Dynamic Ego Networks

Growth, Stability, Change and Inequality

David Hachen   Omar Lizardo
Michael Penta   Brandon Sepulvado
Joseph Workman

University of Notre Dame - iCeNSA

Indiana University - Bloomington, November 8, 2013

Overview

Vertex Degree (Centrality) in Ego Networks

- "Range" of contacts that an individual possesses at any one time is fundamental in ego-network research (Burt 1983)
- Range has *criterion validity*:
  - It predicts a variety of (usually beneficial) outcomes
  - It correlates with exogenous markers of social positon (e.g. SES)
- Range has *construct validity*: It is a very important indicator of an individual's position in the social structure
- Thought to be composed of the following dimensions:
  1. Ego Network *Size*
  2. Ego Network *Density* (clustering coefficient)
  3. Ego Network *Diversity* (composition)
- While empirically and conceptually distinct, size is usually correlated with the other two making degree centrality the main component of the network range construct

Inequality in Degree

- Well known that there are differences among people in the size of their social networks
- Degree distributions tend to not be normally distributed
  - Some persons have many other people with whom they interact, others interact with very few persons
  - Long tail, power law like distribution
- What explains this inequality? What processes generate degree heterogeneity? (Gould 2002)
  - Preferential Attachment and Popularity Tournaments
  - Exogenous factors
- But how do we measure degree?

Cross-sectional vs. Dynamic Measures of Degree

- Standard (cross-sectional) ego-network analyses provide a "snapshot" of an individual's range of contacts
  - Degree is simply the number of (self)reported contacts
  - If behavioral data are used, degree is the cumulative count of the number of people with whom a person interacts over some time period
- Inclusion of the temporal (dynamic) component complicates the picture
  - Tie formation and tie decay must be taken into account (Burt 2000)
  - Recent paper by Miritello et al (2013) shows how to use temporally fine-grained dynamic network data to measure various aspects of ego network size
  - Construction of degree centrality indicators become dependent on the temporal window within which individuals are observed

Three Measures of Degree

- Each individual is observed on a sequential number of time units up until the last observation point $\{t_0, t_1, t_2, t_3 ... t_\omega\}$
  - Time units may be hours, days, weeks, months, etc.
  - Time window $\Omega = \{t_0, t_1, t_2, ... t_\omega\}$
- In this dynamic context we can distinguish at least three empirically and conceptually distinct notions of vertex degree
  - **Cumulative Degree** $k_i(T_k)$: Count of the number of *unique contacts* that a given individual has interacted with up until a given time point T_k
  - **Effective Degree** $\kappa_i(t_k)$: The number of *active contacts* that an individual has the *potential* to interact with at a given time point t_k
  - **Active Degree** $\kappa_i^*(t_k)$: The number of *active contacts* than an individual *does in fact interact with* at a given time point t_k

Measuring Degree: The Censoring Problem

- **Left-censoring problem**: When an interaction between ego (i) and an alter (j) occurs for the first time at time $t_k \in \Omega$ we do not know if it is a new tie or the reactivation of an old tie
- **Right-censoring problem**: Similarly, when an interaction occurs for the last time at time $t_k \in \Omega$ we do not know if the tie has been deactivated or whether it was just not active for the remainder of Ω
- **Solution**: Use time windows that are before the first and after the last time point in Ω to establish whether the first interaction in Ω indicates a new tie and the last one a deactivation

Measuring Degree: The Censoring Problem

- With left censoring, begin by identifying all interactions with alters within $\Omega = \{t_0, t_1, t_2, ...t_\omega\}$ and count the number of ties that began prior to Ω (i.e. the tie was active in the time window before Ω):
  - This is value of $\kappa_i(t_0)$
  - In NetSense data where there is no left censoring, $\kappa_i(t_0) = 0$
- For all other interactions, note the time point $t_a \in \Omega$ when the interaction first occurs:
  - t_a is the *formation* date of the tie
- For all ties where there is at least one interaction within Ω determine if another interaction occurs after Ω:
  - If so, this is a *persistent* tie (does not decay within Ω)
  - If not, determine latest time point $t_w \in \Omega$ when tie is active
  - t_w is the *deactivation date* of the tie

