USC ANNENBERG
School for Communication & Journalism

A New Measure of Structural Holes for Dynamic and Adaptive Networks

Poong Oh

Motivation of This Study

- **Research topic**: Evolution of Inter-organizational Networks

- **Assumptions:**
  - **Adding and/or Deleting** partnerships (an adaptive/dynamic network); Not randomly; but in a manner to **maximize the benefits** from the network.
  - **Maximizing network benefit** = building an **effective network** as large as possible ← Structural holes (Burt, 1992)

- **Research Questions**:
  - What does the final stage of the network look like?
  - What is the *equilibrium state* of the network, if any?

- Running simulation → The results didn't make any sense
  - Errors in the simulation code?
  - Or Problems with the structural hole argument???

Burt's Structural Holes Argument
(Background)

Network Matters

- Market production equation:

(Profit) = (Investment) X (Rate of Return)

$10 million = $100 million X 0.1

Network Matters

- Market production equation:

(Profit) = (Investment) X (Rate of Return)

- Under the **perfect competition** assumption, *Rate of Return* is treated as **constant**; thus, profit is determined only by investment.

- **Perfect production factor mobility**: In the long run, factors of production are perfectly mobile allowing free long term adjustments to changing market conditions; from low-yield to high-yield investment

Network Matters

- Market production equation:

(Profit) = (Investment) X (Rate of Return)

- **Competition is *never* perfect**:
  - "**Social structure renders** competition imperfect by creating entrepreneurial opportunities for certain players and not for others" (Burt, 1992; p. 8)
  - Social structure as "**impediments** to reallocating" financial and human capitals (p. 10).
  - In short, **entrepreneurial opportunities** and/or **competitive advantage** are **unevenly distributed** across the actors

Network Matters

- Market production equation:

(Profit) = (Investment) X (Rate of Return)

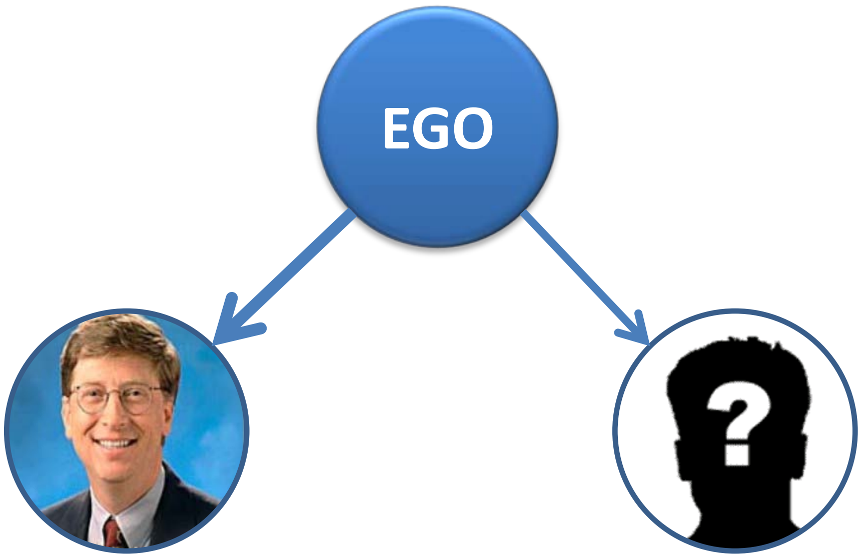
Financial & Human Capital

Social Capital
"the better connected, the higher return"
(Burt 2005)

Q: What is the **better connection**?

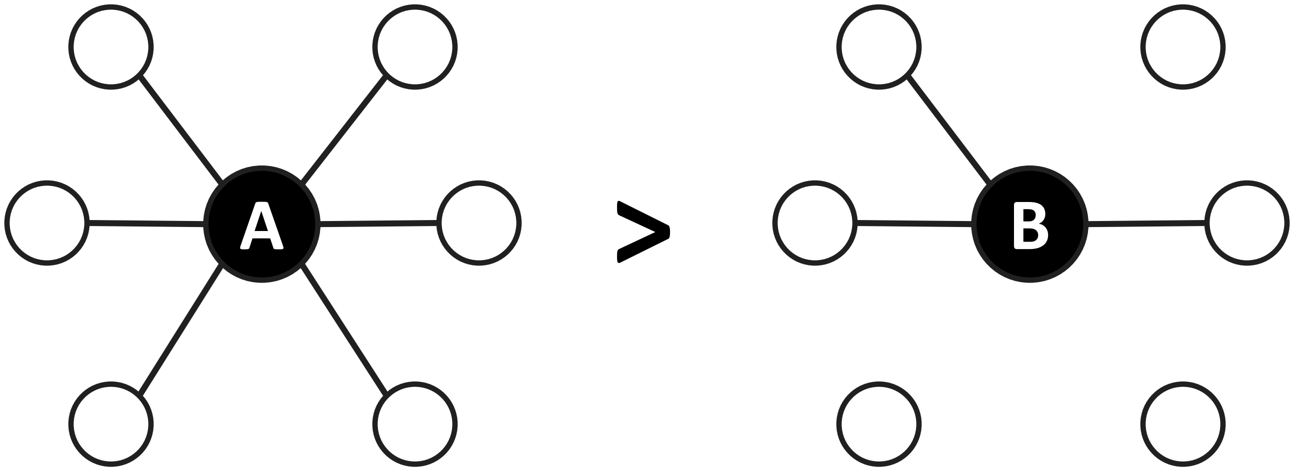
What is the better connection?

- "Doing business with **wealthy clients**… has **a higher margin** than doing business with **poor clients**" (Burt 1992, p. 12)
- **TO-WHOM question** is not interesting; but **HOW question** is
- **Generalization beyond specific individuals**



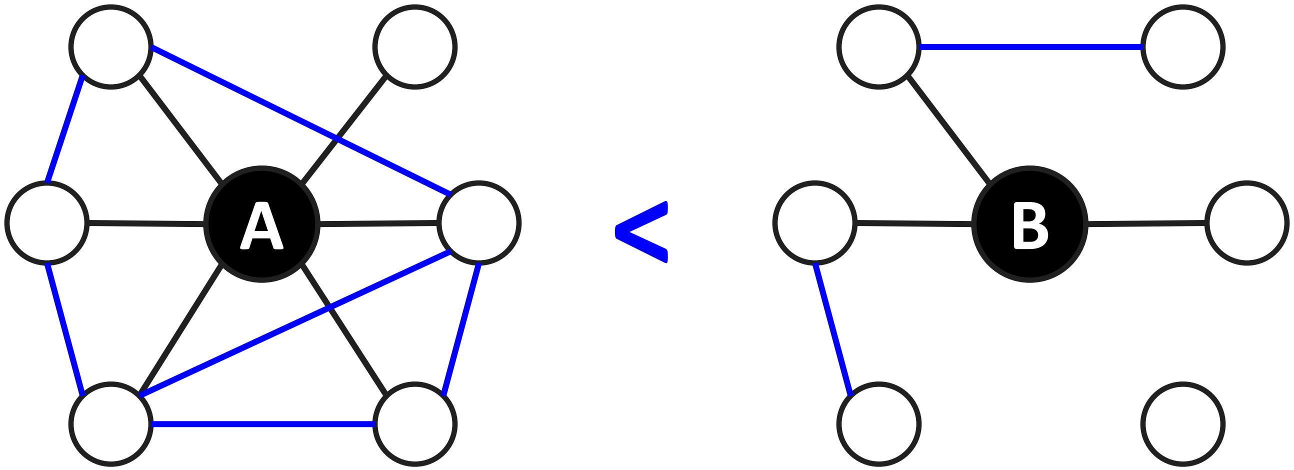
What is the better connection?

- **Network size** matters?



What is the better connection?

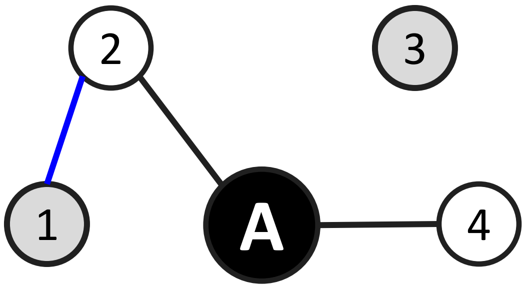
- **Network size** matters?



- Yes, it does. But when cost (network time and energy invested) is considered, **the less redundant contacts are better** (i.e., maximal efficiency; *cf*. network closure, Burt, 2005)

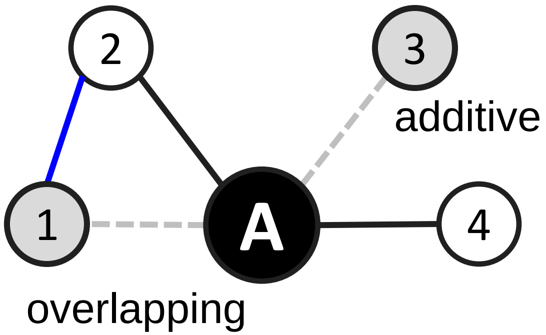
Structural Holes

- **Network positions connecting nonredundant contacts**
- Network benefits in structural holes are "**additive** rather than **overlapping**" (p. 18)



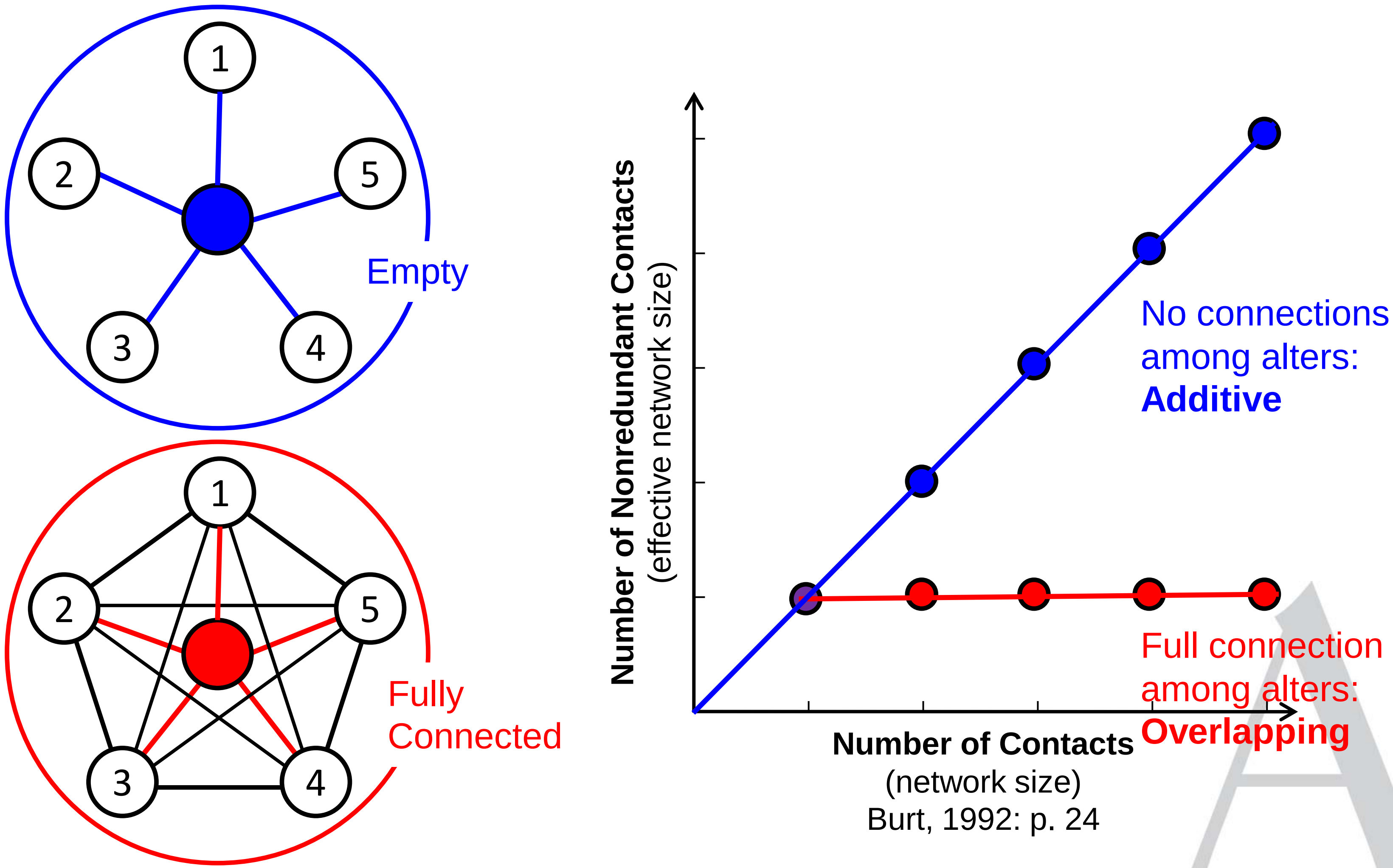
Structural Holes

- **Network positions connecting nonredundant contacts**
- Network benefits in structural holes are "**additive** rather than **overlapping**" (p. 18)



