Network Modeling: Event Networks

Viviana Amati   Jürgen Lerner   Bobo Nick

Dept. Computer & Information Science
University of Konstanz

Winter 2011/2012
(last updated: February 1, 2012)

Where are we?

Recall.

Introduction

- availability of massive network datasets
- individuals are (seemingly?) influenced by their friends
- challenge: social selection vs. social influence

Static network models no time-information

- can generate networks looking like social networks
- can test/validate social correlation
- by design, cannot resolve selection vs. influence

SAOM for networks observed at several time points

- can resolve the selection vs. influence question, but
- computationally expensive; few actors / time steps
- inappropriate for typical "Web-based" network data that comes as sequences of **interaction events**.

Recall.

Introduction

- availability of massive network datasets
- individuals are (seemingly?) influenced by their friends
- challenge: social selection vs. social influence

Static network models no time-information

- can generate networks looking like social networks
- can test/validate social correlation
- by design, cannot resolve selection vs. influence

SAOM for networks observed at several time points

- can resolve the selection vs. influence question, but
- computationally expensive; few actors / time steps
- inappropriate for typical “Web-based” network data that comes as sequences of **interaction events**.

Recall.

Introduction

- availability of massive network datasets
- individuals are (seemingly?) influenced by their friends
- challenge: social selection vs. social influence

Static network models no time-information

- can generate networks looking like social networks
- can test/validate social correlation
- by design, cannot resolve selection vs. influence

SAOM for networks observed at several time points

- can resolve the selection vs. influence question, but
- computationally expensive; few actors / time steps
- inappropriate for typical "Web-based" network data that comes as sequences of **interaction events.**

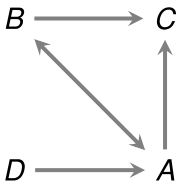
Event network data.

Data: time-stamped interaction among social actors

Examples: email, usenet, chat, telephone calls, collaboration in wikis, social bookmark systems, social networking sites, . . .

t_1	A	`calls`	B
t_2	B	`calls`	C
t_3	A	`calls`	C
t_4	B	`calls`	A
t_5	A	`calls`	C
t_6	D	`calls`	A

$\dots$



$T_1^{(ac)} = t_3,\ T_2^{(ac)} = t_5, \dots$

Model the **event times** on each dyad dependent on

- previous events on the same/reverse/incident dyads;
- exogenous actor and dyad characteristics.

Data input format.

Given a set V of actors (vertices of the network).

Input data: sequence of events $E = (e_1, \ldots, e_N)$, where each event $e \in E$ is a tuple $e = (u, v, t)$ with

- $u \in V$ **source** actor (initiator);
- $v \in V$ **target** actor (recipient);
- $t \in \mathbb{R}$ **time**, when e happened;
- sometimes there is additional information coding the **type** or strength of the event.

Set of dyads: $D = V \times V \setminus \{(v, v)\,;\; v \in V\}$ (directed graphs).

Assume that no events happen at the same point in time.

Sequence E is in increasing order with respect to time (from past to future).

Data input format.

Given a set V of actors (vertices of the network).

Input data: sequence of events $E = (e_1, \ldots, e_N)$, where each event $e \in E$ is a tuple $e = (u, v, t)$ with

- $u \in V$ **source** actor (initiator);
- $v \in V$ **target** actor (recipient);
- $t \in \mathbb{R}$ **time**, when e happened;
- sometimes there is additional information coding the **type** or strength of the event.

Set of dyads: $D = V \times V \setminus \{(v, v)\,;\ v \in V\}$ (directed graphs).

Assume that no events happen at the same point in time.

Sequence E is in increasing order with respect to time (from past to future).

Specification of the probability density.

We want to specify a probability density f for sequences of events $E = (e_1, \ldots, e_N)$.

Remark: f should be indexed by N; this is ignored to keep notation concise.

We decompose f into conditional probability densities

$$f(E) = f(e_1) \cdot f(e_2|e_1) \cdot f(e_3|e_1, e_2) \ldots f(e_N|e_1, \ldots, e_{N-1})$$

Motivation: the past events $e_1, \ldots, e_{i-1}$ influence the distribution of the next event e_i.

Thus, it suffices to specify the conditional distribution of the next event e_i, given the past events $e_1, \ldots, e_{i-1}$

$$f(e_i|e_1, \ldots, e_{i-1})$$

Waiting time T_{uv} to the next event.

Let $e_i = (u, v, t)$ be the next event, let $E_{<i} = (e_1, \ldots, e_{i-1})$ denote the sequence of past events, and assume for simplicity that the last previous event e_{i-1} happened at $t = 0$.

For the moment, take it for granted that e_i happens on the dyad (u, v). Let T_{uv} denote the random variable for the time of e_i.

Want to specify a model in which certain aspects of $E_{<i}$ stochastically cause e_i to happen rather early or rather late. For instance,

- if there are many events from u to v in the past, we expect T_{uv} to be rather small (short waiting time);
- if u and v have past events with a common third actor w, we again expect T_{uv} to be rather small, etc.

Waiting time T_{uv} to the next event.

Let $e_i = (u, v, t)$ be the next event, let $E_{<i} = (e_1, \ldots, e_{i-1})$ denote the sequence of past events, and assume for simplicity that the last previous event e_{i-1} happened at $t = 0$.

For the moment, take it for granted that e_i happens on the dyad (u, v). Let T_{uv} denote the random variable for the time of e_i.

Want to specify a model in which certain aspects of $E_{<i}$ stochastically cause e_i to happen rather early or rather late. For instance,

- if there are many events from u to v in the past, we expect T_{uv} to be rather small (short waiting time);
- if u and v have past events with a common third actor w, we again expect T_{uv} to be rather small, etc.

Waiting time T_{uv} to the next event.

Let $e_i = (u, v, t)$ be the next event, let $E_{<i} = (e_1, \ldots, e_{i-1})$ denote the sequence of past events, and assume for simplicity that the last previous event e_{i-1} happened at $t = 0$.

For the moment, take it for granted that e_i happens on the dyad (u, v). Let T_{uv} denote the random variable for the time of e_i.

Want to specify a model in which certain aspects of $E_{<i}$ stochastically cause e_i to happen rather early or rather late. For instance,

- if there are many events from u to v in the past, we expect T_{uv} to be rather small (short waiting time);
- if u and v have past events with a common third actor w, we again expect T_{uv} to be rather small, etc.

Some functions involving the distribution of T_{uv}.