Computing Cumulative, Effective and Active Degree

- **Effective Degree** κ_i: Beginning at t_0, $\kappa_i(t_1)$ is computed by adding the number of newly formed ties, $n_{\alpha,i}(t_1)$, and subtracting the number of deactivated ties $n_{\omega,i}(t_1)$
  - $\kappa_i(t_k) = \kappa_i(t_k - 1) + [n_{\alpha,i}(t_k) - n_{\omega,i}(t_k)]$
- **Cumulative Degree** $k_i(T_k)$: Total number of alters a person has interacted with up until time k
  - $k_i(t_k) = \sum_j a_{ij}(t_k)$
  - Where $a_{ij}(t_k) = 1$ when the ij tie appears for the first time in Ω
  - Cumulative Degree: $k_i(T_w) = \sum_{\mathbf{t_0}}^{\mathbf{t_w}} k_i(t_k)$
- **Active Degree** κ_i^*:
  - $\kappa_i^*(t_k) = \sum_j a_{ij}^*(t_k)$
  - Where $a_{ij}^*(t_k) = 1$ if i interacts with j at t_k

Effective Degree Growth

- $G_i(t_k) = n_{\alpha,i}(t_k) - n_{\omega,i}(t_k)$
- Changes in $\kappa_i(t_k)$ at the ego level are solely a function of the difference between the number of ties formed $n_{\alpha,i}(t_k)$ and number of ties deactivated $n_{\omega,i}(t_k)$
- This implies that: $\kappa_i(t_k) = \kappa_i(t_k - 1) + G_i(t_k)$
- For any given time period Δ that begins at t_a and ends at t_w:
  1. Whenever $\sum_{t_a}^{t_w} n_{\alpha,i}(t) \neq \sum_{t_a}^{t_w} n_{\omega,i}(t)$, we should observe fluctuations in κ_i within persons over time
  2. Whenever $\sum_{t_a}^{t_w} n_{\alpha,i}(t) \simeq \sum_{t_a}^{t_w} n_{\omega,i}(t)$ we should observe little variation in κ_i within persons over time
  3. If (2) holds, then $\kappa_i(t_k) \simeq \bar{\kappa}_i$, where $\bar{\kappa}_i = \frac{1}{T} \sum_{t_a}^{t_w} \kappa_i(t)$ and T is the number of time points in Δ

Churn (Turnover)

- Effective Degree: Size of pool of people with whom an ego can interact
- Churn captures extent of change in who is in this pool of people
- When $\kappa_i(t_k) fairly stable$ (because $\kappa_i(t_k) \simeq \bar{\kappa}_i$ and $n_{\alpha,i}(t_k) \simeq n_{\omega,i}(t_k)$) then churn is just count of number of new (or deactivated) ties:
  - $C_i(t_k) = n_{\alpha,i}(t_k) = n_{\omega,i}(t_k)$
- When effective degree is changing within persons over time ($\kappa_i(t_k) \neq \bar{\kappa}_i$), then churn is given either by:
  - Growth case ($n_{\alpha,i}(t_k) > n_{\omega,i}(t_k)$): $C_i(t_k) = \sum_{t_0}^{t_w} n_{\omega,i}(t)$
  - Decline case ($n_{\alpha,i}(t_k) < n_{\omega,i}(t_k)$): $C_i(t_k) = \sum_{t_0}^{t_w} n_{\alpha,i}(t)$
- When normalized by effective degree, gives us the average proportion of the effective ego-network that has changed in the time period:
  - $C_i^*(t_k) = \frac{C_i(t_k)}{\bar{\kappa}_i}$

Summary

- **Cumulative Degree**: $k_i(T_k) = \kappa_i(0) + \sum_{t_0}^{t_w} n_{\alpha,i}(t_k)$
  - When there is churn, cumulative degree overstates effective degree
- **Effective degree**: $\kappa_i(t_k)$, and when relatively stable $\overline{\kappa}_i$
- **Active Degree**: $\kappa_i^*(t_k) = \sum_j a_{ij}^*(t_k)$
  - When normalized by effective degree, active degree gives the *proportion* of the total network that the individual activates at any given point in time:
  - $\kappa_i^{**}(t_k) = \frac{\kappa_i^*(t_k)}{\kappa_i(t_k)}$
  - When there is heterogeneity in levels of (communication) activity across ties and/or persons, active degree understates effective degree