- **Effective Network Size**
- **Network Constraint**

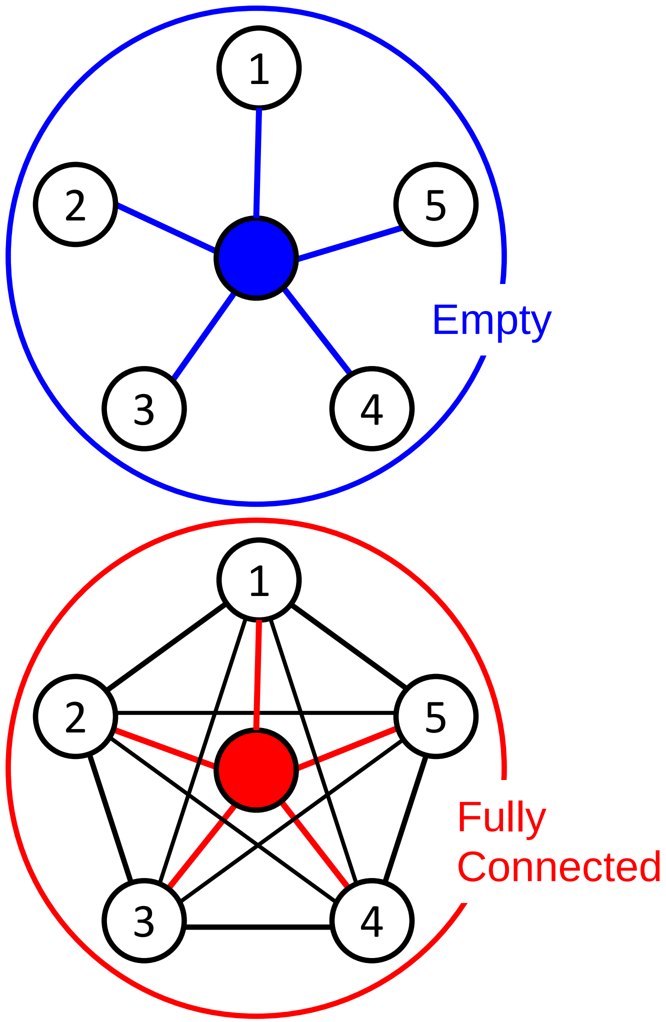
Structural Holes

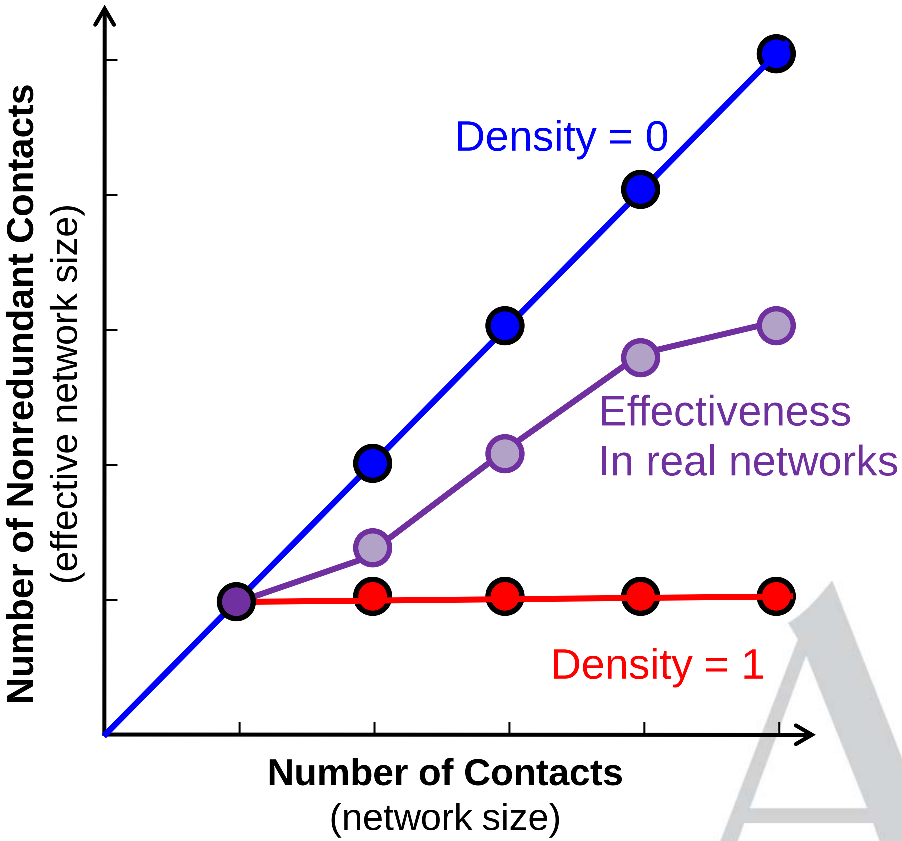


Burt, 1992: p. 24

USC | ANNENBERG
School for Communication & Journalism

Structural Holes

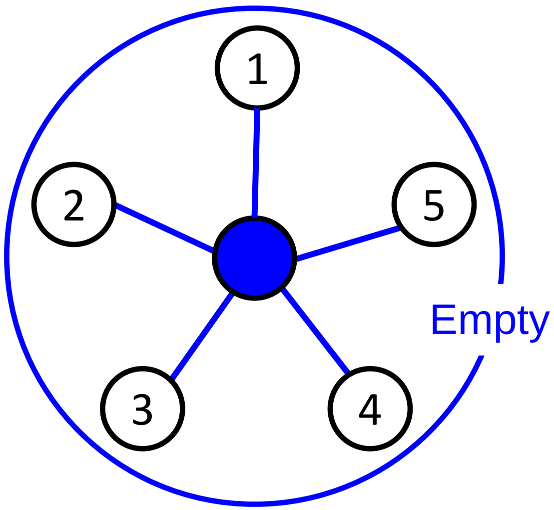




Burt, 1992: p. 24

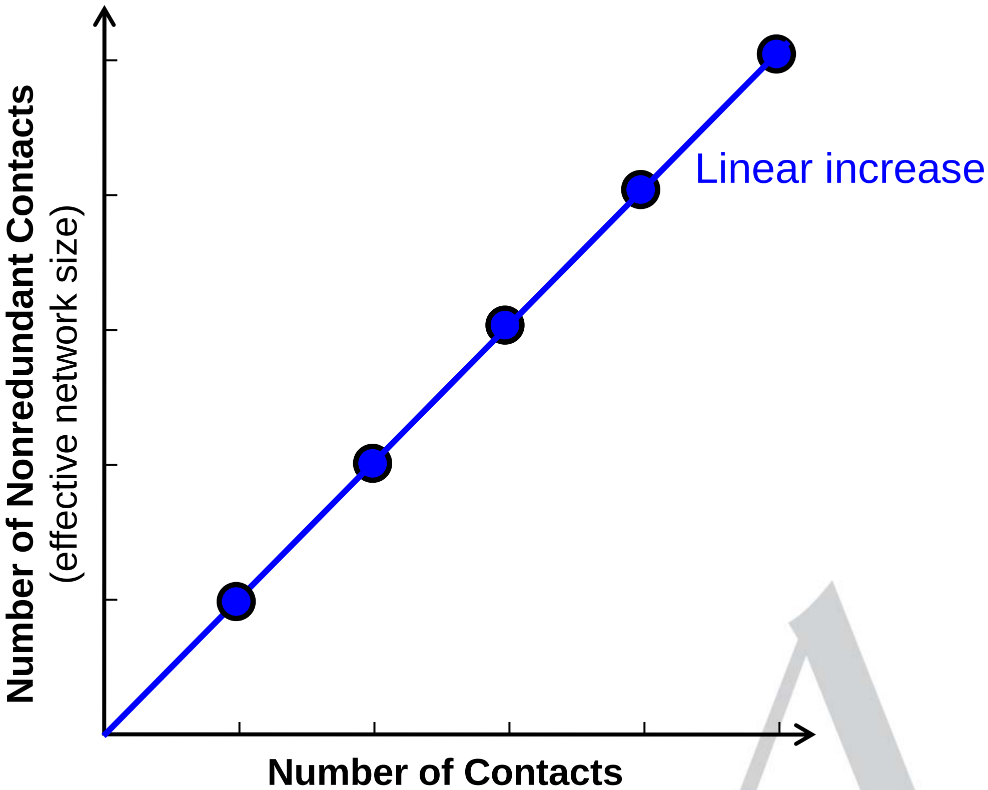
Poong's Argument
(Introduction of heterogeneity index)

Hidden Assumptions #1



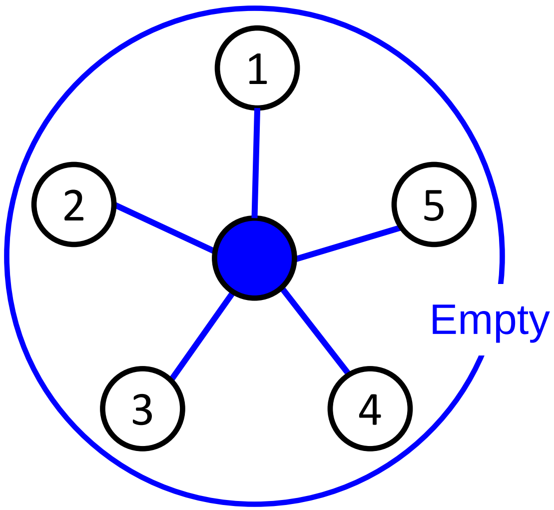
 # 1: When X_{jk} = 0,
$U(X_{ij}, X_{ik}) = U(X_{ij})+U(X_{ik})$;

e.g., Completely different opportunities from disconnected others

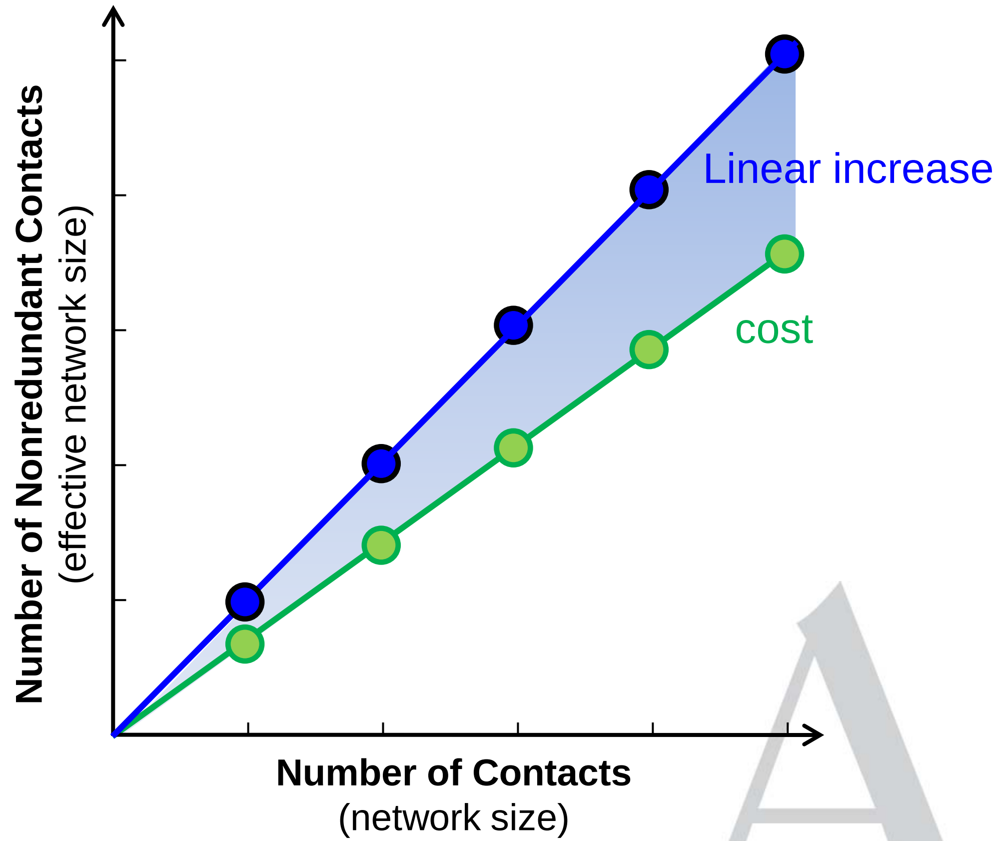


Burt, 1992: p. 24

Hidden Assumptions #1

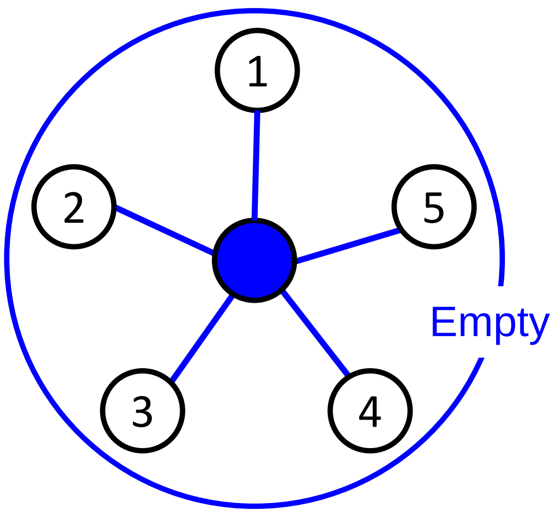


When **unit cost < benefit**, a rational actor will add new ties, until connected to all others

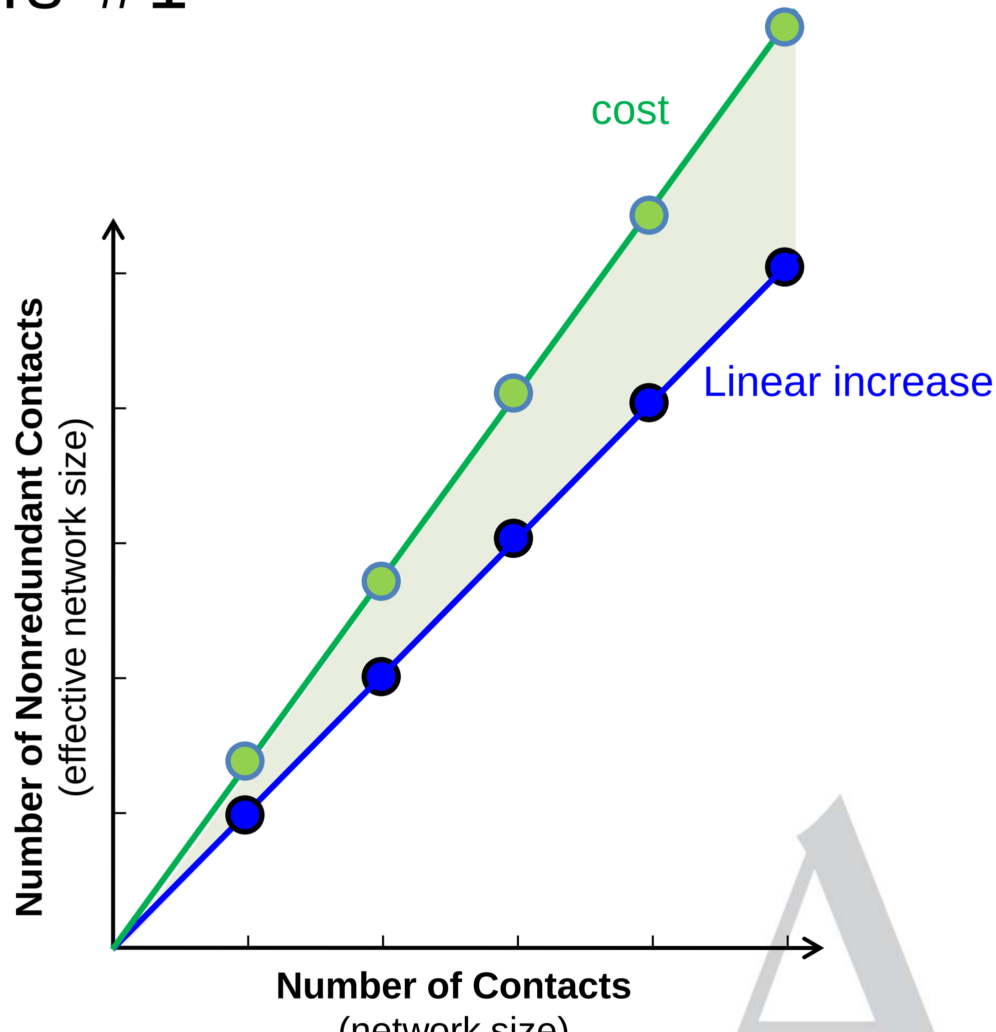


Burt, 1992: p. 24

Hidden Assumptions #1

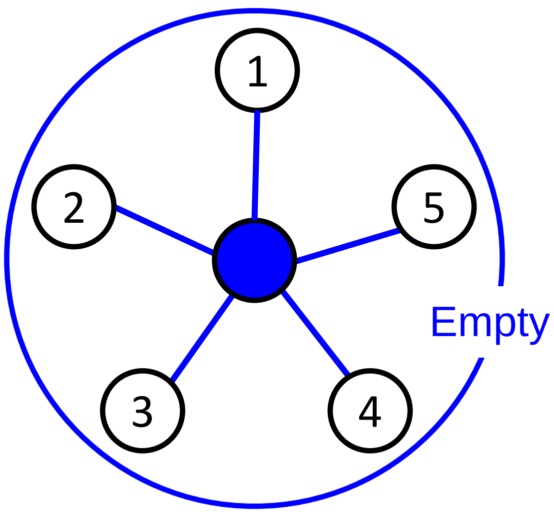


When **unit cost > benefit**, a rational actor will initiate no ties at all, simply follow market mechanism: **transaction cost theory** (Williamson, 1979)