Survival function (probability that e_i has not happened before or at t)

$$t \mapsto S(t) := P(t \leq T_{uv})$$

Failure function (probability that e_i has happened at t or earlier)

$$t \mapsto F(t) := P(T_{uv} \leq t) = 1 - S(t)$$

Probability density for T_{uv}

$$t \mapsto f(t) := \lim_{\Delta t \to 0} \frac{P(t \leq T_{uv} \leq t + \Delta t)}{\Delta t} = \frac{d}{dt} F(t)$$

Hazard function (conditional probability density of e_i happening at t, given that it did not happen before)

$$t \mapsto \lambda(t) := \frac{f(t)}{S(t)} = \lim_{\Delta t \to 0} \frac{P(t \leq T_{uv} \leq t + \Delta t \,|\, t \leq T_{uv})}{\Delta t}$$

Each of these functions determines the other three; specifying λ yields the most intuitive interpretation.

To give some intuition about the meaning of these functions,

consider the case where T refers to the time at which an individual dies.

Want to model the distribution of T dependent on risk-behavior of the individuals (e. g., smoking, drug using, climbing).

Intuition about S and F.

Survival function

$$t \mapsto S(t) := P(t \leq T)$$

Probability of being alive at time t. Expected to be lower for risk-takers.

Failure function

$$t \mapsto F(t) := P(T \leq t) = 1 - S(t)$$

Probability of being dead at time t. Expected to be higher for risk-takers.

Both functions have constraints imposed by their meaning:

- $S(0) = 1$;
- S is monotonically decreasing;
- $\lim_{t \to \infty} S(t) = 0$

Intuition about the probability density f.

Probability density for T

$$t \mapsto f(t) := \lim_{\Delta t \to 0} \frac{P(t \leq T \leq t + \Delta t)}{\Delta t} = \frac{d}{dt} F(t)$$

More intuitive if we consider the probability for a fixed time unit $\Delta t = 1$ (e. g., year).

$$P(T = t) \qquad \text{probability of dying in year } t$$

Is this higher or lower for risk-takers?

Intuition about the hazard function λ.

Hazard function

$$t \mapsto \lambda(t) := \frac{f(t)}{S(t)} = \lim_{\Delta t \to 0} \frac{P(t \leq T \leq t + \Delta t \,|\, t \leq T)}{\Delta t}$$

More intuitive if we consider the probability for a fixed time unit $\Delta t = 1$ (e. g., year).

$$P(T = t \,|\, t \leq T)$$

Probability of dying in year t restricted to those individual that reach this age.

Expected to be higher for risk-takers.

Only constraint: $\lambda(t) > 0$.

The hazard λ determines the density f.

Probability density function

$$
\begin{aligned}
f(t) &= \lim_{\Delta t \to 0} \frac{Pr(t \le T \le t + \Delta t)}{\Delta t} \\
&= \lim_{\Delta t \to 0} \frac{F(t + \Delta t) - F(t)}{\Delta t} = \frac{d}{dt} F(t) = -\frac{d}{dt} S(t)
\end{aligned}
$$

It is

$$
\lambda(t) = \frac{f(t)}{S(t)} = -\frac{\frac{d}{dt} S(t)}{S(t)} = -\frac{d}{dt} \log S(t)
$$

Considering that $S(0) = 1$ it follows

$$
S(t) = \exp\left(-\int_0^t \lambda(\tau)\, d\tau\right) .
$$

Thus, $f(t) = \lambda(t) \cdot \exp\left(-\int_0^t \lambda(\tau)\, d\tau\right)$

The hazard λ determines the density f.

Probability density function

$$
\begin{aligned}
f(t) &= \lim_{\Delta t \to 0} \frac{Pr(t \leq T \leq t + \Delta t)}{\Delta t} \\
&= \lim_{\Delta t \to 0} \frac{F(t + \Delta t) - F(t)}{\Delta t} = \frac{d}{dt} F(t) = -\frac{d}{dt} S(t)
\end{aligned}
$$

It is

$$\lambda(t) = \frac{f(t)}{S(t)} = -\frac{\frac{d}{dt} S(t)}{S(t)} = -\frac{d}{dt} \log S(t)$$

Considering that $S(0) = 1$ it follows

$$S(t) = \exp\left(-\int_0^t \lambda(\tau)\, d\tau\right) .$$

Thus, $f(t) = \lambda(t) \cdot \exp\left(-\int_0^t \lambda(\tau)\, d\tau\right)$

The hazard λ determines the density f.

Probability density function

$$f(t) = \lim_{\Delta t \to 0} \frac{Pr(t \leq T \leq t + \Delta t)}{\Delta t}$$
$$= \lim_{\Delta t \to 0} \frac{F(t + \Delta t) - F(t)}{\Delta t} = \frac{d}{dt} F(t) = -\frac{d}{dt} S(t)$$

It is

$$\lambda(t) = \frac{f(t)}{S(t)} = -\frac{\frac{d}{dt} S(t)}{S(t)} = -\frac{d}{dt} \log S(t)$$

Considering that $S(0) = 1$ it follows

$$S(t) = \exp\left(-\int_0^t \lambda(\tau)\, d\tau\right) .$$

Thus, $f(t) = \lambda(t) \cdot \exp\left(-\int_0^t \lambda(\tau)\, d\tau\right)$

The hazard λ determines the density f.

Probability density function

$$f(t) = \lim_{\Delta t \to 0} \frac{Pr(t \leq T \leq t + \Delta t)}{\Delta t}$$
$$= \lim_{\Delta t \to 0} \frac{F(t + \Delta t) - F(t)}{\Delta t} = \frac{d}{dt} F(t) = -\frac{d}{dt} S(t)$$

It is

$$\lambda(t) = \frac{f(t)}{S(t)} = -\frac{\frac{d}{dt} S(t)}{S(t)} = -\frac{d}{dt} \log S(t)$$

Considering that $S(0) = 1$ it follows

$$S(t) = \exp\left(-\int_0^t \lambda(\tau)\, d\tau\right) .$$

Thus, $f(t) = \lambda(t) \cdot \exp\left(-\int_0^t \lambda(\tau)\, d\tau\right)$

The hazard λ determines the density f.

