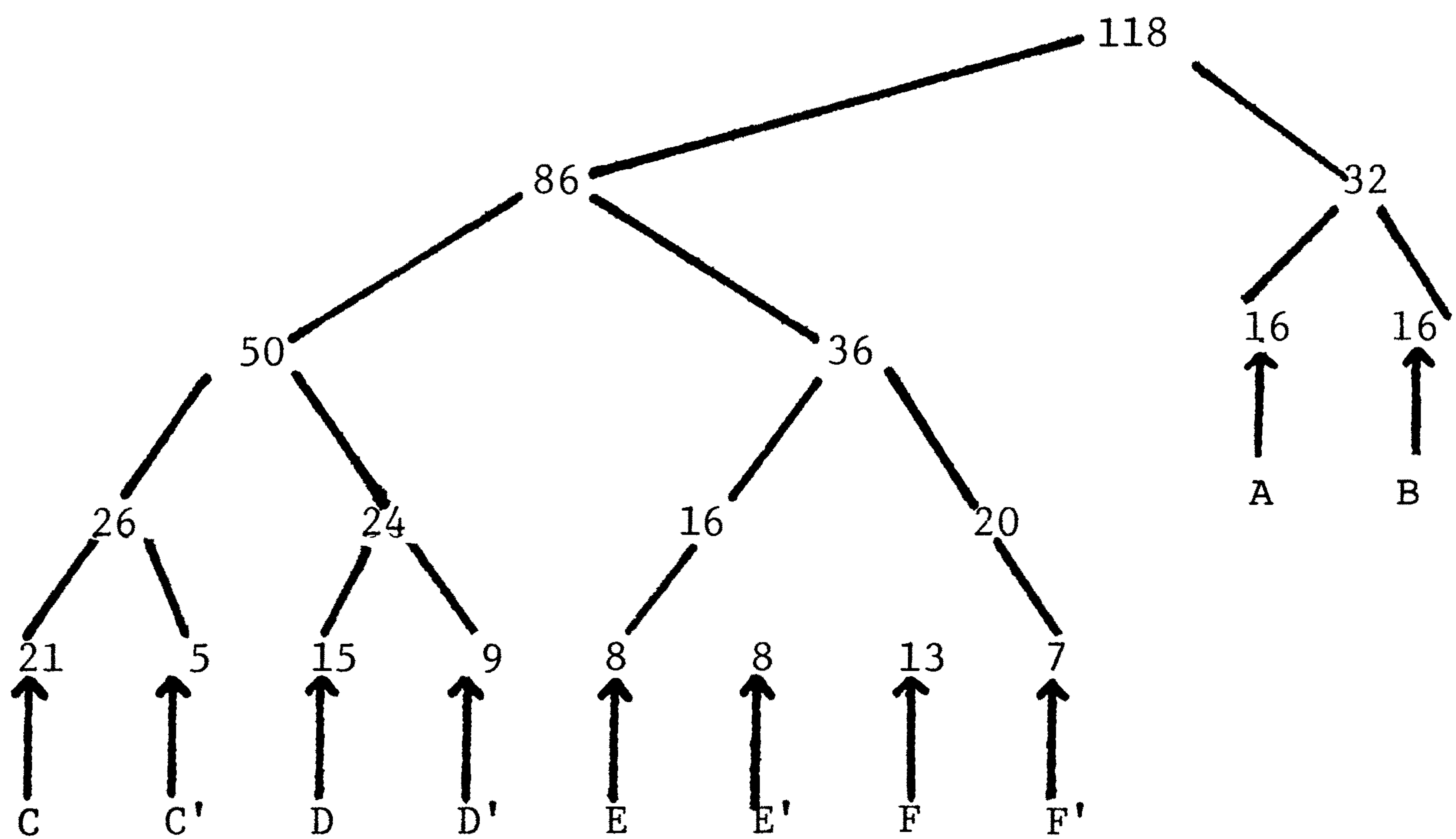


FIG. 1.—Blockmodel "tree" of hierarchical decompositions: 118 nations circa 1965. Letters designate block "names" in 10-block model. See table 1 for list of countries located in each block.