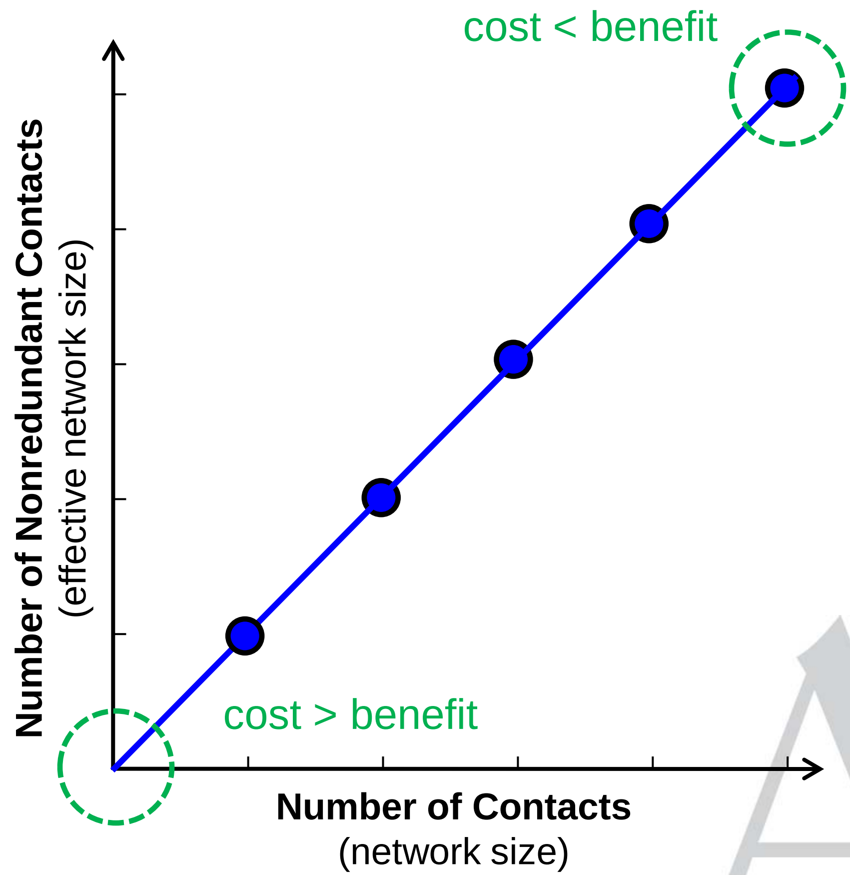


Burt, 1992: p. 24

Hidden Assumptions #1



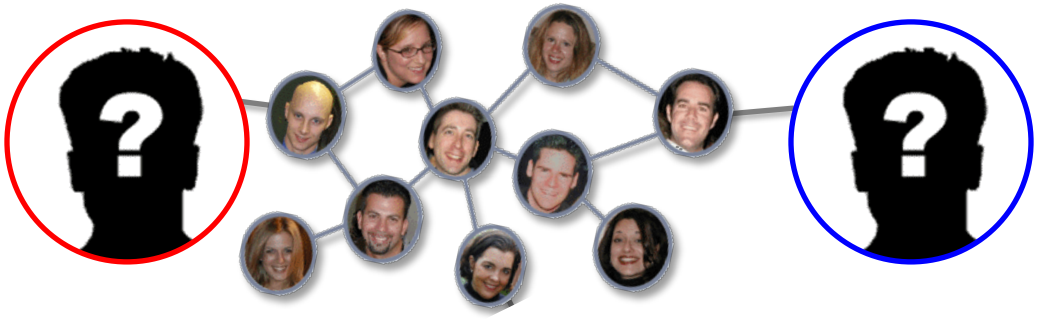
One possible equilibrium state, depending on;
If cost < benefit, all contacts
If cost > benefit, no contact



Burt, 1992: p. 24

Hidden Assumptions #1

- **Do disconnected alters provide completely different opportunities?**

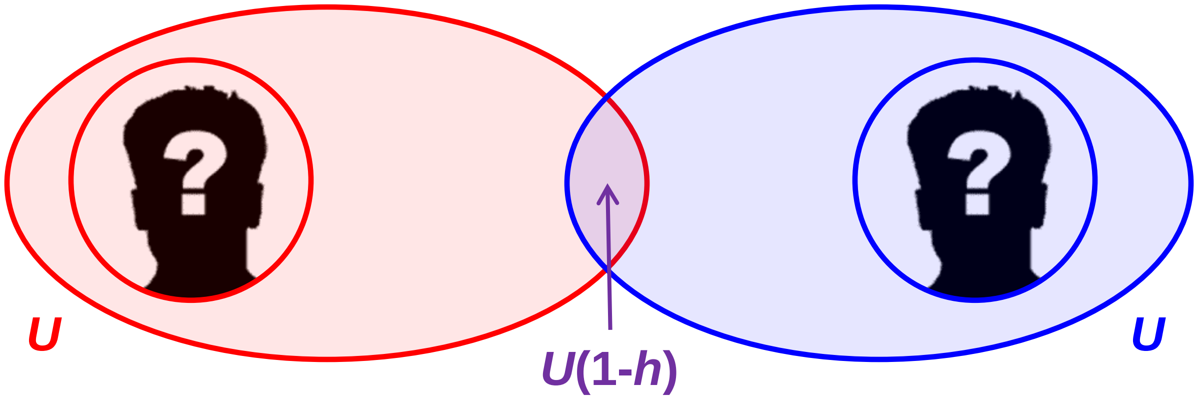


  - Possibility of indirect connections ("six degrees of separation")
  - Redundancy by *structural equivalence* (Burt, 1992: p. 18)

The lack of direct connection does not imply ***difference***

Hidden Assumptions #1

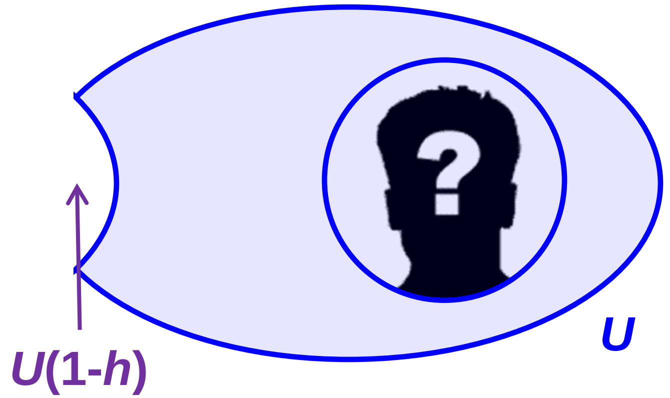
- **Do disconnected alters provide completely different opportunities?**



- Heterogeneity index, ***h***. [0, 1]
  - ***h*** = 0: no difference at all; complete overlap
  - ***h*** = 1: complete difference → Burt's formula

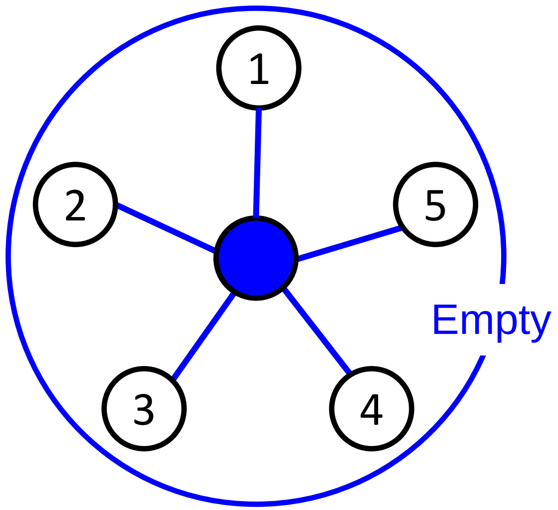
Hidden Assumptions #1

- **Do disconnected alters provide completely different opportunities?**



- Heterogeneity index, ***h***. [0, 1]
  - ***h*** = 0: no difference at all; complete overlap
  - ***h*** = 1: complete difference → Burt's formula
  - $U(X_{ij}, X_{ik}) = U(X_{ij}) + U(X_{ik}) - (1 - h)\cdot U(X_{ik}) = U(X_{ij}) + h\cdot U(X_{ik})$
  - No matter how small the overlap (1 – *h*) is, it makes a **big difference**, unless it's not zero

Hidden Assumptions #1



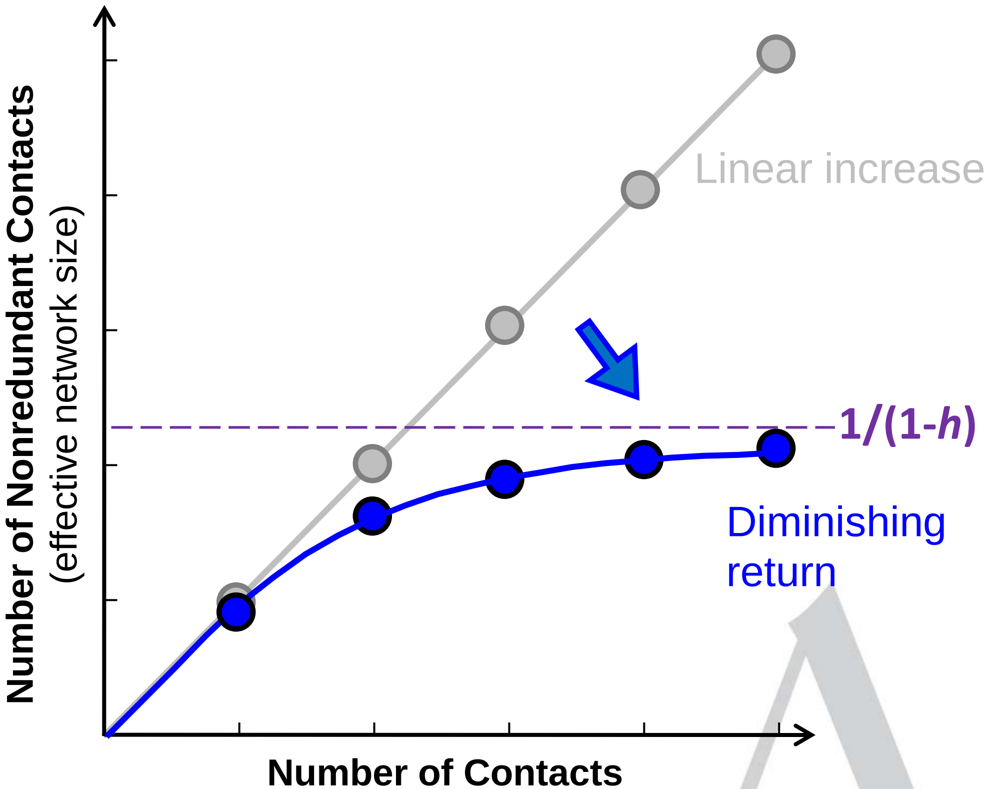
U

$U + Uh$

$U + Uh + Uh^2$

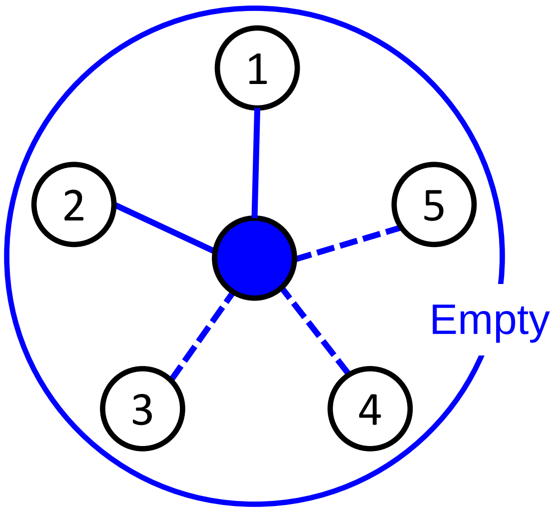
$U + Uh + Uh^2 + Uh^3$

...