Probability density function

$$
\begin{aligned}
f(t) &= \lim_{\Delta t \to 0} \frac{Pr(t \leq T \leq t + \Delta t)}{\Delta t} \\
&= \lim_{\Delta t \to 0} \frac{F(t + \Delta t) - F(t)}{\Delta t} = \frac{d}{dt} F(t) = -\frac{d}{dt} S(t)
\end{aligned}
$$

It is

$$\lambda(t) = \frac{f(t)}{S(t)} = -\frac{\frac{d}{dt} S(t)}{S(t)} = -\frac{d}{dt} \log S(t)$$

Considering that $S(0) = 1$ it follows

$$S(t) = \exp\left(-\int_0^t \lambda(\tau)\, d\tau \right) .$$

Thus, $f(t) = \lambda(t) \cdot \exp\left(-\int_0^t \lambda(\tau)\, d\tau \right)$

The hazard λ determines the density f.

Probability density function

$$\begin{aligned} f(t) &= \lim_{\Delta t \to 0} \frac{Pr(t \leq T \leq t + \Delta t)}{\Delta t} \\ &= \lim_{\Delta t \to 0} \frac{F(t + \Delta t) - F(t)}{\Delta t} = \frac{d}{dt} F(t) = -\frac{d}{dt} S(t) \end{aligned}$$

It is

$$\lambda(t) = \frac{f(t)}{S(t)} = -\frac{\frac{d}{dt} S(t)}{S(t)} = -\frac{d}{dt} \log S(t)$$

Considering that $S(0) = 1$ it follows

$$S(t) = \exp\left(-\int_0^t \lambda(\tau)\, d\tau\right) .$$

Thus, $f(t) = \lambda(t) \cdot \exp\left(-\int_0^t \lambda(\tau)\, d\tau\right)$

The hazard λ determines the density f.

Probability density function

$$f(t) = \lim_{\Delta t \to 0} \frac{Pr(t \leq T \leq t + \Delta t)}{\Delta t}$$
$$= \lim_{\Delta t \to 0} \frac{F(t + \Delta t) - F(t)}{\Delta t} = \frac{d}{dt} F(t) = -\frac{d}{dt} S(t)$$

It is

$$\lambda(t) = \frac{f(t)}{S(t)} = -\frac{\frac{d}{dt} S(t)}{S(t)} = -\frac{d}{dt} \log S(t)$$

Considering that $S(0) = 1$ it follows

$$S(t) = \exp\left(-\int_0^t \lambda(\tau)\, d\tau \right) \quad .$$

Thus, $f(t) = \lambda(t) \cdot \exp\left(-\int_0^t \lambda(\tau)\, d\tau \right)$

The hazard λ determines the density f.

Probability density function

$$
\begin{aligned}
f(t) &= \lim_{\Delta t \to 0} \frac{Pr(t \le T \le t + \Delta t)}{\Delta t} \\
&= \lim_{\Delta t \to 0} \frac{F(t + \Delta t) - F(t)}{\Delta t} = \frac{d}{dt} F(t) = -\frac{d}{dt} S(t)
\end{aligned}
$$

It is

$$\lambda(t) = \frac{f(t)}{S(t)} = -\frac{\frac{d}{dt} S(t)}{S(t)} = -\frac{d}{dt} \log S(t)$$

Considering that $S(0) = 1$ it follows

$$S(t) = \exp\left(-\int_0^t \lambda(\tau)\, d\tau\right) \ .$$

Thus, $f(t) = \lambda(t) \cdot \exp\left(-\int_0^t \lambda(\tau)\, d\tau\right)$

Specifying the hazard function λ.

Hazard function (conditional probability density of e_i happening at t, given that it did not happen before)

$$t \mapsto \lambda(t) := \frac{f(t)}{S(t)}$$

The hazard is modeled as a function of statistics s_ℓ

$$\lambda(t) = \exp\left(\sum_{\ell=1}^{k} \theta_\ell \cdot s_\ell(t)\right)$$

with parameters θ_ℓ specifying increase/decrease of the hazard.

The statistics s_ℓ quantify specific aspects of the past events $E_{<i}$ with respect to the current dyad (u, v). For instance,

- average frequency of past events from u to v;
- number of third actors w such that both u and v had an event with w, etc.

Probability density of the event $e_i = (u, v, t)$.

Remember: so far we derived the density conditional on the fact that e_i happens on the dyad (u, v); but it could also have happened on any other dyad $(u', v') \neq (u, v)$.

The probability that the next event on (u', v') does not happen before t is

$$S_{u'v'}(t) = \exp\left(-\int_0^t \lambda_{u'v'}(\tau)\, d\tau \right)$$

Thus the joint density of the event e_i is

$$f(e_i | E_{<i}) = \lambda_{uv}(t) \cdot S_{uv}(t) \cdot \prod_{u'v' \neq uv} S_{u'v'}(t) = \lambda_{uv}(t) \cdot \prod_{u'v' \in D} S_{u'v'}(t) \quad .$$

Probability density of the event sequence.

The joint probability density for sequence of dyadic events $E = (e_1, \dots, e_N)$ with $e_i = (u_i, v_i, t_i)$ is

(assuming that the observation started at time $t_0 < t_1$)

$$\begin{aligned} f(e_1, \dots, e_N) &= \prod_{i=1}^{N} f(e_i | E_{<i}) \\ &= \prod_{i=1}^{N} \lambda_{u_i v_i}(t_i) \cdot \prod_{uv \in D} S_{uv}^{(t_{i-1})}(t_i) \\ &= \prod_{i=1}^{N} \lambda_{u_i v_i}(t_i) \cdot \exp\left(- \sum_{uv \in D} \int_{t_{i-1}}^{t_i} \lambda_{uv}(\tau)\, d\tau \right) \\ \text{with} \quad \lambda_{uv}(\tau) &= \exp\left(\sum_{\ell=1}^{k} \theta_\ell \cdot s_\ell(\tau; u, v) \right) . \end{aligned}$$

It remains to specify the statistics as a function of the previous events $E_{<i}$.

Defining the statistics $s_\ell(t, u, v)$.

The statistics $s_\ell(t, u, v)$ are a function of the past events

$$E_{<t} := \{e = (u', v', t') \in E\,;\ t' < t\}$$

Note that $s_\ell(t, u, v)$ might depend on events that happen on other dyads $(u', v') \neq (u, v)$.

General definition applies two steps:

1. a dynamic network $G_t = (V, D, a_t)$ is built from $E_{<t}$;
   - V is the set of actors and D the set of dyads;
   - $a_t : D \to \mathbb{R}$ encodes past activity on dyads;
   - when an event (u, v, t) happens, the value of $a_t(u, v)$ changes;
   - when time moves forward, all edge weights a_t decrease;
2. the statistics $s_\ell(t, u, v)$ are functions of the edge weights a_t.

Definition of the edge weights a_t.