NetSense Study

- Two-year data collection effort (now extended to 4),funded by NSF and subsidized by Sprint
- $\sim$ 200 first-year students who began attending Notre Dame in the Fall of 2011 were given a smartphone and free voice/text/data plans
- Capture all communicative interactions: sender number, receiver number, time, date, communication form (voice, text)
- Random allocation of students to same-sex dorms
- Survey data: Prior to arrival students completed two surveys (and then subsequently every 4 months)
  - Demographic/Attitudes/Tastes/Behaviors survey
  - Network Survey: asked about their top 20 contacts

Data

- 161 students
  - $\sim$ 35 students dropped out of study before week 70
- 70 weeks: all of first year and first semester second year
- $\sim$ 23 million communication events (voice or text)
- $\sim 46,000$ ties (weighted edges): communication between someone in the study and another person (possibly also in the study)
- $\sim 23,000$ mutual ties: at least 1 communication from i to j and from j to i
- Vast majority of ties and communication are with other people who are not in the NetSense study

Types of Ties

- Communication events pick up activity with both pre-existing (prior to arrival at ND) ties and new ties
- **Pre-existing Ties**:
  - **Family ties** (3%): Identified from repeated name-generator surveys (top-20 contacts)
  - **Non-family ties** (6%): Ties to high school friends and others that remain active after arrival: Identified from repeated name-generator surveys (top-20 contacts)
- **New (Post-Arrival) ND Ties** (52%, $N \approx 11,000$): Ties formed after arrival on campus
  - No left-censoring: first time an interaction occurs is when the tie begins
- Residual Category: unknown type of ties (39%)

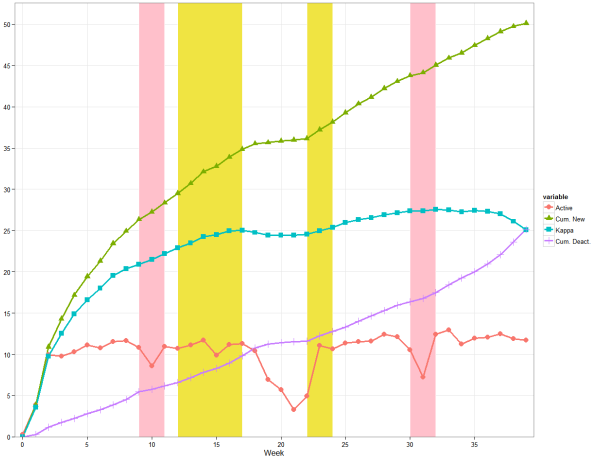
Time Window (Ω)

- Focus of the study is on the evolution of each student's ND ego-network during their first academic year (AY1), a 39 week period
- Remaining 31 weeks include Summer 2012 and Fall Semester 2012 are used to determine if tie deactivates in AY1
- If a tie is active in AY1 but not active in the Summer 2012 or Fall 2012, then it is considered to have been deactivated in AY1
- Of the 11k post-arrival ties, 78% (N $\simeq 9k$) activated during AY1
- Remainder activated in either Summer 2012 or Fall 2012 and are not considered in this analysis
- Of the 9k ties activated in AY1, 58% (N $\simeq 5k$) were *not* active in F12, so these ties are considered to have been deactivated in AY1

Academic Year Time Periods

- Period 1: Formative Period (weeks 1-12, includes week 10 which is Fall Break)
- Period 2: Fall Semester (weeks 13-18)
- Period 3: Winter Break (weeks 19-22)
- Period 4: Beginning Spring Semester (weeks 23-25)
- Period 5: Spring Semester (weeks 26-36)

Average Effective Degree, Active Degree, Cumulative Degree, and Cumulative Deactivated Ties (N = 161)



Weekly average values of $\kappa_i(t_k)$, $k_i(T_k)$,$k_i^*(t_k)$ and $n_{\omega,i}(T_k)$, n = 161