Snyder, David and Edward L. Kick. 1979. "Structural Position in the World System and Economic Growth, 1955–1970: A Multiple-Network Analysis of Transnational Interactions." *American Journal of Sociology* 84: 1096–1126.

Results from CONCOR

LISTING OF COUNTRIES IN EACH BLOCK: FOUR-NETWORK BLOCKMODEL
(Trade, Interventions, Diplomats, and Treaties)
FOR 118 NATIONS CIRCA 1965

Block	Nations
A	Chad, Congo (Brazzaville), Congo (Kinshasa), Uganda, Burundi, Rwanda, Somalia, Ethiopia, Malagasy Republic, Morocco, Algeria, Tunisia, Libya, Sudan, United Arab Republic,* Yemen*
B	Mali,* Mauritania,* Ghana,* Upper Volta,* Senegal, Dahomey, Niger, Ivory Coast, Republic of Guinea, Liberia, Sierra Leone, Togo, Cameroun, Nigeria, Gabon, Central African Republic
C	Canada, United States, United Kingdom, Netherlands, Belgium, Luxembourg, France, Switzerland, Spain, Portugal, West Germany, Austria, Italy, Yugoslavia, Greece, Sweden, Norway, Denmark, South Africa, Japan, Australia
C′	Venezuela, Peru, Argentina, Uruguay, South Korea
D	Cuba, Ireland, East Germany, Hungary, Cyprus, Bulgaria, Rumania, USSR, Kenya, Iran, Turkey, Iraq, Lebanon, Jordan, Israel
D′	Finland, Saudi Arabia, Taiwan, India, Pakistan, Burma, Ceylon, Malaysia, Philippines
E	Panama, Colombia, Ecuador, Brazil, Bolivia, Paraguay, Chile, North Vietnam
E′	Haiti, Dominican Republic, Mexico, Guatemala, Honduras, El Salvador, Nicaragua, Costa Rica
F	Jamaica, Trinidad and Tobago, Poland, Czechoslovakia, Malta, China (People's Republic), Mongolian Republic, Nepal, Thailand, Cambodia, Laos, New Zealand, Iceland
F′	Albania, Syria, Kuwait, Afghanistan, North Korea, South Vietnam, Indonesia

* Starred countries in blocks A and B cluster together in further "splits" of the data, although they are not shown as separate blocks in the analyses.

Snyder, David and Edward L. Kick. 1979. "Structural Position in the World System and Economic Growth, 1955–1970: A Multiple-Network Analysis of Transnational Interactions." *American Journal of Sociology* 84: 1096–1126.

TRADE (Mean Density = .3410)