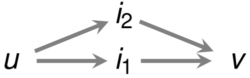
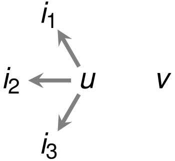
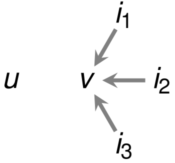
The weight $a_t(u,v)$ at time t on dyad (u,v) is a function of those events happening on (u,v) before t.

$$E_{<t;uv} = \{e = (u',v',t') \in E\,;\; u' = u, v' = v, t' < t\}$$

The *halflife* $T_{1/2} > 0$ controls how fast the influence of past events diminuishes. The edge weights are defined by

$$a_t(u,v) = \sum_{e \in E_{<t;uv}} 1 \cdot \exp\left(-(t - t_e) \cdot \frac{\ln(2)}{T_{1/2}}\right)$$

Definition of selected statistics $s_\ell(t, u, v)$.

name	formula	$u \longrightarrow v$ depends on
`inertia`	$a_t(u, v)$	$u \longrightarrow v$
`recipr.`	$a_t(v, u)$	$u \longleftarrow v$
`trans.`	$\sqrt{\sum_{i \in V} a_t(u, i) \cdot a_t(i, v)}$	
`outDegSrc`	$\sum_{i \in V} a_t(u, i)$	
`inDegTgt`	$\sum_{i \in V} a_t(i, v)$	

Application example:
conflict and cooperation in political event networks.

Data from the Kansas Event Data System.

Tool that generates sequences of events from news reports.

Event $e = (u, v, t, w)$ encodes

- u **source** actor (initiator);
- v **target** actor (recipient);
- t **timestamp**, the day on which e happened;
- w **weight** from -10 (most hostile) to $+10$ (most cooperative).

Examples of event-types and their associated weights:

OPTIMIST COMMENT	0.4	PESSIMIST COMMENT	-0.4
VISIT	1.9	ACCUSE	-2.0
PROMISE	4.0	REJECT	-4.0
AGREE	6.0	THREATEN	-6.0
EXTEND MIL AID	8.3	MILITARY DEMO	-7.6
MERGE, INTEGRATE	10.0	MILITARY ENGAGEMENT	-10.0

Data from the Kansas Event Data System.

Tool that generates sequences of events from news reports.

Event $e = (u, v, t, w)$ encodes

- u **source** actor (initiator);
- v **target** actor (recipient);
- t **timestamp**, the day on which e happened;
- w **weight** from -10 (most hostile) to $+10$ (most cooperative).

Examples of event-types and their associated weights:

OPTIMIST COMMENT	0.4	PESSIMIST COMMENT	-0.4
VISIT	1.9	ACCUSE	-2.0
PROMISE	4.0	REJECT	-4.0
AGREE	6.0	THREATEN	-6.0
EXTEND MIL AID	8.3	MILITARY DEMO	-7.6
MERGE, INTEGRATE	10.0	MILITARY ENGAGEMENT	-10.0

Data from the Kansas Event Data System.

http://eventdata.psu.edu/

Sequences of daily events from news reports.

region	time period	actors	events
GULF	1979/04/15 – 1999/03/31	202	304,000
LEVANT	1991/05/05 – 2007/01/31	699	171,000
BALKANS	1989/04/02 – 2003/07/31	325	78,000
TURKEY	1992/01/03 – 2006/07/31	429	20,000

```
800923 IRQ IRN 223 MIL ENGAGEME
800923 IRN IRQ 223 MIL ENGAGEME
800924 USA IRN 121 CRITICIZE
800924 USA IRQ 121 CRITICIZE
800924 USA IRN 192 CUT ROUTINE
800924 USA IRQ 192 CUT ROUTINE
800924 UNO IRN 095 PLEAD
```

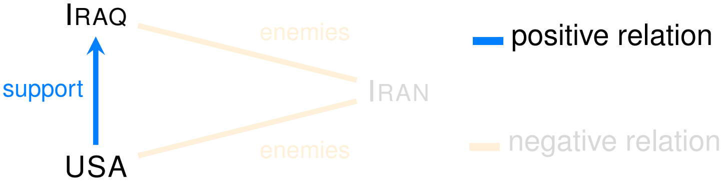
```
800924 UNO IRQ 095 PLEAD
800924 IRQ IRN 222 NONMIL DESTR
800924 IRN IRQ 223 MIL ENGAGEME
800924 IRQ IRN 122 DENIGRATE
800924 IRN IRQ 122 DENIGRATE
800924 IRQ IRN 031 MEET
...
```

Examplary hypothesis to be tested.

Structural balance theory do actors collaborate with the enemies of their enemies, fight the enemies of their friends, . . .

Anecdotal illustration of structural balance.

In the 1980s the USA provided support to the Iraq, although Iraq is not a typical ally of the USA.

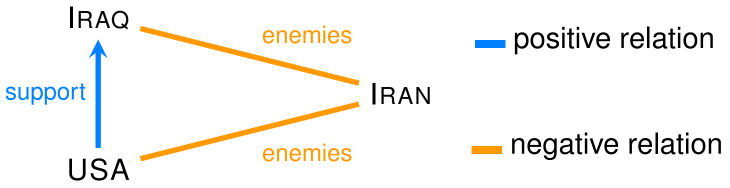


Potential explanation: USA supported enemy of an enemy.

Here: statistical tests of structural balance theory in event data.

Anecdotal illustration of structural balance.

In the 1980s the USA provided support to the Iraq, although Iraq is not a typical ally of the USA.

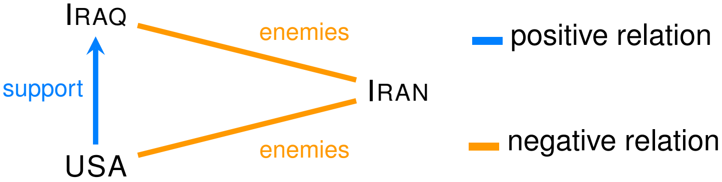


Potential explanation: USA supported enemy of an enemy.

Here: statistical tests of structural balance theory in event data.

Anecdotal illustration of structural balance.

In the 1980s the USA provided support to the Iraq, although Iraq is not a typical ally of the USA.



Potential explanation: USA supported enemy of an enemy.

Here: statistical tests of structural balance theory in event data.

Hypotheses derived from structural balance theory.

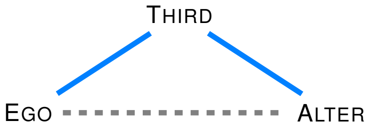
(Heider, 1946; Cartwright and Harary, 1956)

- $\overbrace{\text{The friend of my friend}}^{\textit{explanatory variable}}$ is my $\overbrace{\text{friend}}^{\textit{dep. var.}}$.
- The enemy of my friend is my enemy.
- The friend of my enemy is my enemy.
- The enemy of my enemy is my friend.