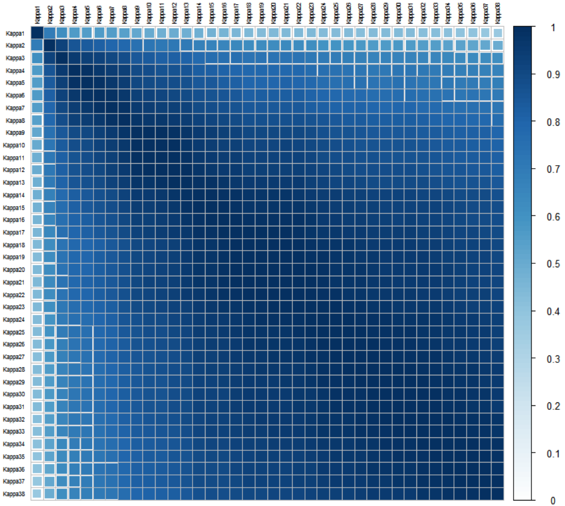
Cumulative and Effective Degree

- By week 39, the average student has interacted with about 50 other ND students (cumulative degree)
- Good deal of variation around this average ($\sigma_{k_i} = 25.2,\ CV = 0.50$)
- κ_i (effective degree) is fairly stable after an initial growth period
  - All students begin with $\kappa_i(t_0) = 0$
  - By t_{12} κ_i begins to stabilize for the typical student at around 24
  - For the remainder of the Fall semester ($13 \geq t \leq 18$) $\kappa_i(t)$ is around 24.5
  - κ_i increases at the beginning of the Spring semester and then stays fairly constant: $\overline{\kappa_i}(t_{26}, t_{36}) = 27.2$
  - There is a decline in κ_i at end of the AY1 (weeks 37-39) to around 25 ($\sigma_{\kappa_i} = 14.7,\ CV = 0.59$) as some ties created in first year are not active in Fall 2012

Cumulative and Effective Degree

- Relative stability of $\kappa_i(t_k)$ from end of Fall semester through Spring semester is a result of the fact that: $n_{\alpha,i}(13,39) \simeq n_{\omega,i}(13,39)$
  - $n_{\alpha,i}(13,39) = 20.8, n_{\omega,i}(13,39) = 18.6$
- $k_i(t_k)$ overstates effective degree because it conflates degree and turnover
- Typical student formed about 50 ties during AY1, but also deactivated about 25 ties, resulting in a $\kappa_i(39) \simeq 25$ and a turnover rate of 100% for AY1

Intertemporal Correlation of κ_i



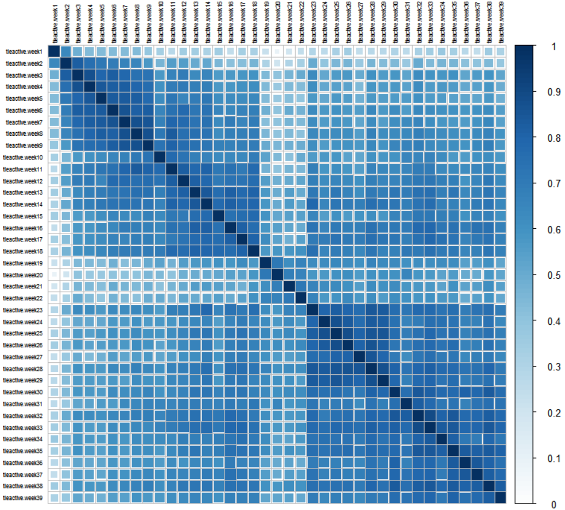
Cumulative and Effective Degree

- Effective degree ($\kappa_i(t_k)$) and cumulative degree ($k_i(t_k)$) are correlated
  - E.g., for weeks 26-36, r = .77
- Students with higher effective degrees are more likely to form new ties
- Association is not perfect: There are people with lower degree and lots of new ties, and people with higher degree and fewer new ties

Active Degree

- $\kappa_i^*(t_k)$ shows strong period effects
- During breaks, $\kappa_i^*(t)$ drops as students communicate less with their ND friends
- $\kappa_i^*(t_k)$ picks up during semesters and remains fairly constant for the average student at around 11 students per week
- $\kappa_i^*(t_k)$ is less stable than $\kappa_i(t_k)$
- $\kappa_i^*(t_k)$ is on average much lower that $\kappa_i(t_k)$
- It is also less heterogeneous ($\sigma_{\kappa_i^*} = 3.9,\ CV = 0.36$)
- During the semester the typical ND student interacts each week with between 40-50% of the ND contacts that they *could* interact with (e.g., $\kappa^{**}(t_{26}, t_{36}) = \frac{11.5}{26.3} = 0.43$)