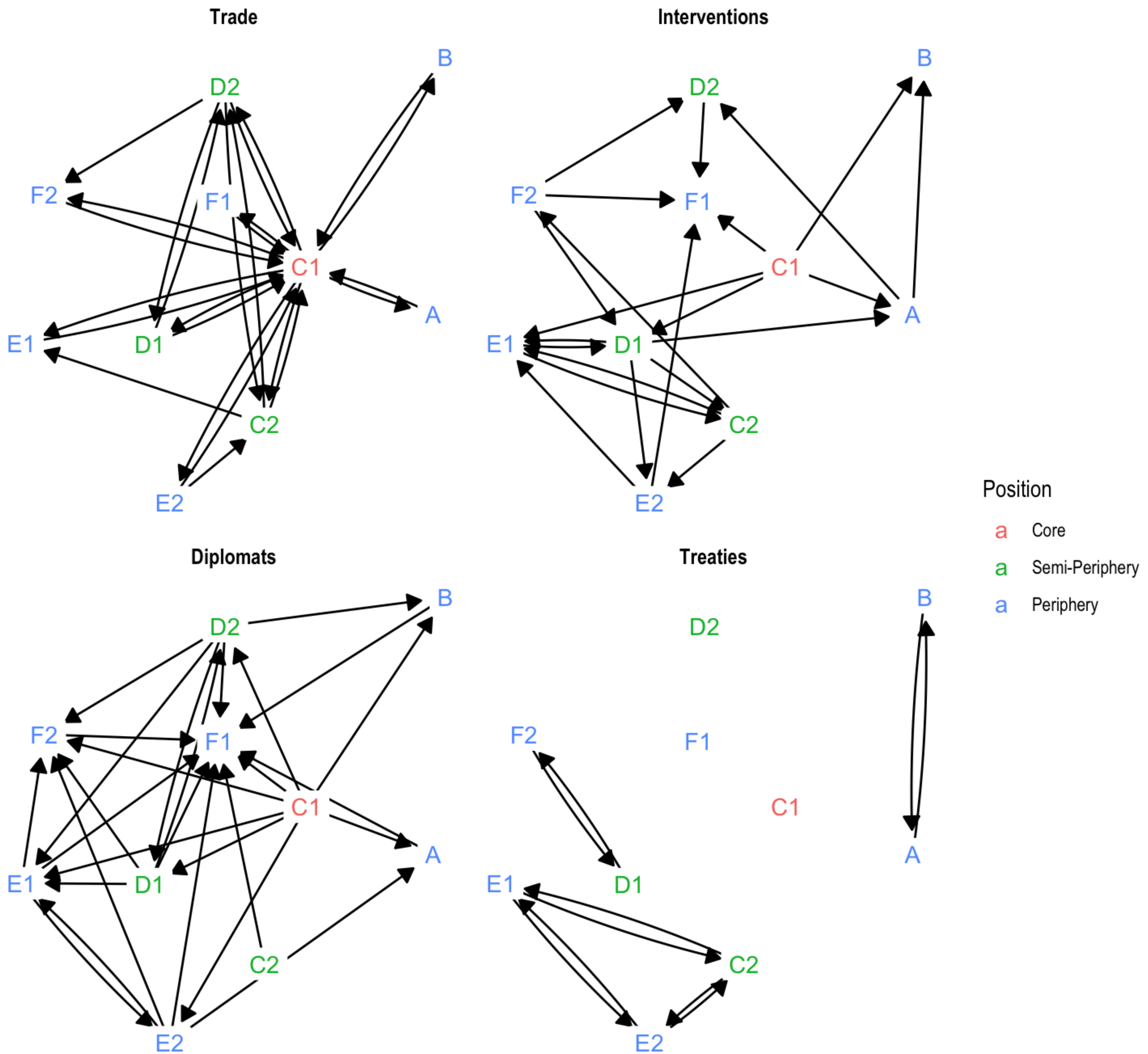
	C	C'	D	D'	E	E'	F	F'	A	B
C	.943	.895	.908	.889	.744	.714	.630	.612	.622	.500
C'	.829	.560	.320	.422	.250	.525	.215	.114	.087	.100
D	.854	.267	.613	.474	.033	.133	.282	.286	.271	.108
D'	.884	.533	.459	.617	.153	.194	.308	.349	.215	.069
E	.679	.225	.042	.139	.422	.172	.106	0	.023	0
E'	.643	.525	.125	.125	.125	.297	.135	.036	.039	.023
F	.634	.215	.262	.265	.029	.135	.183	.165	.144	.087
F'	.551	.171	.229	.317	.018	.018	.132	.082	.045	.009
A	.604	.037	.225	.181	0	.039	.115	.045	.141	.063
B	.533	.025	.092	.042	0	.008	.038	0	.070	.160

Snyder, David and Edward L. Kick. 1979. "Structural Position in the World System and Economic Growth, 1955–1970: A Multiple-Network Analysis of Transnational Interactions." *American Journal of Sociology* 84: 1096–1126.