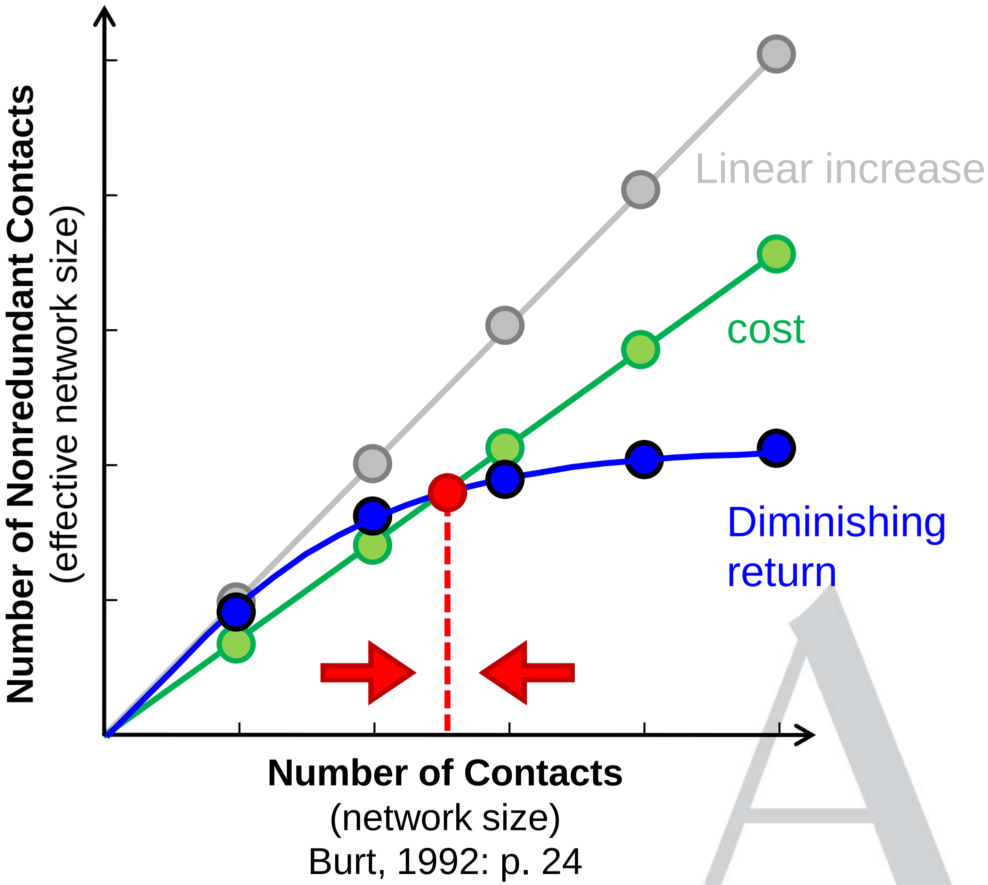


Burt, 1992: p. 24

Hidden Assumptions #1



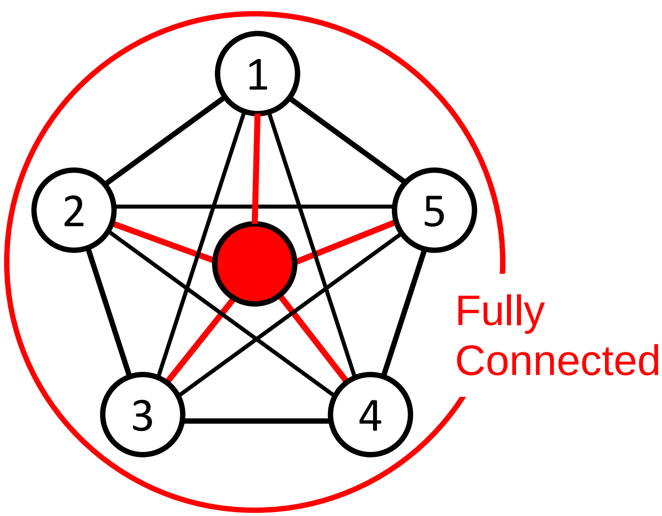
A rational actor won't build more than 2 (in this example): i.e., "**adjusted equilibrium state**" by *h*

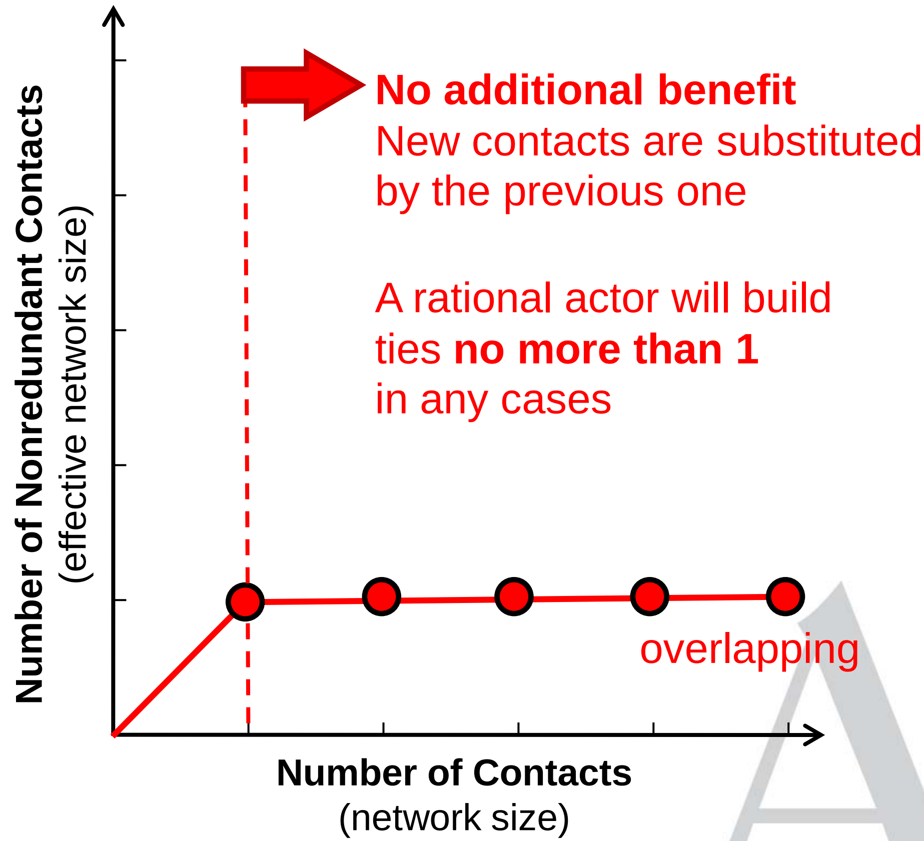


Hidden Assumptions #2

 # 2: When X_{jk} = 1,
$U(X_{ikj}) = U(X_{ij})$

i.e., the benefit from an indirect tie is ***substitutable*** for that from a direct tie

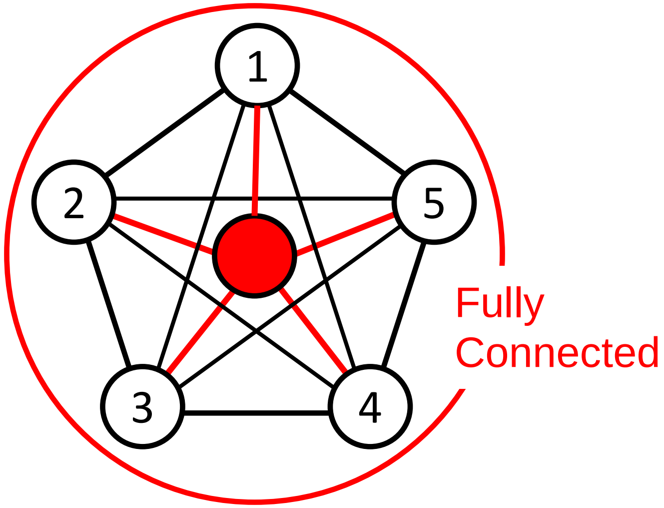


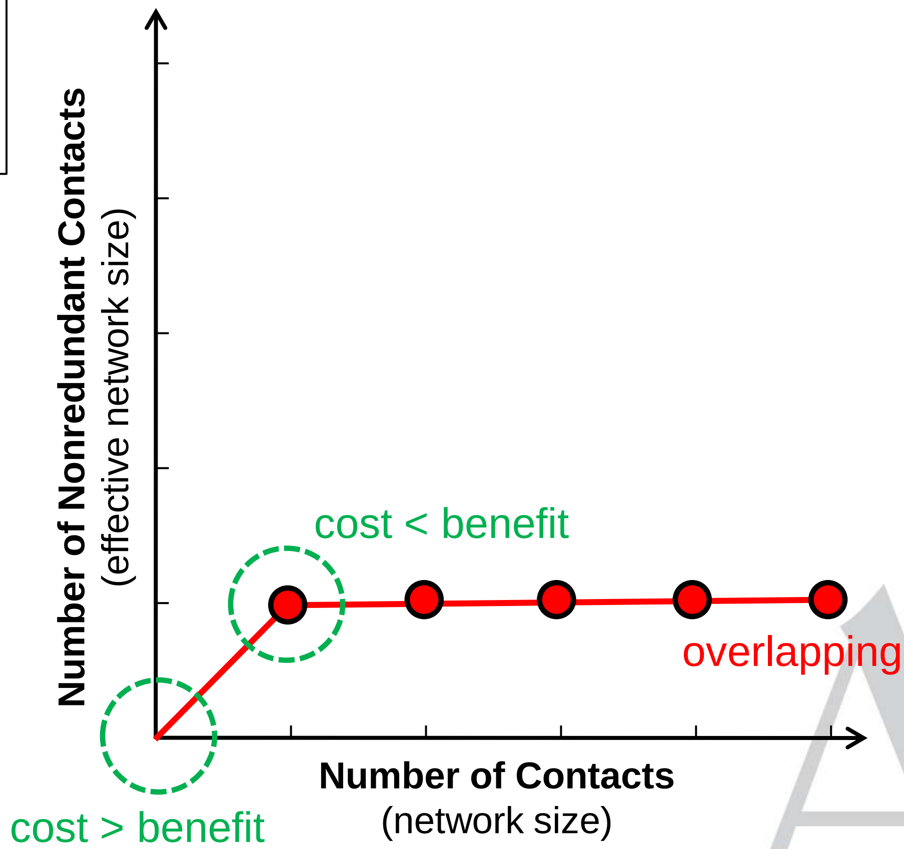


Burt, 1992: p. 24

Hidden Assumptions #2

One possible equilibrium state, depending on;
If cost < benefit, 1 contact
If cost > benefit, no contact



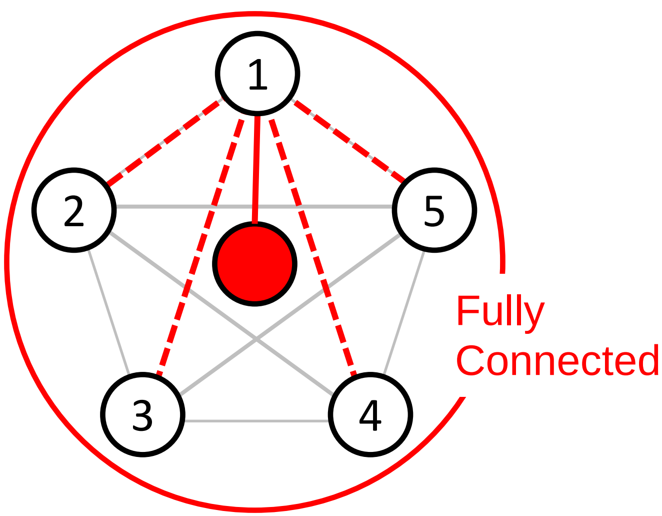


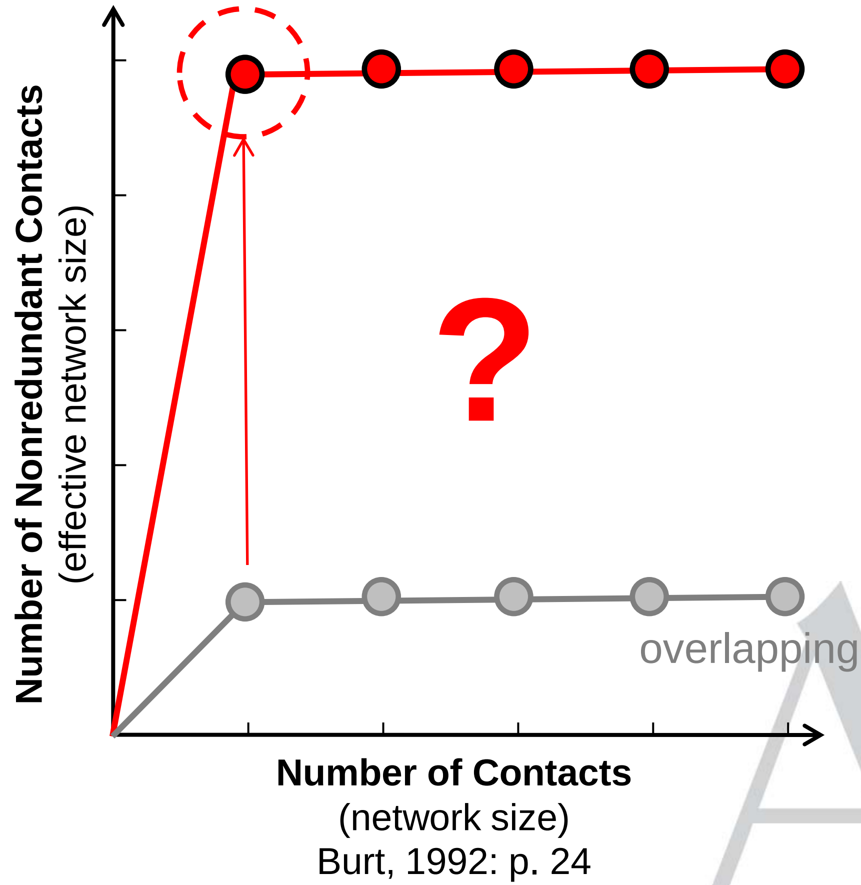
Burt, 1992: p. 24

Hidden Assumptions #2

 # 2: When X_{jk} = 1,
$U(X_{ikj}) = U(X_{ij})$

i.e., the benefit from an indirect tie is ***substitutable*** for that from a direct tie