Hypotheses derived from structural balance theory.

(Heider, 1946; Cartwright and Harary, 1956)

- $\overbrace{\textbf{The friend of my friend}}^{\textit{explanatory variable}}$ **is my** $\overbrace{\textbf{friend}}^{\textit{dep. var.}}$.
- The enemy of my friend is my enemy.
- The friend of my enemy is my enemy.
- The enemy of my enemy is my friend.

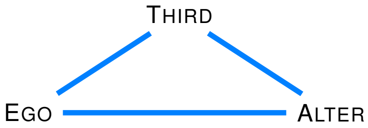


$\Rightarrow$ higher probability for cooperation, lower probability for conflict

Hypotheses derived from structural balance theory.

(Heider, 1946; Cartwright and Harary, 1956)

- **$\overbrace{\text{The friend of my friend}}^{\textit{explanatory variable}}$ is my $\overbrace{\text{friend}}^{\textit{dep. var.}}$.**
- The enemy of my friend is my enemy.
- The friend of my enemy is my enemy.
- The enemy of my enemy is my friend.

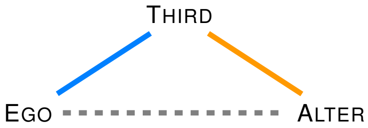


$\Rightarrow$ higher probability for cooperation, lower probability for conflict

Hypotheses derived from structural balance theory.

(Heider, 1946; Cartwright and Harary, 1956)

- The $\overbrace{\text{friend of my friend}}^{\textit{explanatory variable}}$ is my $\overbrace{\text{friend}}^{\textit{dep. var.}}$.
- **The enemy of my friend is my enemy.**
- The friend of my enemy is my enemy.
- The enemy of my enemy is my friend.

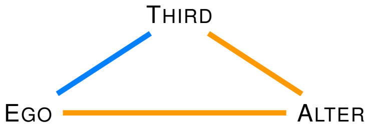


⇒ lower probability for cooperation, higher probability for conflict

Hypotheses derived from structural balance theory.

(Heider, 1946; Cartwright and Harary, 1956)

- $\overbrace{\text{The friend of my friend}}^{\textit{explanatory variable}}$ is my $\overbrace{\text{friend}}^{\textit{dep. var.}}$.
- **The enemy of my friend is my enemy.**
- The friend of my enemy is my enemy.
- The enemy of my enemy is my friend.

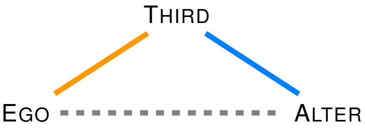


⇒ lower probability for cooperation, higher probability for conflict

Hypotheses derived from structural balance theory.

(Heider, 1946; Cartwright and Harary, 1956)

- $\overbrace{\text{The friend of my friend}}^{\textit{explanatory variable}}$ is my $\overbrace{\text{friend}}^{\textit{dep. var.}}$.
- The enemy of my friend is my enemy.
- **The friend of my enemy is my enemy.**
- The enemy of my enemy is my friend.

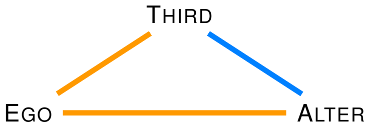


⇒ lower probability for cooperation, higher probability for conflict

Hypotheses derived from structural balance theory.

(Heider, 1946; Cartwright and Harary, 1956)

- $\overbrace{\text{The friend of my friend}}^{\textit{explanatory variable}}$ is my $\overbrace{\text{friend}}^{\textit{dep. var.}}$.
- The enemy of my friend is my enemy.
- **The friend of my enemy is my enemy.**
- The enemy of my enemy is my friend.

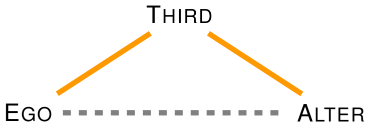


$\Rightarrow$ lower probability for cooperation, higher probability for conflict

Hypotheses derived from structural balance theory.

(Heider, 1946; Cartwright and Harary, 1956)

- $\overbrace{\text{The friend of my friend}}^{\textit{explanatory variable}}$ is my $\overbrace{\text{friend}}^{\textit{dep. var.}}$.
- The enemy of my friend is my enemy.
- The friend of my enemy is my enemy.
- **The enemy of my enemy is my friend.**

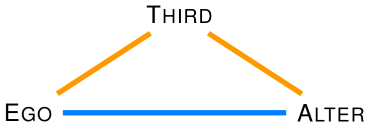


$\Rightarrow$ higher probability for cooperation, lower probability for conflict

Hypotheses derived from structural balance theory.

(Heider, 1946; Cartwright and Harary, 1956)

- $\overbrace{\text{The friend of my friend}}^{\textit{explanatory variable}}$ is my $\overbrace{\text{friend}}^{\textit{dep. var.}}$.
- The enemy of my friend is my enemy.
- The friend of my enemy is my enemy.
- **The enemy of my enemy is my friend.**



⇒ higher probability for cooperation, lower probability for conflict

Decomposition into event rate and conditional type.

Probability density for one weighted event $e = (u, v, t, w)$

$$f(e \big| G_{<t}; \theta) = f_\lambda(u, v, t \big| G_{<t}; \theta^{(\lambda)}) \cdot f_\mu(w \big| u, v; G_{<t}; \theta^{(\mu)}) \ .$$

Time-to-event on dyad (u, v) exp. dist. with rate

$$\lambda_{uv} \quad = \quad \exp\left(\sum_{h=1}^{k'} \beta_h \cdot s'_h(u, v; G_{<t})\right)$$

Conditional event weight normaly distributed around mean

$$\mu_{uv} \quad = \quad \sum_{h=1}^{k} \alpha_h \cdot s_h(u, v; G_{<t})$$

Estimation of weight parameters by linear regression;
(independent of model for the event rate!).