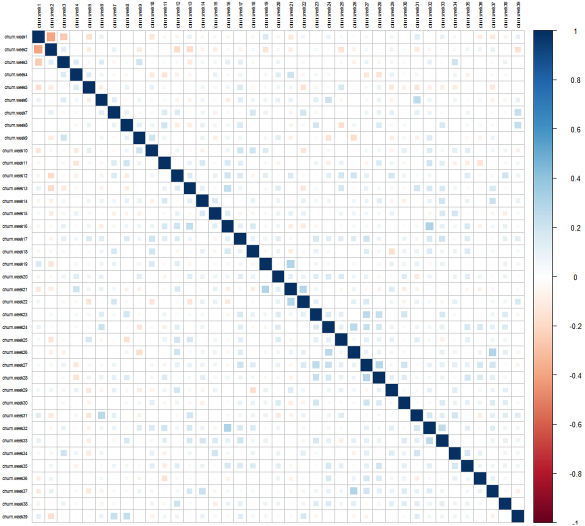
Intertemporal Correlation of κ_i^*



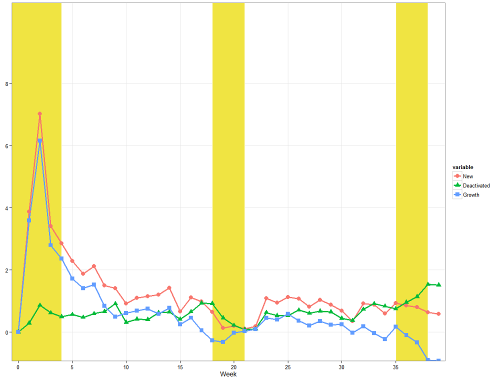
Churn

- Need to normalize counts of new ties by degree to obtain indicator of churn, $C_i^*(t_{26}, t_{36}) = \frac{n_{\alpha,i}(t_{26},t_{36})}{\overline{\kappa}_i(t_{26},t_{36})}$
- Distribution of churn in Spring semester is normal with a few positive outliers
    - $\mu_{C_i^*} = \frac{9.0}{26.89} = .34$, $\sigma_{C_i^*} = .17$
- Captures the proportion of a person's network that turns over in this 11 week period
- No correlation with effective degree ($r = -0.07$)
- Small positive correlation turnover in weeks 13-18 and weeks 26-36 ($r = 0.26$)
- Churn is less stable and more episodic than degree

Intertemporal Correlation of C_i^*



Growth



Weekly average values of $n_{\alpha,i}(t_k)$, $n_{\omega,i}(t_k)$ and $G_i(t_k) = n_{\alpha,i}(t_k) - n_{\omega,i}(t_k)$, n = 161

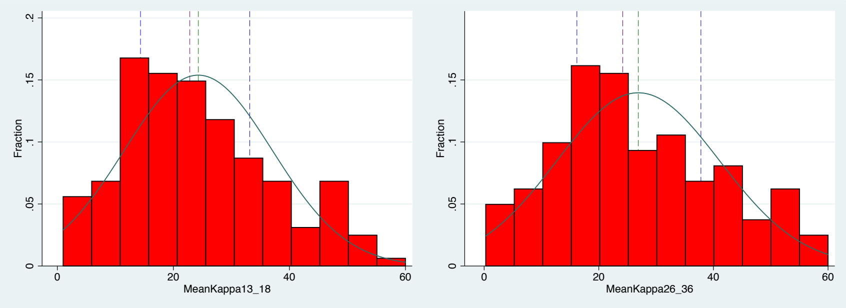
Growth

- During the initial formation period growth is relatively high $\mu_{G_i(t_0,t_{10})} = 0.75$
- After that growth declines towards zero $\mu_{G_i(t_{11},t_{39})} = 0.16$
- However, there is a lot of heterogeneity across individuals in relation to this average $\sigma_{G_i(t_{11},t_{39})} = 0.39$
- In the spring semester average growth is even lower $\mu_{G_i(t_{26},1_{36})} = 0.13$
- Although highly heterogeneous $\sigma_{G_i(t_{26},t_{36})} = 1.17$
- After the formation period, about a fifth of students experience a decline in effective degree, about a tenth remain stable, and the rest have positive weekly growth