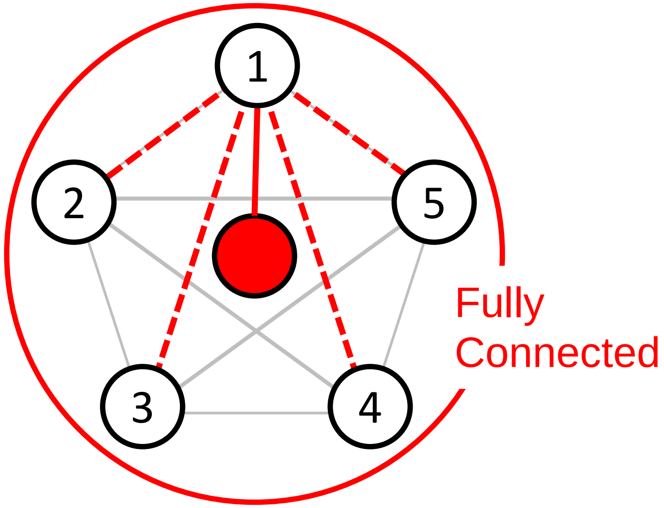




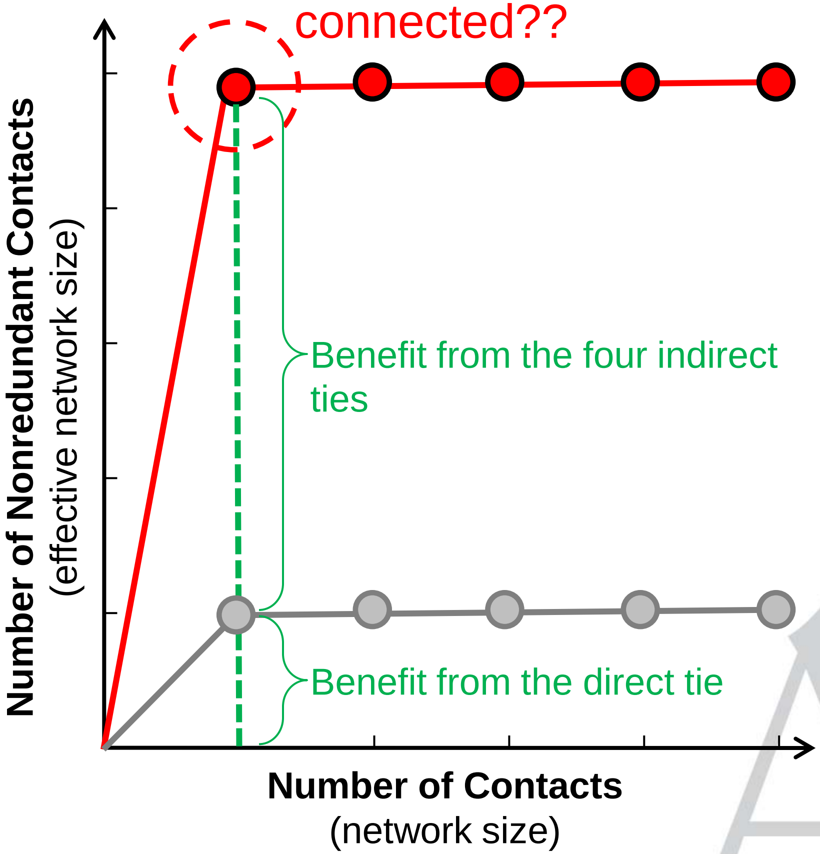
Hidden Assumptions #2

 # 2: When X_{jk} = 1,
$U(X_{ikj}) = U(X_{ij})$

i.e., the benefit from an indirect tie is ***substitutable*** for that from a direct tie



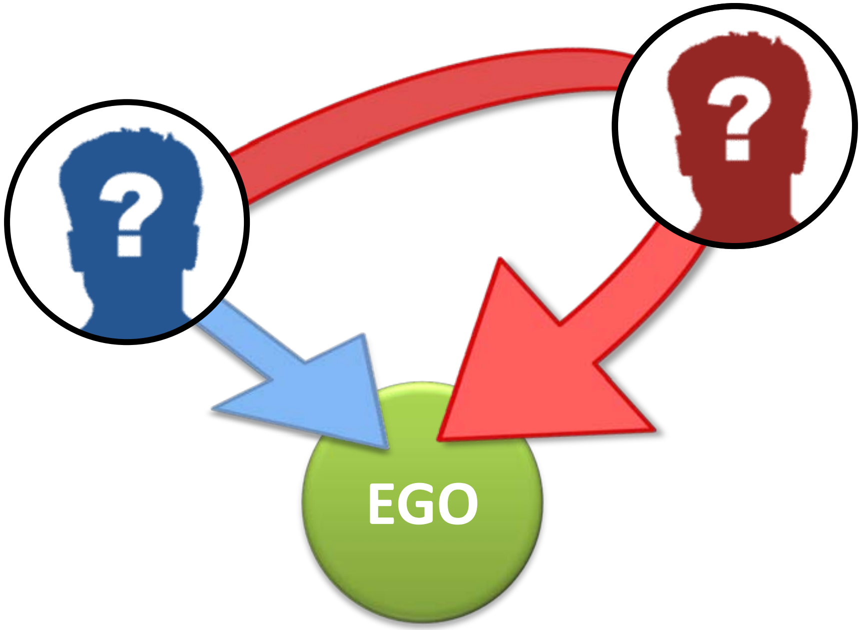
The maximal benefit from a first contact, since all others are Indirectly connected??



Burt, 1992: p. 24

Hidden Assumptions #2

- **Are indirect ties completely substitutable for direct ties?**



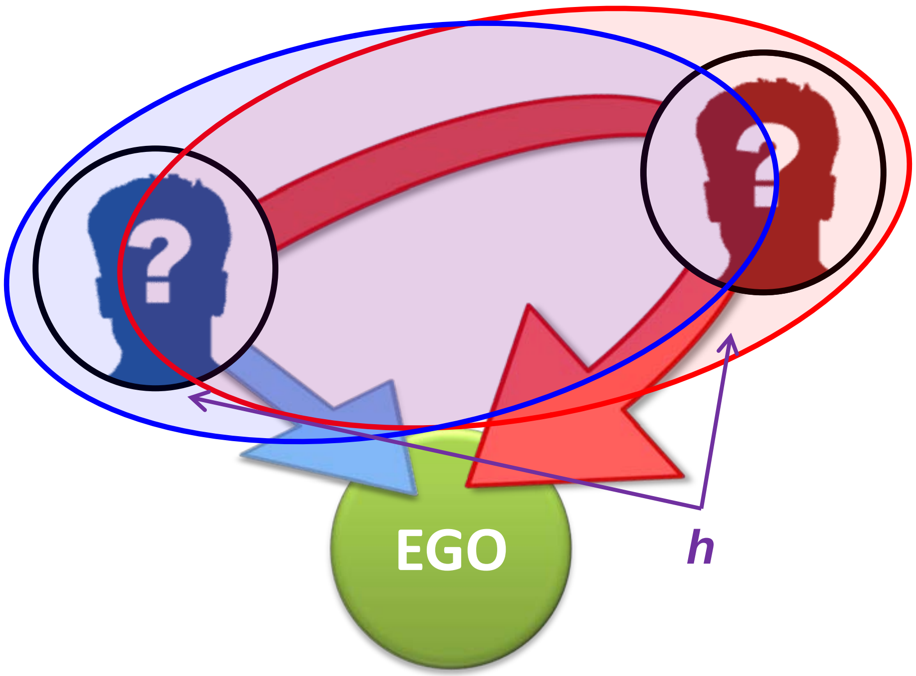
Possibility of **omission** and **distortion** (e.g., rumor spread)

The presence of the tie between the two persons does not necessarily imply that they offer ***identical opportunities***;

They differ from each other to a degree of ***h***

Hidden Assumptions #2

- **Are indirect ties completely substitutable for direct ties?**



Possibility of **omission** and **distortion** (e.g., rumor transmission)

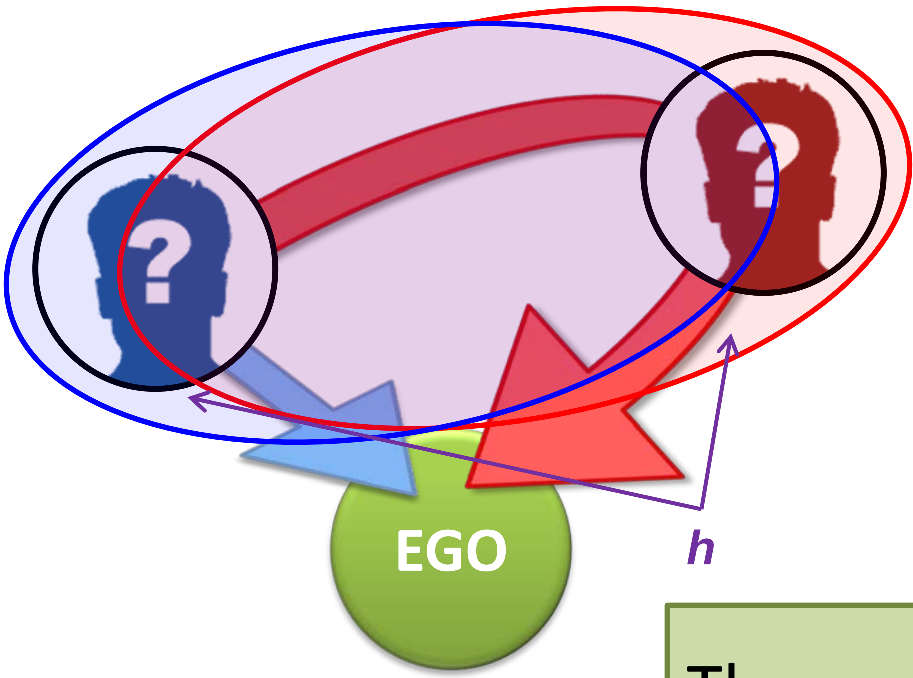
The presence of the tie between the two persons does not necessarily imply that they offer ***identical opportunities***;

They differ from each other to a degree of ***h***

$$U(X_{ijk}) = U(X_{ij}) - U(X_{ik}) + h \cdot U(X_{ik}) = U(X_{ij}) + (1 - h) \cdot U(X_{ik})$$

Hidden Assumptions #2

- **Are indirect ties completely substitutable for direct ties?**



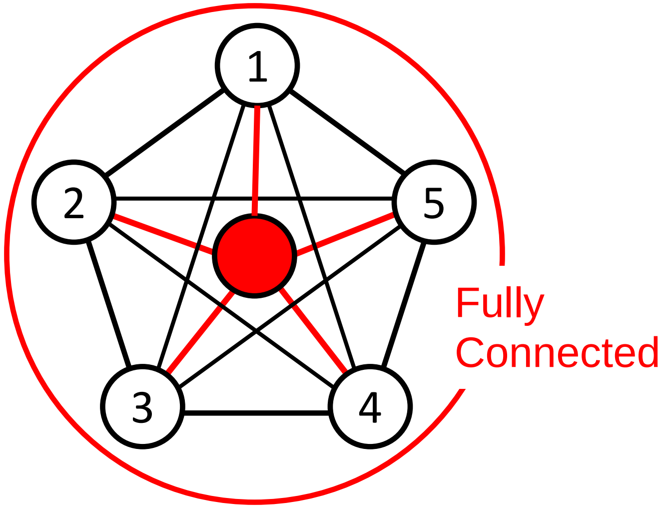
Possibility of **omission** and **distortion** (e.g., rumor transmission)

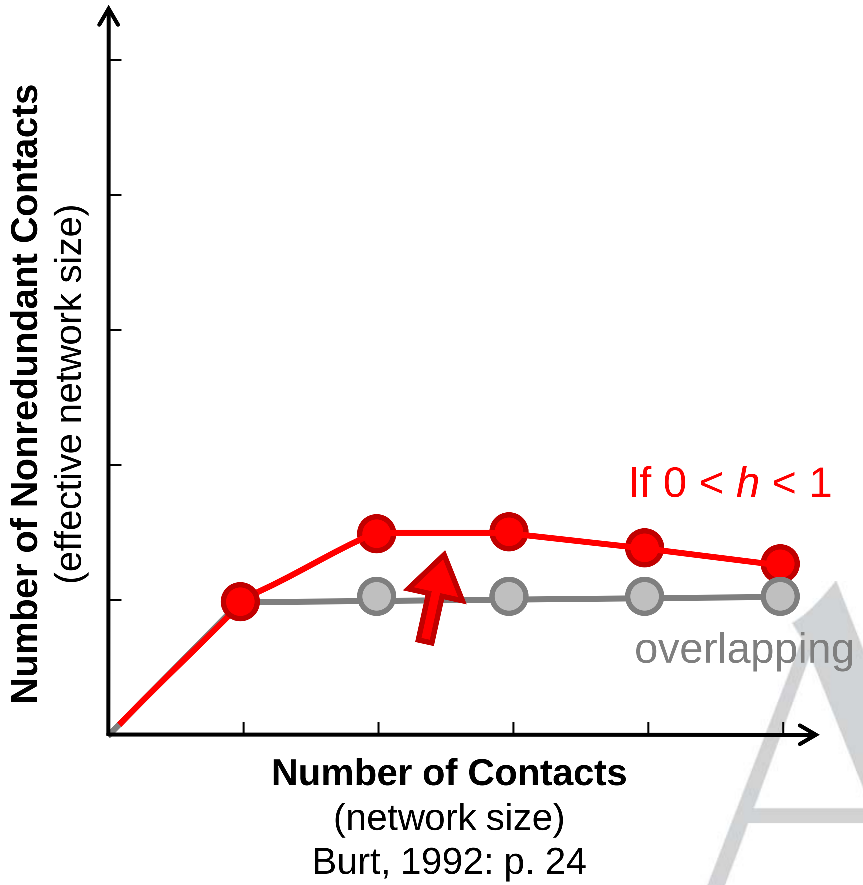
The presence of the tie between the two persons does not necessarily imply that they offer ***identical opportunities***;