TRADE

	C	C'	D	D'	E	E'	F	F'	A	B
C	1	1	1	1	1	1	1	1	1	1
C'	1	1	0	1	0	1	0	0	0	0
D	1	0	1	1	0	0	0	0	0	0
D'	1	1	1	1	0	0	0	1	0	0
E	1	0	0	0	1	0	0	0	0	0
E'	1	1	0	0	0	0	0	0	0	0
F	1	0	0	0	0	0	0	0	0	0
F'	1	0	0	0	0	0	0	0	0	0
A	1	0	0	0	0	0	0	0	0	0
B	1	0	0	0	0	0	0	0	0	0

Snyder, David and Edward L. Kick. 1979. "Structural Position in the World System and Economic Growth, 1955–1970: A Multiple-Network Analysis of Transnational Interactions." *American Journal of Sociology* 84: 1096–1126.

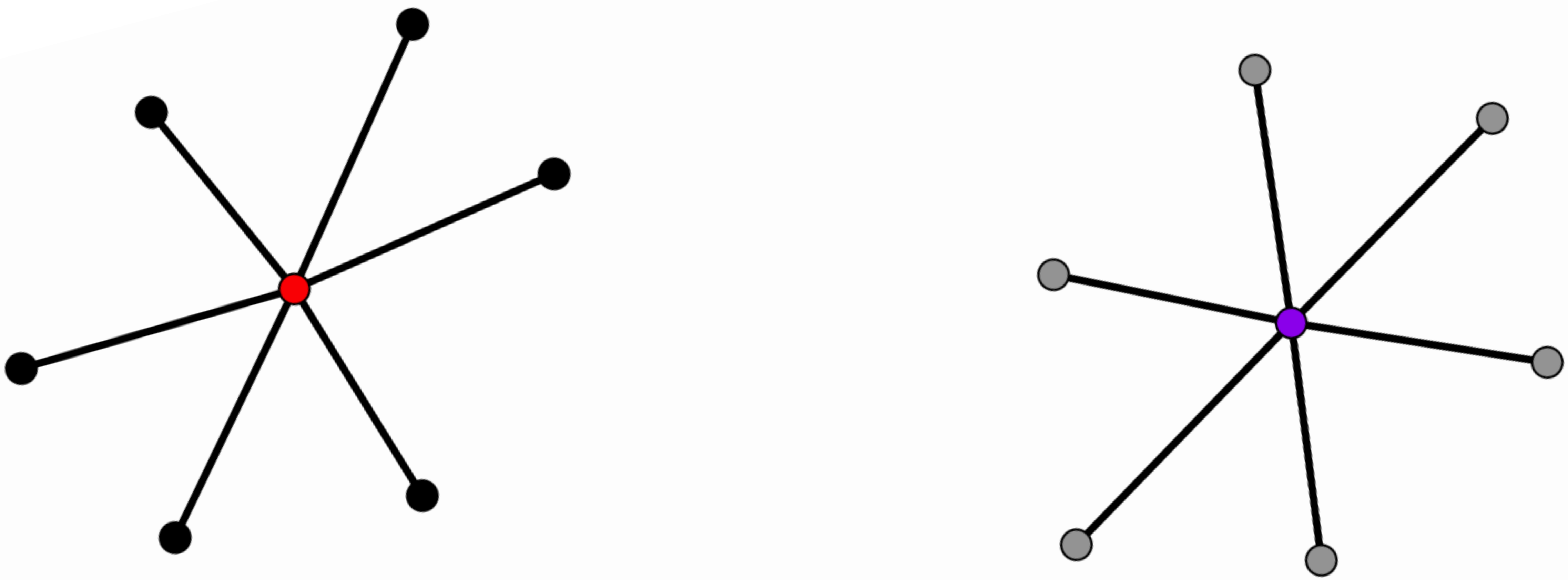

Trade
Interventions
Diplomats
Treaties
A
B
C1
C2
D1
D2
E1
E2
F1
F2
Position
a Core
a Semi-Periphery
a Periphery

Let’s find out how to do this all in R!

Two Criticisms of Classical Blockmodeling

- Certain approaches are mathematically unsound
  - It isn’t clear why CONCOR works or what exactly it is doing
- All approaches are theoretically unsound
  - Structural equivalence doesn’t match what we have in mind when we talk about roles

Illustrating Structural Equivalence



Illustrating Regular Equivalence