Effective Degree Distributions

- Two relatively stable time periods: Fall (weeks 13-18) and Spring (weeks 26-36)
- Fall: $\mu_{\overline{\kappa_i}(t_{13},t_{18})} = 24.2$, $\sigma_{\overline{\kappa_i}(t_{13},t_{18})} = 12.5$
  - Between subject: $\sigma^{b}_{\overline{\kappa_i}(t_{13},t_{18})} = 12.35$
  - Within subject: $\sigma^{w}_{\overline{\kappa_i}(t_{13},t_{18})} = 1.70$
  - $\rho = .98$
- Spring: $\mu_{\overline{\kappa_i}(t_{26},t_{36})} = 26.9$, $\sigma_{\overline{\kappa_i}(t_{26},t_{36})} = 14.0$
  - Between subject: $\sigma^{b}_{\overline{\kappa_i}(t_{26},t_{36})} = 13.86$
  - Within subject: $\sigma^{w}_{\overline{\kappa_i}(t_{26},t_{36})} = 1.98$
  - $\rho = .98$
- $r_{\overline{\kappa_i}(t_{13},t_{18}),\overline{\kappa_i}(t_{26},t_{36})} = .94$

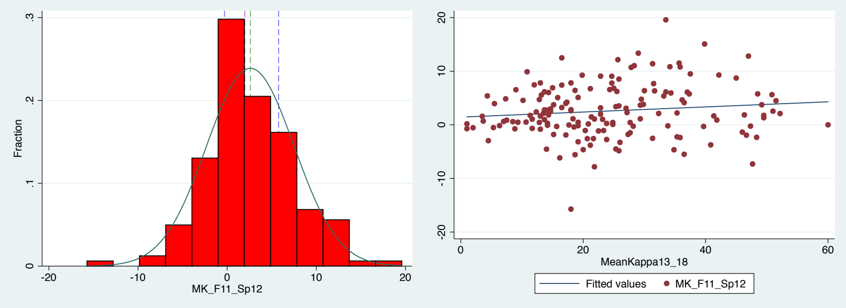
Effective Degree Distributions



Effective Degree Distribution for Two Periods: Fall 2011 (Weeks 13-18) and Spring 2012 (Weeks 26-36), n = 161

Change in Effective Degree

- Change score distribution $\mu = 2.6$, $\sigma = 4.9$
- Effective degree in Fall 2011 does *not* predict level of change in effective degree from Fall 2011 to Spring 2012
  - Implication: No indication of preferential attachment



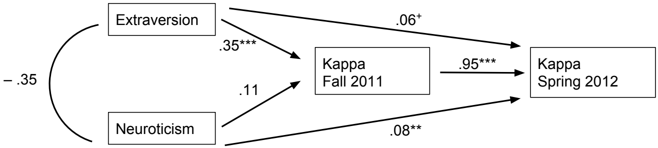
Change in Effective Degree Distribution and Scatter Plot of Change by Degree in Fall 2011, n = 161

Predicting Effective Degree

- Exogenous factors: demographic, background, activities, health, attitudes, aspirations, personality
- Measured prior to arriving at ND
- Path Model: Exogenous factors –> $\overline{\kappa_i}(t_{13}, t_{18})$ –> $\overline{\kappa_i}(t_{26}, t_{36})$
  - Alternatively can predict change scores