They differ from each other to a degree of ***h***

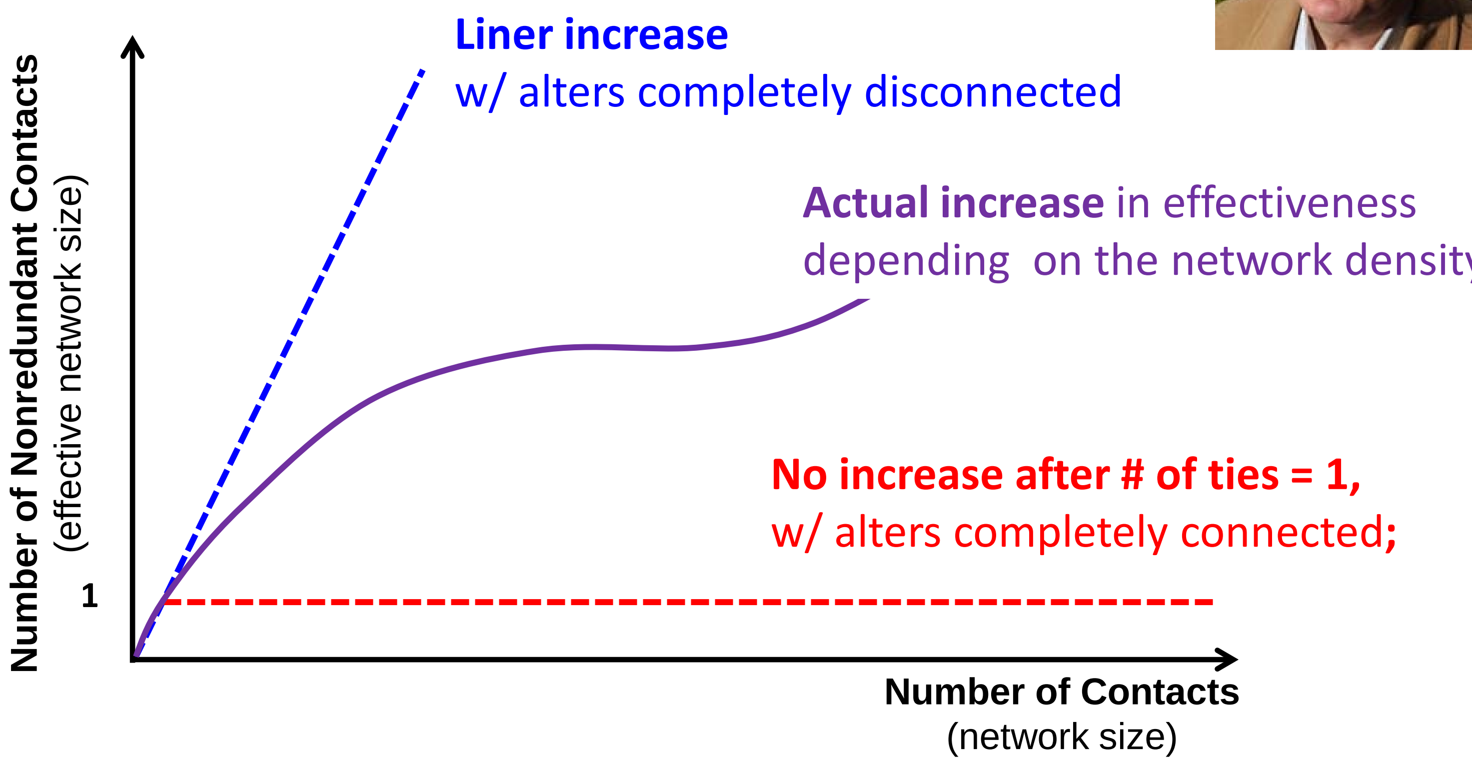
The connection does not guarantee ***sameness***

Hidden Assumptions #2





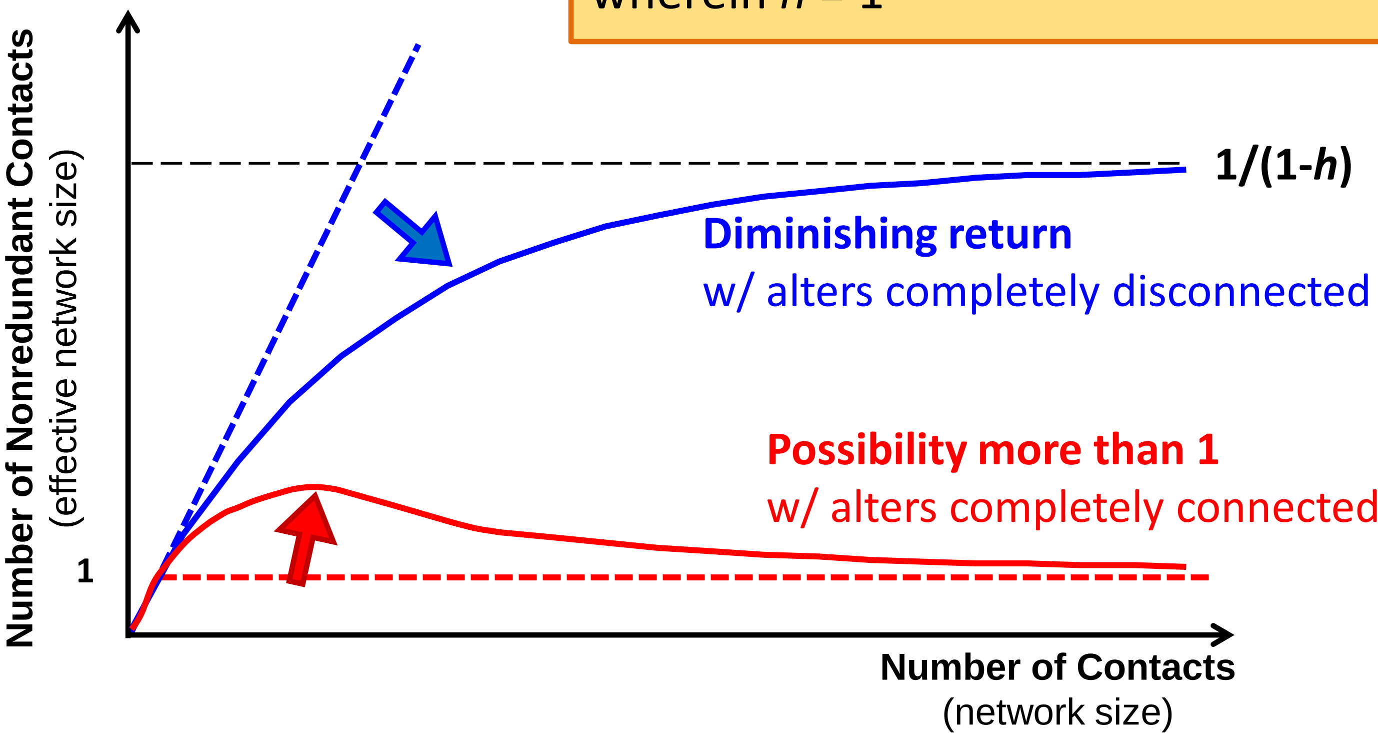
Put these together…



Put these together…

Introduction of ***h***

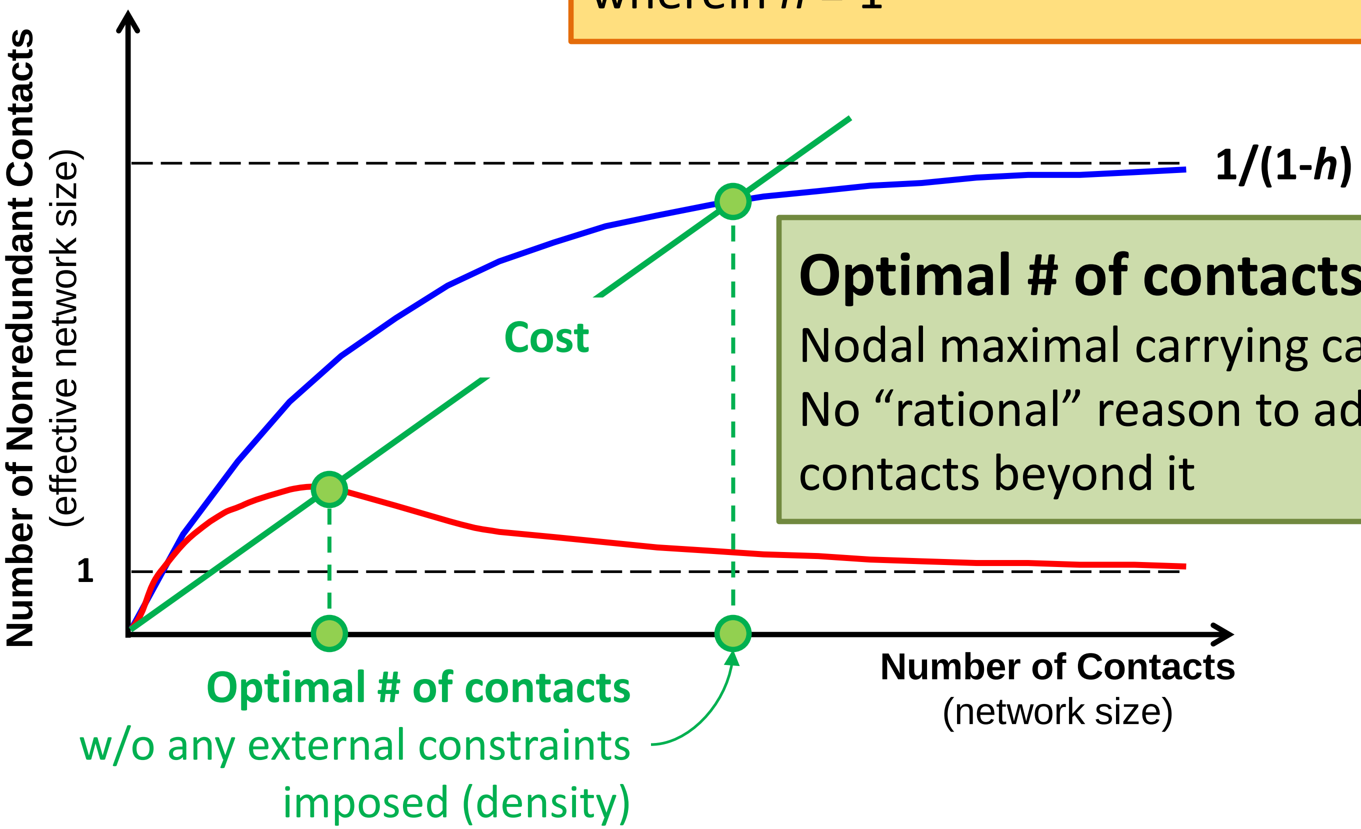
Not rejection of Burt's formulation;
But ***generalization*** beyond extreme cases wherein *h* = 1



Put these together…

Introduction of ***h***

Not rejection of Burt's formulation;
But ***generalization*** beyond extreme cases wherein *h* = 1



Optimal # of contacts

Nodal maximal carrying capacity
No "rational" reason to add contacts beyond it

Theoretical Connection to Nodal Carrying Capacity

Nodal carrying capacity

- "Recent work by Leskovec, Kleinberg, and Faloutsos (2005, 2007) has identified a network process called *densification* wherein the average number of links per node, what might be considered the *nodal carrying capacity*, tends to grow as the number of nodes in the network grows over time."

Monge, et al., (2008), p. 456

Degree Power Law: Densification

(Leskovec, et al., 2007)

$$e(t) \propto n(t)^{\alpha}$$

- *e*: the number of links; *k*: constant
- Exponent α: $1 < \alpha < 2$ in the real-world networks

- They argue that as the network size increases…
  (1) Average number of links (not density) increases
  (2) Average distance among nodes decreases

- By *densification*, they mean *increasing closeness* not network density

Degree Power Law: Densification

(Leskovec, et al., 2007)

- Rewritten as

$$e = k \cdot n^{\alpha}$$

- k: constant
- e/n = average # of links
- $e/\{n(n-1)\}$ = network density; e/n^2 for large networks

Degree Power Law: Densification

(Leskovec, et al., 2007)

$$e = k \cdot n^{\alpha} \; (1 < \alpha < 2)$$

- When $\alpha = 1$; $e = k \cdot n$
- $e/n = k$
- Average # of links is ***constant***, regardless of the network size, *n*
- $e/\{n(n-1)\} = k/n$
- Density *decreases*
- **"Cost-dependent"**

- When $\alpha = 2$; $e = k \cdot n^2$
- $e/n = k \cdot n$
- Average # of links ***increases*** *proportionally* to the network size, *n*
- $e/\{n(n-1)\} = k$
- Density remains *constant*
- **"Cost-free"**

- Real-world networks
- *e*/*n* **increases** as *n* increases, but **not proportionally/linearly**

WHY NOT? Moderated by "heterogeneity"

Basic Statistics Quiz

n_1

n_2

$N = n_2 + n_2$

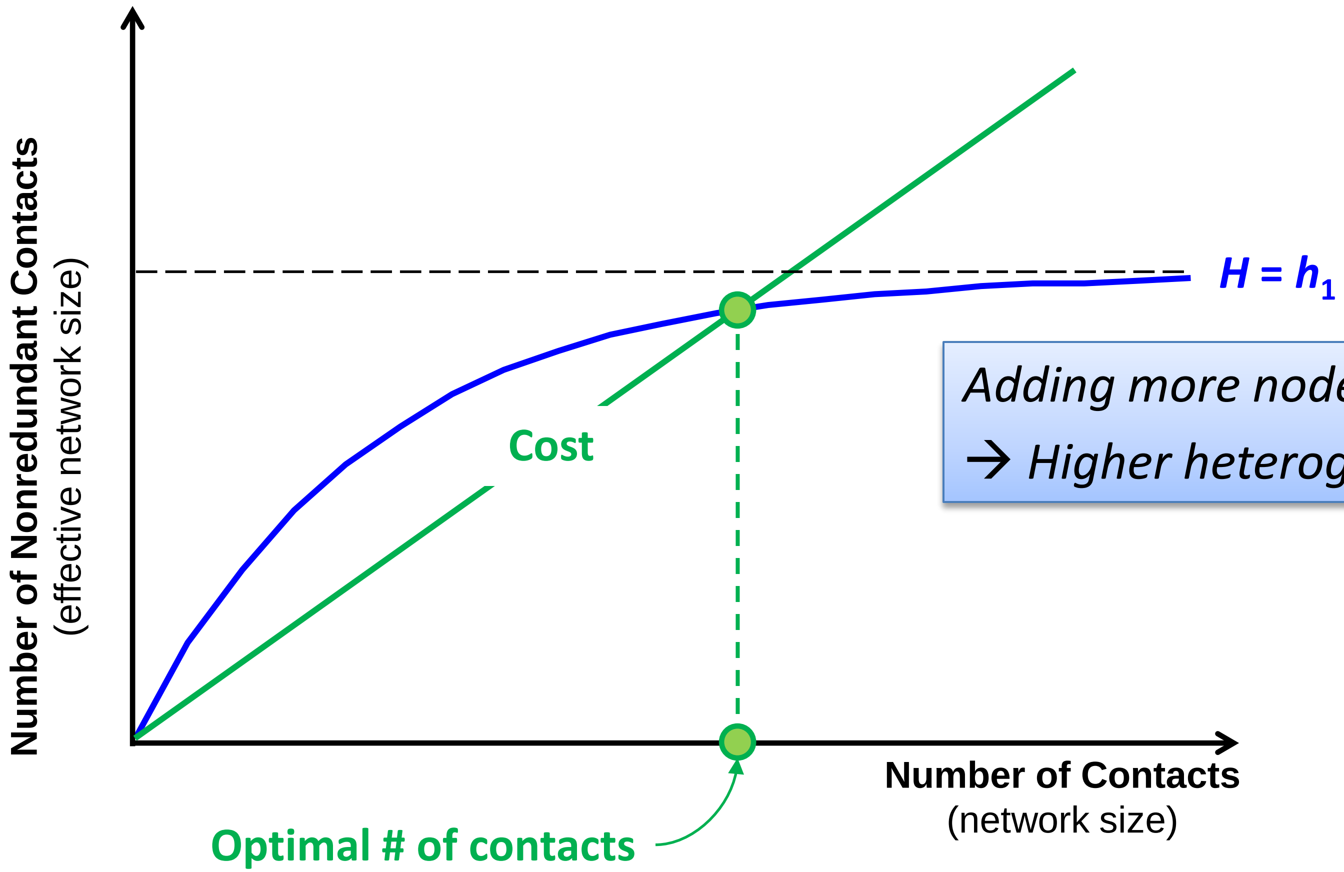
- A random sample (size = n_1)
- The variable of interest: X
- V_1 = Var(X)