Statistics (I): inertia and reciprocity.

event weight $\quad \mu_{uv} = \sum_{h=1}^{k} \alpha_h \cdot s_h(u, v; G_{<t})$

explanatory variables: dependent var.: $u \longrightarrow v$;

dependence on dyad history

$\texttt{inertia}^{\pm}(u, v; G_{<t}) \quad u \rightrightarrows v \quad u \rightrightarrows v$

dependence on reverse dyad history

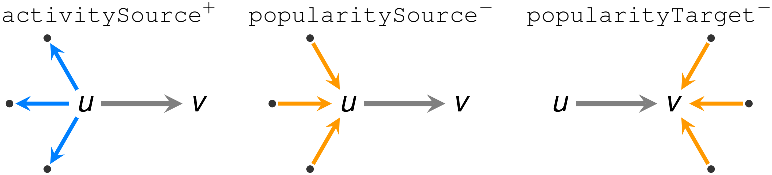
$\texttt{reciprocity}^{\pm}(u, v; G_{<t}) \quad u \leftrightarrows v \quad u \leftrightarrows v$

Statistics (II): activity and popularity.

event weight $\mu_{uv} = \sum_{h=1}^{k} \alpha_h \cdot s_h(u, v; G_{<t})$

explanatory variables: dependent var.: $u \longrightarrow v$;

accounting for differences in roles and positions (degree)



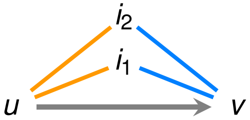
Varying source/target, positive/negative, and popularity/activity yields eight statistics.

Statistics (III): structural balance.

$$\text{event weight} \quad \mu_{uv} = \sum_{h=1}^{k} \alpha_h \cdot s_h(u, v; G_{<t})$$

explanatory variables: $\color{blue}{\blacksquare}$ $\color{orange}{\blacksquare}$ dependent var.: $u \longrightarrow v$;

$$\texttt{friendOfEnemy}(u, v; G_{<t}) = \sqrt{\sum_{\text{actors: } i} a^{-}(u, i; t) \cdot a^{+}(i, v; t)}$$



Similar for friends of friends, enemies of friends, and enemies of enemies.

Statistics (IV): covariates.

$$\text{event weight} \quad \mu_{uv} = \sum_{h=1}^{k} \alpha_h \cdot s_h(u, v; G_{<t})$$

Events on a dyad (u, v) depend on various characteristics:

- u and v are allies or not;
- geographical distance between u and v; joint border
- democracy level of u and v;
- capability scores (size, military and economic power)
- ...

These are available for state actors (countries)
$\Rightarrow$ restrict network / estimate model only for GULF conflict.

Conditional weight parameters.

statistics	LEVANT	BALKANS	GULF	TURKEY
`inertia`$^+$	**0.272 (0.113)**	**0.718 (0.081)**	**0.252 (0.012)**	**4.184 (0.693)**
`inertia`$^-$	**-0.088 (0.008)**	**-0.368 (0.030)**	**-0.092 (0.003)**	**-0.657 (0.230)**
`reciprocity`$^+$	0.048 (0.120)	-0.129 (0.100)	**0.132 (0.014)**	0.665 (0.758)
`reciprocity`$^-$	**-0.137 (0.013)**	**-0.225 (0.036)**	**-0.096 (0.003)**	**-0.729 (0.307)**
`friendOfFriend`	**1.557 (0.073)**	**0.886 (0.122)**	**0.223 (0.023)**	**2.216 (0.620)**
`friendOfEnemy`	-0.061 (0.044)	**-0.818 (0.084)**	**-0.134 (0.013)**	0.518 (0.574)
`enemyOfFriend`	-0.069 (0.040)	**-0.679 (0.081)**	**-0.157 (0.013)**	0.091 (0.566)
`enemyOfEnemy`	**-0.305 (0.015)**	**0.198 (0.051)**	**0.060 (0.007)**	**-3.110 (0.439)**
`activitySource`$^+$	**0.135 (0.019)**	**0.222 (0.022)**	**0.061 (0.003)**	**1.140 (0.169)**
`activitySource`$^-$	**-0.107 (0.003)**	**-0.058 (0.010)**	**-0.013 (0.001)**	**-1.272 (0.097)**
`activityTarget`$^+$	0.008 (0.017)	**0.231 (0.021)**	**0.042 (0.003)**	**1.264 (0.163)**
`activityTarget`$^-$	**-0.045 (0.003)**	**-0.017 (0.008)**	0.001 (0.001)	**-0.488 (0.095)**
`popularitySource`$^+$	**0.078 (0.017)**	0.033 (0.018)	**-0.025 (0.004)**	**0.396 (0.166)**
`popularitySource`$^-$	**-0.017 (0.004)**	-0.009 (0.012)	**0.005 (0.001)**	-0.123 (0.090)
`popularityTarget`$^+$	**0.127 (0.014)**	**0.058 (0.014)**	**-0.028 (0.004)**	0.080 (0.156)
`popularityTarget`$^-$	**-0.045 (0.003)**	**-0.061 (0.010)**	**0.004 (0.001)**	**-0.685 (0.078)**
constant	**-0.087 (0.002)**	**-0.038 (0.002)**	**-0.078 (0.001)**	**-0.013 (0.004)**

blue: more friendly **orange: more hostile**

Conditional weight parameters.

statistics	LEVANT	BALKANS	GULF	TURKEY
inertia$^+$	**0.272 (0.113)**	**0.718 (0.081)**	**0.252 (0.012)**	**4.184 (0.693)**
inertia$^-$	**-0.088 (0.008)**	**-0.368 (0.030)**	**-0.092 (0.003)**	**-0.657 (0.230)**
reciprocity$^+$	0.048 (0.120)	-0.129 (0.100)	**0.132 (0.014)**	0.665 (0.758)
reciprocity$^-$	**-0.137 (0.013)**	**-0.225 (0.036)**	**-0.096 (0.003)**	**-0.729 (0.307)**
friendOfFriend	1.557 (0.073)	0.886 (0.122)	0.223 (0.023)	2.216 (0.620)
friendOfEnemy	-0.061 (0.044)	-0.818 (0.084)	-0.134 (0.013)	0.518 (0.574)
enemyOfFriend	-0.069 (0.040)	-0.679 (0.081)	-0.157 (0.013)	0.091 (0.566)
enemyOfEnemy	-0.305 (0.015)	0.198 (0.051)	0.060 (0.007)	-3.110 (0.439)
activitySource$^+$	0.135 (0.019)	0.222 (0.022)	0.061 (0.003)	1.140 (0.169)
activitySource$^-$	-0.107 (0.003)	-0.058 (0.010)	-0.013 (0.001)	-1.272 (0.097)
activityTarget$^+$	0.008 (0.017)	0.231 (0.021)	0.042 (0.003)	1.264 (0.163)
activityTarget$^-$	-0.045 (0.003)	-0.017 (0.008)	0.001 (0.001)	-0.488 (0.095)
popularitySource$^+$	0.078 (0.017)	0.033 (0.018)	-0.025 (0.004)	0.396 (0.166)
popularitySource$^-$	-0.017 (0.004)	-0.009 (0.012)	0.005 (0.001)	-0.123 (0.090)
popularityTarget$^+$	0.127 (0.014)	0.058 (0.014)	-0.028 (0.004)	0.080 (0.156)
popularityTarget$^-$	-0.045 (0.003)	-0.061 (0.010)	0.004 (0.001)	-0.685 (0.078)
constant	-0.087 (0.002)	-0.038 (0.002)	-0.078 (0.001)	-0.013 (0.004)

inertia$^+$ $u \rightrightarrows v$, inertia$^-$ $u \rightrightarrows v$

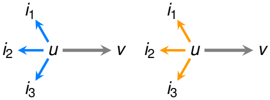
reciprocity$^+$ $u \rightleftarrows v$, reciprocity$^-$ $u \rightleftarrows v$