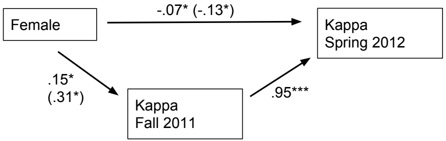
Extroversion

- Extroversion: 5 point scale (8 items), $\mu = 2.62$, $\sigma = .61$
- Extroversion strong predictor of effective degree
  - Net of other factors, a 1 SD increase in Extroversion yields .35 of a SD increase in $\overline{\kappa_i}(t_{13}, t_{18})$ ($\sim$ 6.7 friends for a 1 unit difference in Extroversion)
  - Net of Extroversion's effect on $\overline{\kappa_i}(t_{13}, t_{18})$ there is no direct effect of Extroversion on $\overline{\kappa_i}(t_{26}, t_{36})$
  - Because Extroversion is negatively correlated with Neuroticism, and Neuroticism has a small positive effect on $\overline{\kappa_i}(t_{26}, t_{36})$, the total effect of Extroversion on $\overline{\kappa_i}(t_{26}, t_{36})$ (.32) is less than the sum of its direct effect (.06) and its indirect effect through $\overline{\kappa_i}(t_{13}, t_{18})$ (.33)



Gender

- Women have .31 SD of $\overline{\kappa_i}(t_{13}, t_{18})$ ($\sim$ 4) more ND ties than men
- However men catch up so that by Spring women have .16 SD of $\overline{\kappa_i}(t_{26}, t_{36})$ ($\sim$ 2) more ND ties than men



Ego-Network Degree

- Cumulative, Effective, and Active Degree conceptually and empirically distinct in a dynamic ego-network context
- Cumulative Degree indexes a person's "history" of sociability but not their capacity to maintain ties nor their relative extent of tie activation
- Effective Degree best captures the notion of range:
  1. It is stable at relatively long time-scales
  2. Closest conceptually to Dunbar's fixed inter-individual "ceiling" of sociability
- Active degree captures the "strength of engagement" dimension
  1. Gap between active and effective degree a useful index of the relative prevalence of non-decayed, weaker ties
  2. Operational way of differentiating the inner-layer "support clique" from the active network "outer layer" (Hill and Dunbar 2003)

Degree Inequality

- In the NetSense data, differences in effective degree emerge within the first 13 weeks
- These differences persist through the Spring semester given the high correlation between effective degree in the Fall and Spring semesters
- The most important predictor of initial degree (and, therefore, subsequent degree) is Extroversion
- No evidence of preferential attachment
  - Change in effective degree is not a function of initial level of effective degree
- The effective degree distribution is not long-tailed
- The effective degree distribution is generated by differences among persons in their tendency to be social

Churn

- Differences across persons in their effective degree says nothing about who is in a person's ego-network
- While the size of a person's network may be a relatively stable quantity, who is in that network can change
- Through churn (turnover), growth (new ties) and decline (deactivation of ties) the composition of ego-networks change
- Churn likely to be more episodic and less person-specific

Thanks!

Operational criteria for classification as post-arrival tie

1. Named as ND student in ego-network surveys
2. Contact is shared by two unrelated NetSense study participants
3. Phone number of contact associated with ND email address
4. Phone number appear in the address books of unrelated NetSense study participants

Discriminant and Criterion Validity of Different Measures

	k_i	$\bar{\kappa}_i$	κ_i^*	G_i	Deact.	C_i	κ_i^{**}
Women	2.469	4.993*	3.499*	-0.0721	4.198*	-0.0217	0.0638*
	(1.21)	(2.32)	(3.71)	(-1.23)	(2.48)	(-0.41)	(3.52)
ND First	1.353	2.220	1.786*	0.0226	0.811	-0.0239	0.0475*
	(0.62)	(0.96)	(1.79)	(0.36)	(0.45)	(-0.36)	(2.32)
Outgoing	7.887*	10.38*	3.616*	0.0538	6.595*	-0.0644	-0.0446*
	(4.13)	(5.35)	(4.20)	(0.78)	(3.32)	(-1.11)	(-2.50)
Constant	19.97*	22.15*	8.710*	0.179*	15.66*	0.880*	0.391*
	(9.59)	(10.41)	(9.71)	(2.70)	(9.47)	(14.97)	(21.87)
N	156	156	156	156	156	156	156

* $p < 0.05$

Extroversion Scale

- I am someone who...
  - is talkative
  - is reserved
  - tends to be quiet
  - has an assertive personality
  - is sometimes shy, inhibited
  - is outgoing, sociable

Neuroticism Scale

- I am someone who...
  - is depressed, blue
  - is relaxed, handles stress well
  - can be tense
  - worries a lot
  - is emotionally stable, not easily upset
  - can be moody
  - remains calms in tense situations
  - gets nervous easily