- Add another random sample (size = n_2) selected from the same population
- Now the total size N is $n_1 + n_2$
- **What's the variance of X?**

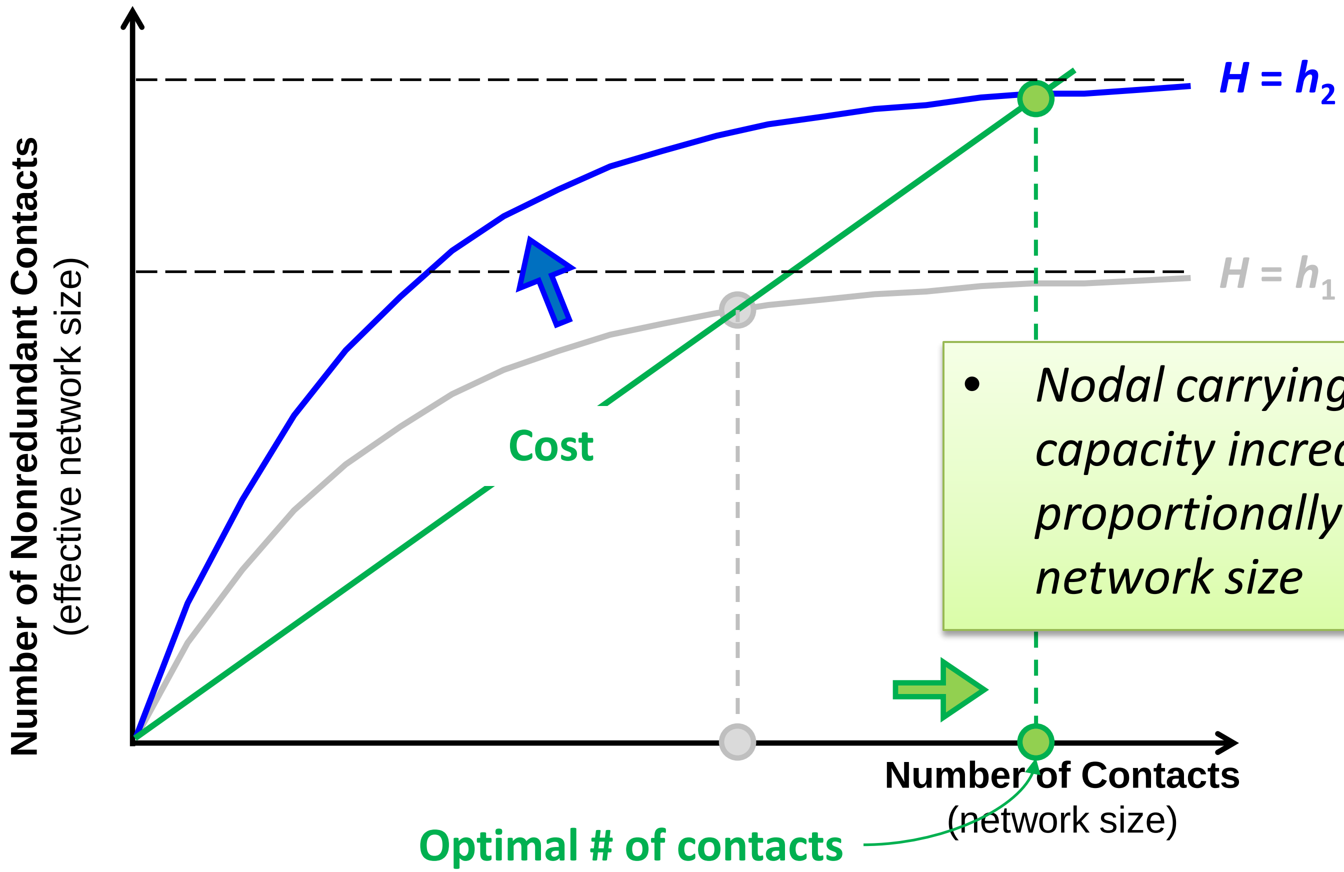
- $V = V_1 (n_1 + n_2)/ n_1$

- As the size increases, the variance (or *heterogeneity*) increases

As net size increases, *h* also increases



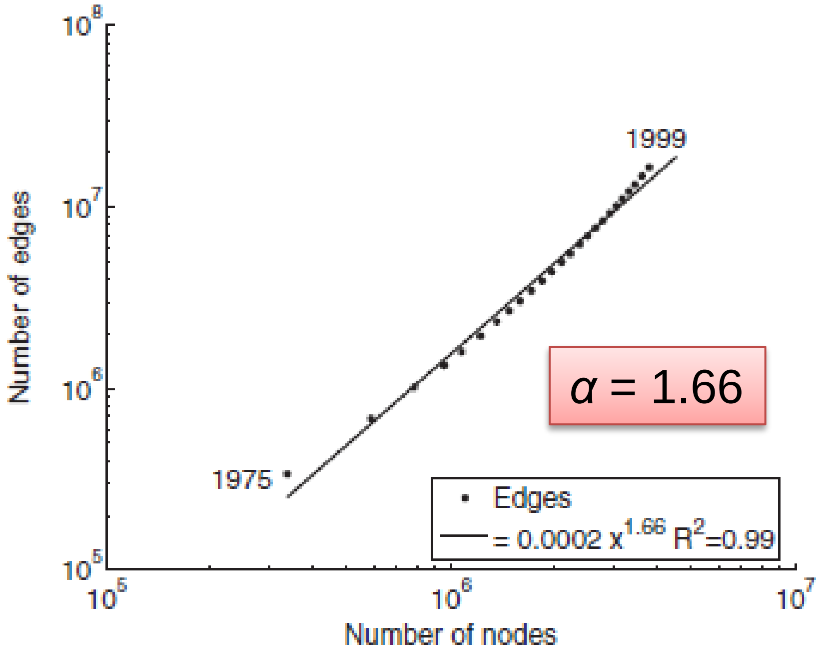
As net size increases, *h* also increases



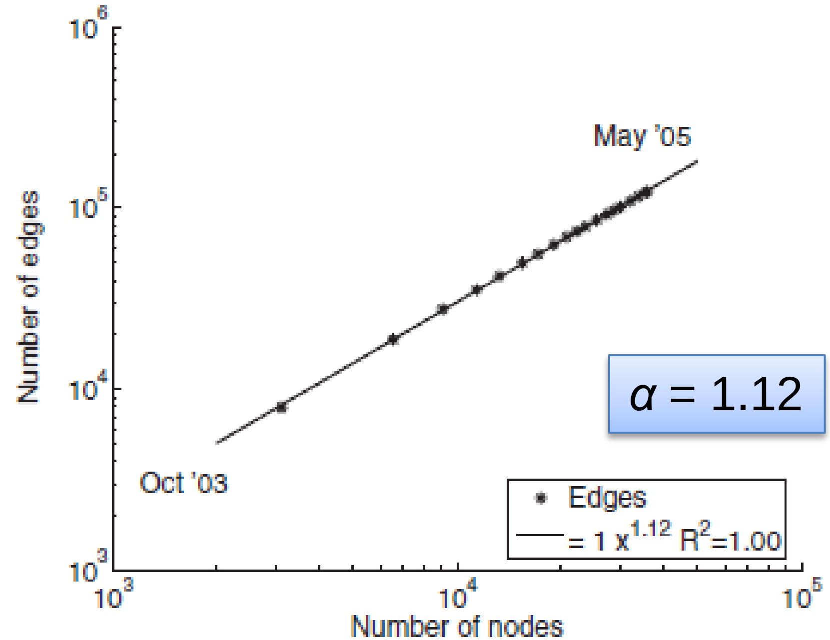
- *Nodal carrying capacity increases not proportionally to the network size*

Difference in growth rate

(Leskovec, et al., 2007: Figure 2, p. 10)