Conditional weight parameters.

statistics	LEVANT	BALKANS	GULF	TURKEY
inertia$^+$	0.272 (0.113)	0.718 (0.081)	0.252 (0.012)	4.184 (0.693)
inertia$^-$	-0.088 (0.008)	-0.368 (0.030)	-0.092 (0.003)	-0.657 (0.230)
reciprocity$^+$	0.048 (0.120)	-0.129 (0.100)	0.132 (0.014)	0.665 (0.758)
reciprocity$^-$	-0.137 (0.013)	-0.225 (0.036)	-0.096 (0.003)	-0.729 (0.307)
friendOfFriend	**1.557 (0.073)**	**0.886 (0.122)**	**0.223 (0.023)**	**2.216 (0.620)**
friendOfEnemy	-0.061 (0.044)	**-0.818 (0.084)**	**-0.134 (0.013)**	0.518 (0.574)
enemyOfFriend	-0.069 (0.040)	**-0.679 (0.081)**	**-0.157 (0.013)**	0.091 (0.566)
enemyOfEnemy	**-0.305 (0.015)**	**0.198 (0.051)**	**0.060 (0.007)**	**-3.110 (0.439)**
activitySource$^+$	0.135 (0.019)	0.222 (0.022)	0.061 (0.003)	1.140 (0.169)
activitySource$^-$	-0.107 (0.003)	-0.058 (0.010)	-0.013 (0.001)	-1.272 (0.097)
activityTarget$^+$	0.008 (0.017)	0.231 (0.021)	0.042 (0.003)	1.264 (0.163)
activityTarget$^-$	-0.045 (0.003)	-0.017 (0.008)	0.001 (0.001)	-0.488 (0.095)
popularitySource$^+$	0.078 (0.017)	0.033 (0.018)	-0.025 (0.004)	0.396 (0.166)
popularitySource$^-$	-0.017 (0.004)	-0.009 (0.012)	0.005 (0.001)	-0.123 (0.090)
popularityTarget$^+$	0.127 (0.014)	0.058 (0.014)	-0.028 (0.004)	0.080 (0.156)
popularityTarget$^-$	-0.045 (0.003)	-0.061 (0.010)	0.004 (0.001)	-0.685 (0.078)
constant	-0.087 (0.002)	-0.038 (0.002)	-0.078 (0.001)	-0.013 (0.004)

Conditional weight parameters.

statistics	LEVANT	BALKANS	GULF	TURKEY
`inertia`$^+$	0.272 (0.113)	0.718 (0.081)	0.252 (0.012)	4.184 (0.693)
`inertia`$^-$	-0.088 (0.008)	-0.368 (0.030)	-0.092 (0.003)	-0.657 (0.230)
`reciprocity`$^+$	0.048 (0.120)	-0.129 (0.100)	0.132 (0.014)	0.665 (0.758)
`reciprocity`$^-$	-0.137 (0.013)	-0.225 (0.036)	-0.096 (0.003)	-0.729 (0.307)
`friendOfFriend`	1.557 (0.073)	0.886 (0.122)	0.223 (0.023)	2.216 (0.620)
`friendOfEnemy`	-0.061 (0.044)	-0.818 (0.084)	-0.134 (0.013)	0.518 (0.574)
`enemyOfFriend`	-0.069 (0.040)	-0.679 (0.081)	-0.157 (0.013)	0.091 (0.566)
`enemyOfEnemy`	-0.305 (0.015)	0.198 (0.051)	0.060 (0.007)	-3.110 (0.439)
`activitySource`$^+$	**0.135 (0.019)**	**0.222 (0.022)**	**0.061 (0.003)**	**1.140 (0.169)**
`activitySource`$^-$	**-0.107 (0.003)**	**-0.058 (0.010)**	**-0.013 (0.001)**	**-1.272 (0.097)**
`activityTarget`$^+$	0.008 (0.017)	**0.231 (0.021)**	**0.042 (0.003)**	**1.264 (0.163)**
`activityTarget`$^-$	**-0.045 (0.003)**	**-0.017 (0.008)**	0.001 (0.001)	**-0.488 (0.095)**
`popularitySource`$^+$	**0.078 (0.017)**	0.033 (0.018)	**-0.025 (0.004)**	**0.396 (0.166)**
`popularitySource`$^-$	**-0.017 (0.004)**	-0.009 (0.012)	**0.005 (0.001)**	-0.123 (0.090)
`popularityTarget`$^+$	**0.127 (0.014)**	**0.058 (0.014)**	**-0.028 (0.004)**	0.080 (0.156)
`popularityTarget`$^-$	**-0.045 (0.003)**	**-0.061 (0.010)**	**0.004 (0.001)**	**-0.685 (0.078)**
constant	-0.087 (0.002)	-0.038 (0.002)	-0.078 (0.001)	-0.013 (0.004)



Conditional weight parameters.

statistics	LEVANT	BALKANS	GULF	TURKEY
inertia$^+$	**0.272 (0.113)**	**0.718 (0.081)**	**0.252 (0.012)**	**4.184 (0.693)**
inertia$^-$	**-0.088 (0.008)**	**-0.368 (0.030)**	**-0.092 (0.003)**	**-0.657 (0.230)**
reciprocity$^+$	0.048 (0.120)	-0.129 (0.100)	**0.132 (0.014)**	0.665 (0.758)
reciprocity$^-$	**-0.137 (0.013)**	**-0.225 (0.036)**	**-0.096 (0.003)**	**-0.729 (0.307)**
friendOfFriend	**1.557 (0.073)**	**0.886 (0.122)**	**0.223 (0.023)**	**2.216 (0.620)**
friendOfEnemy	-0.061 (0.044)	**-0.818 (0.084)**	**-0.134 (0.013)**	0.518 (0.574)
enemyOfFriend	-0.069 (0.040)	**-0.679 (0.081)**	**-0.157 (0.013)**	0.091 (0.566)
enemyOfEnemy	**-0.305 (0.015)**	**0.198 (0.051)**	**0.060 (0.007)**	**-3.110 (0.439)**
activitySource$^+$	**0.135 (0.019)**	**0.222 (0.022)**	**0.061 (0.003)**	**1.140 (0.169)**
activitySource$^-$	**-0.107 (0.003)**	**-0.058 (0.010)**	**-0.013 (0.001)**	**-1.272 (0.097)**
activityTarget$^+$	0.008 (0.017)	**0.231 (0.021)**	**0.042 (0.003)**	**1.264 (0.163)**
activityTarget$^-$	**-0.045 (0.003)**	**-0.017 (0.008)**	0.001 (0.001)	**-0.488 (0.095)**
popularitySource$^+$	**0.078 (0.017)**	0.033 (0.018)	**-0.025 (0.004)**	**0.396 (0.166)**
popularitySource$^-$	**-0.017 (0.004)**	-0.009 (0.012)	**0.005 (0.001)**	-0.123 (0.090)
popularityTarget$^+$	**0.127 (0.014)**	**0.058 (0.014)**	**-0.028 (0.004)**	0.080 (0.156)
popularityTarget$^-$	**-0.045 (0.003)**	**-0.061 (0.010)**	**0.004 (0.001)**	**-0.685 (0.078)**
constant	**-0.087 (0.002)**	**-0.038 (0.002)**	**-0.078 (0.001)**	**-0.013 (0.004)**

Rate parameters.

statistics	LEVANT	BALKANS	GULF	TURKEY
`inertia`	0.053 (0.016)	-1.415 (0.047)	-0.160 (0.002)	6.564 (0.560)
`reciprocity`	-0.873 (0.023)	0.069 (0.048)	-0.141 (0.002)	-1.399 (0.954)
`triangle`	0.508 (0.007)	3.783 (0.026)	0.604 (0.002)	1.392 (0.395)
`activitySource`	0.658 (0.004)	1.000 (0.010)	0.211 (0.001)	10.593 (0.162)
`activityTarget`	0.600 (0.004)	0.635 (0.007)	0.178 (0.001)	9.274 (0.131)
`popularitySource`	0.699 (0.004)	0.259 (0.013)	0.093 (0.001)	3.820 (0.158)
`popularityTarget`	0.843 (0.003)	1.177 (0.008)	0.128 (0.001)	5.595 (0.119)
constant	-9.939 (0.003)	-9.157 (0.004)	-7.116 (0.002)	-11.163 (0.008)

Conditional weight parameters for GULF conflict.

statistic	event network model	covariate model	joint model
`inertia`$^+$	**0.214 (0.012)**	·	**0.192 (0.012)**
`inertia`$^-$	**-0.085 (0.003)**	·	**-0.071 (0.003)**
`reciprocity`$^+$	**0.124 (0.014)**	·	**0.075 (0.014)**
`reciprocity`$^-$	**-0.082 (0.004)**	·	**-0.052 (0.004)**
`friendOfFriend`	**0.246 (0.027)**	·	**0.138 (0.027)**
`enemyOfFriend`	**-0.206 (0.014)**	·	**-0.119 (0.015)**
`friendOfEnemy`	**-0.224 (0.015)**	·	**-0.137 (0.015)**
`enemyOfEnemy`	**0.113 (0.008)**	·	**0.057 (0.008)**
`activitySource`$^+$	**0.051 (0.003)**	·	**0.009 (0.004)**
`activitySource`$^-$	**-0.008 (0.001)**	·	0.001 (0.002)
`activityTarget`$^+$	**0.040 (0.004)**	·	-0.006 (0.004)
`activityTarget`$^-$	0.002 (0.002)	·	**0.013 (0.002)**
`popularitySource`$^+$	-0.008 (0.005)	·	**0.023 (0.005)**
`popularitySource`$^-$	**0.003 (0.001)**	·	**-0.007 (0.002)**
`popularityTarget`$^+$	**-0.020 (0.005)**	·	0.005 (0.005)
`popularityTarget`$^-$	**0.004 (0.001)**	·	**-0.005 (0.002)**
`lnCapRatio`	·	**0.002 (0.001)**	**-0.007 (0.001)**
`allies`	·	**0.118 (0.003)**	**0.106 (0.003)**
`polityWeakLink`	·	$3.2E^{-4}$ $(1.8E^{-4})$	**-0.001** $(1.9E^{-4})$
`minorPowers`	·	**0.097 (0.003)**	**0.042 (0.004)**
`lnTrade`	·	**0.028 (0.001)**	**0.017 (0.001)**
`contiguity`	·	**-0.093 (0.003)**	**-0.060 (0.003)**
`lnDistance`	·	**0.011 (0.001)**	**0.013 (0.001)**
`lnJointIGO`	·	**-0.097 (0.003)**	**-0.076 (0.003)**
constant	**-0.082 (0.001)**	**-0.017 (0.008)**	-0.002 (0.008)

Rate parameters for GULF conflict.

statistic	event network model	covariate model	joint model
inertia	-0.114 (0.002)	·	$-1.9\mathrm{E}^{-4}$ (0.002)
reciprocity	-0.090 (0.003)	·	0.042 (0.003)
triangle	0.506 (0.002)	·	0.348 (0.003)
activitySource	0.202 (0.001)	·	0.161 (0.001)
activityTarget	0.168 (0.001)	·	0.118 (0.001)
popularitySource	0.094 (0.001)	·	0.073 (0.001)
popularityTarget	0.131 (0.001)	·	0.119 (0.001)
lnCapRatio	·	-0.289 (0.002)	-0.225 (0.002)
allies	·	0.064 (0.006)	-0.223 (0.006)
polityWeakLink	·	-0.137 (0.001)	-0.122 (0.001)
minorPowers	·	-2.726 (0.007)	-1.970 (0.007)
lnTrade	·	0.062 (0.001)	0.142 (0.001)
contiguity	·	1.362 (0.006)	1.310 (0.007)
lnDistance	·	-0.287 (0.002)	-0.343 (0.002)
lnJointIGO	·	1.344 (0.005)	1.313 (0.005)
constant	-6.774 (0.002)	-6.964 (0.017)	-7.530 (0.016)