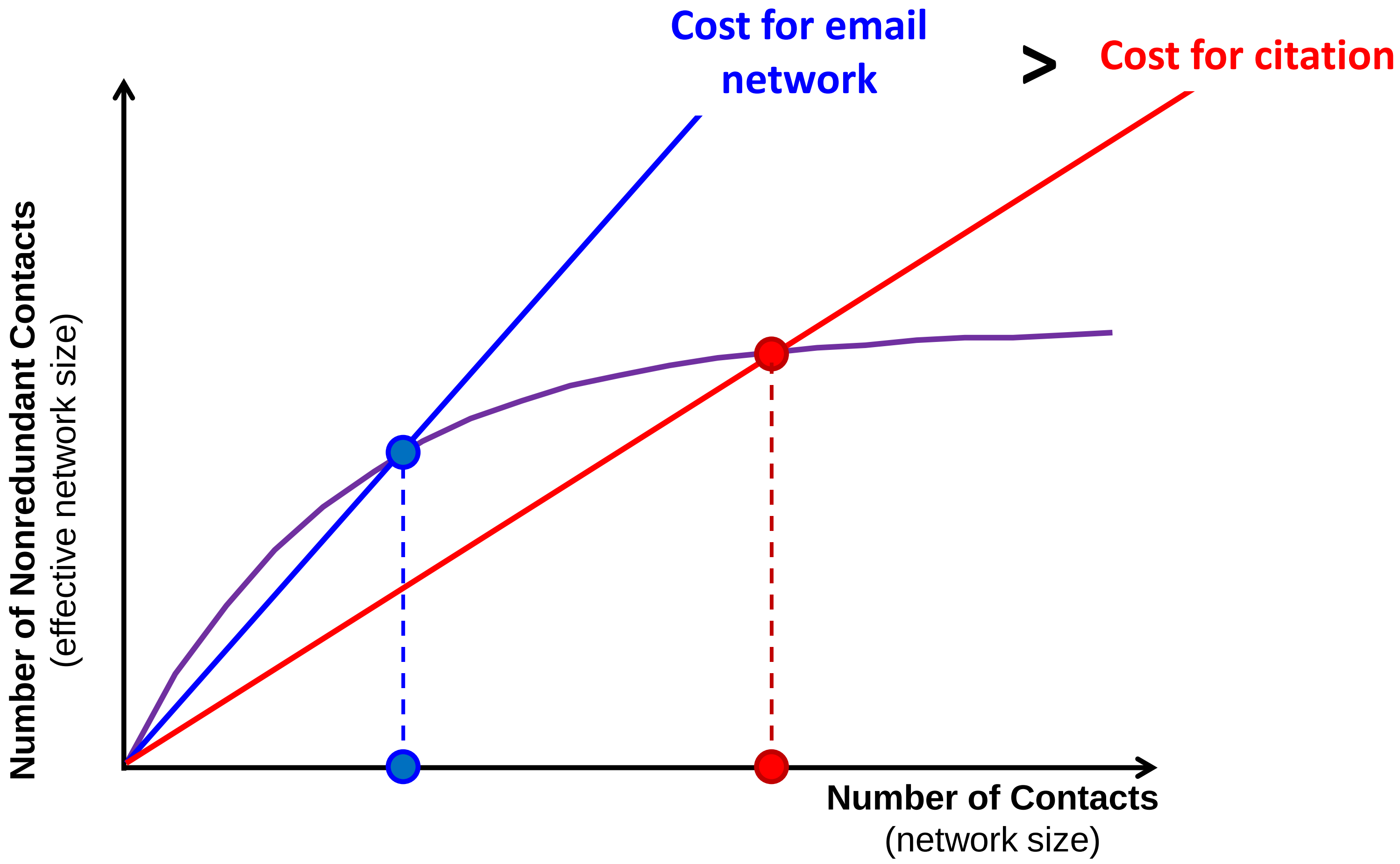
Patent Citation



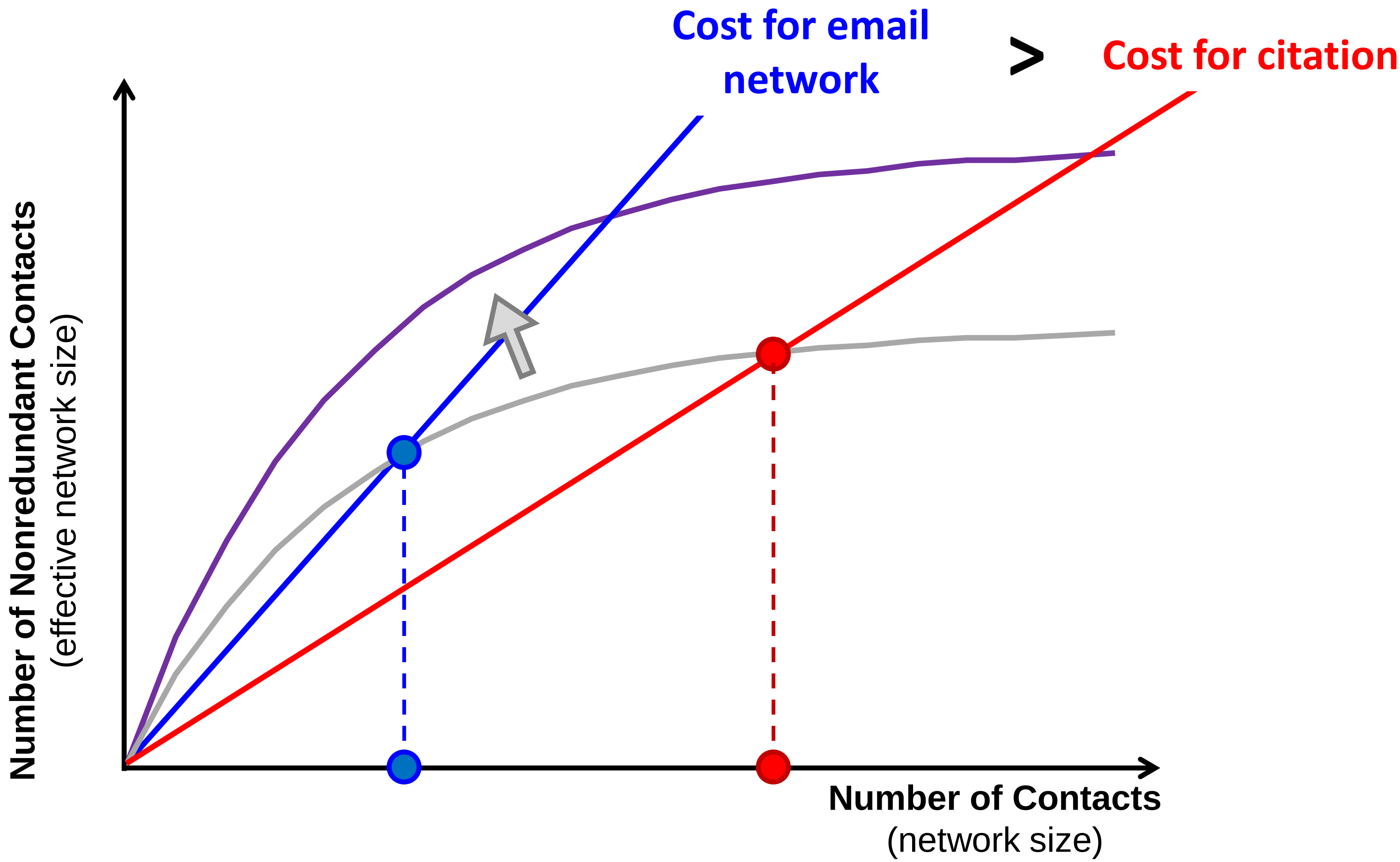
Email Network

- "Different kinds of empirical networks display different link density growth rates" (Monge, et al., 2008, p. 457)
- ***Why do the exponents differ from each other?***

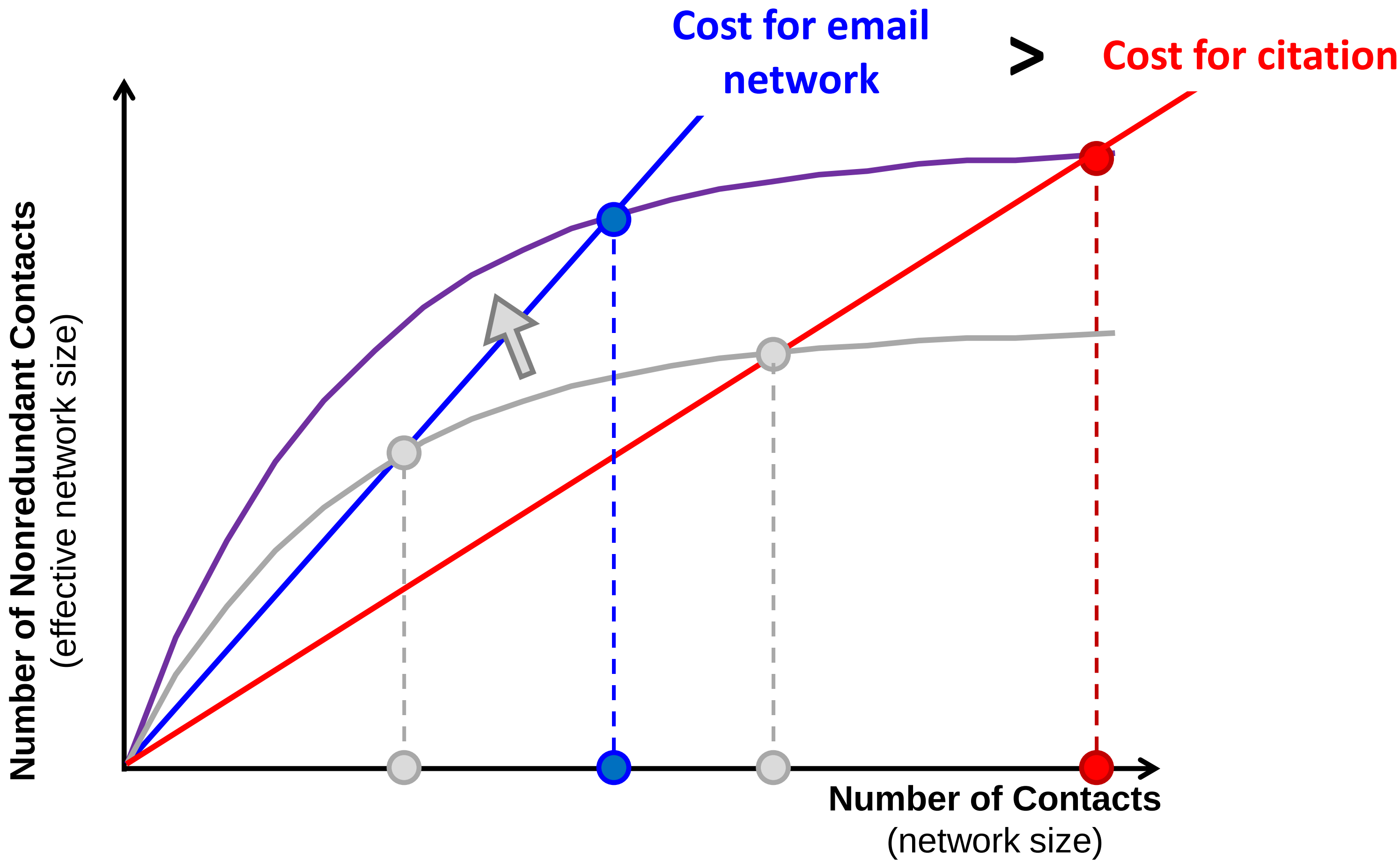
Difference in growth rate



Difference in growth rate

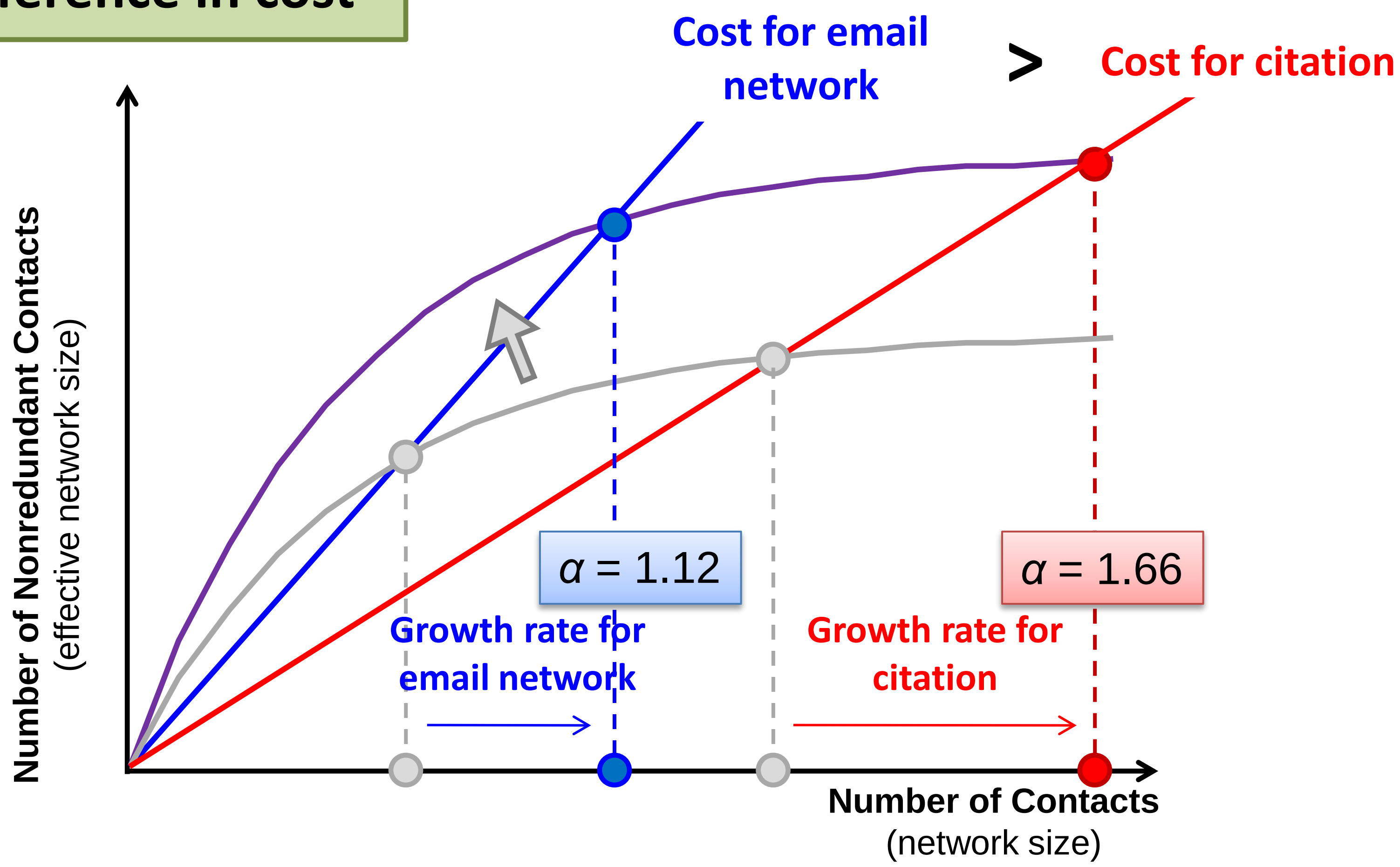


Difference in growth rate



Difference in growth rate

↑ **Difference in cost**



To conclude

- Heterogeneity of network members ***h*** might be **trivial** or **insignificant**

- But, it has a **non-trivial impact** on an adaptive network, where individual members add or delete ties for their maximal profits.

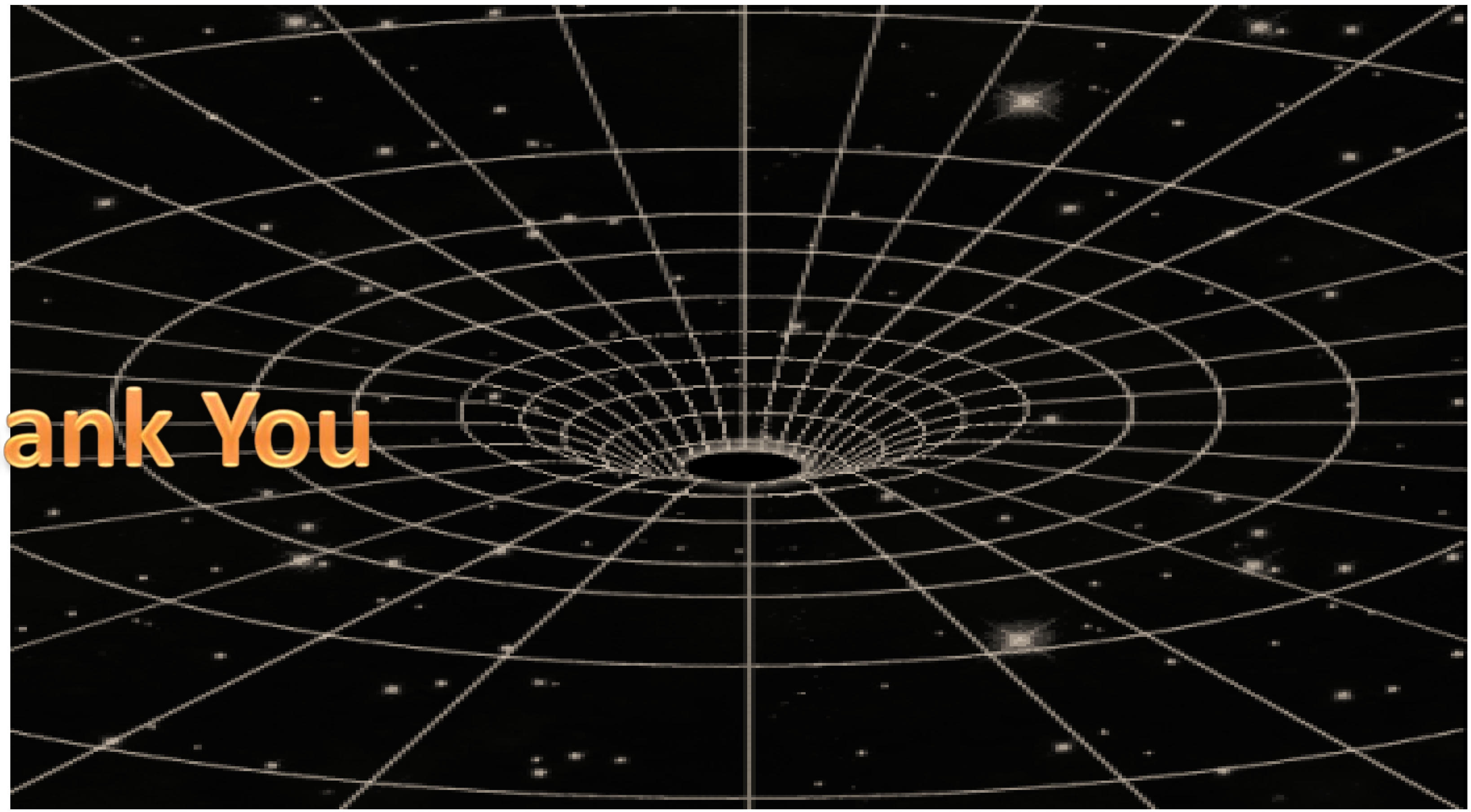
- It allows us to identify "more realistic" equilibrium states, which can be interpreted as **the maximal carrying capacities of individual nodes**

Next Steps

- **How to measure heterogeneity?**
  - **Dyadic level:** e.g., structural equivalence?
  - **Global level:** a single indicator for whole population in a network
    - The determinant of structural equivalence matrix?
    - Variances of variables of interest?

- **Empirical evidence** showing the non-trivial impacts of heterogeneity on the evolution of adaptive and dynamic networks

Let’s get into the holes deeper!



Thank You