Multilevel Models for Social Network Analysis

Paul-Philippe Pare
pare@uwo.ca
Department of Sociology
Centre for Population, Aging, and Health
University of Western Ontario

Pamela Wilcox & Matthew Logan
School of Criminal Justice
University of Cincinnati
Why is the connection not obvious?

• I will argue today that there is a logical and practical connection between multilevel analysis and social network analysis

• Yet, the combined methodological approach is uncommon

• Why?
  – Few scholars have both sets of quantitative skills
  – The two techniques are often taught as separate “advanced statistics” courses in graduate school: students must select one over the other
Game Plan

• Introduction to Multilevel Analysis
• Introduction to Social Network Analysis
• Making the connection
• Empirical example
  – Juvenile delinquency across peer groups: individual differences and peer influences
• Limitations
Multilevel Analysis

• Some studies implicitly or explicitly assume that only individual differences matter, and that social context is either irrelevant, trivial, or indirect
  – Supporters: Individual determinism
    • Social context → Individual differences → Outcomes
  – Detractors: Individualistic fallacy
    • Individual differences + social context → Outcomes
    • Individual differences x social context → Outcomes

• The good thing about multilevel analysis is that we can move beyond the philosophical debate, and test the relative effects of individual and contextual variables on our outcomes of interest
Multilevel Analysis

• In order to test for contextual effects, however, some preliminary steps are necessary:
  – Group differences in the DV (outcome) must be observable. If groups do not vary, there is no point in pursuing contextual effects
  – Theoretically relevant IV (contextual predictors) must be available at the group level. Ultimately, the researcher wants to find statistically significant predictors of the outcome, above and beyond a well-specified individual-level model.
    • Easier said than done
    • Individual predictors are often stronger than contextual predictors. Sometimes individual predictors “wipe away” contextual effects, even if group differences in the DV were observed.
Multilevel Analysis

• While group differences in the DV is a necessary condition to observe contextual effects, it is also a source of statistical bias if left uncorrected

• Why?
  – An important assumption of regression models and similar techniques is the independence of observations/error term
  – When group differences in the DV are observed, it means that people belonging to a group are more similar to each others than randomly selected people
    • Ex. Students in classrooms: some classrooms with all the best students (“enriched” programs), some diversified classrooms, and some classrooms with all the students in difficulties.
  – So, we have a dependence problem
    • If left uncorrected, the SE of our model will be underestimated, and we will observed “fake” significant effects that are the result of group-clustering
Multilevel Analysis

• This is the main reason why multilevel analysis is useful to analyze data where level-1 units (e.g. individuals) are clustered into level-2 groups (e.g. classrooms, neighborhoods, hospitals)
  – Multilevel analysis also provides other advantages (random coefficients, cross-level interactions, group-mean centering, etc.)

• How is the dependence problem corrected?
  – **With the random intercept**: Instead of having a regular intercept ($b_0$) for the whole model like linear regression, multilevel analysis has multiple intercepts for each group.
  – Groups that have higher levels on the DV get higher intercepts, and groups that have lower levels on the DV get lower intercepts
  – So, the dependence problem is solved. The fact that some groups are different from other groups is taken into account by the model.
Social Network Analysis

• SNA in the social sciences has often been used to study:
  – Social Capital (positions in the network providing structural opportunities)
  – Power relations (dominant/subaltern positions)

• More recently, SNA is becoming a popular technique for social scientists in a more general sense
  – Criminology: delinquent peers, gangs, organized crime
  – Internet studies: online communities, Facebook, online dating, etc.
  – Sociology of knowledge: Academic influence (scholars or major publications) in a field/sub-field; Academic “cliques” (groups of scholars who often cite each other work)
Social Network Analysis

• SNA can be based on many different concepts, as long as people are connected to other people for a reason:
  – Friendship
  – Business relationship
  – Hierarchy/power relationship
  – Sexual attraction
  – Time use, etc.

• SNA typically depends on a defined universe for the larger network, otherwise the analysis would become too complicated
  – Workplace
  – Neighborhood
  – School
  – Prison
  – Facebook
  – Online dating website
Social Network Analysis

• SNA is based on surveys administered to people who potentially know each others given the specific survey universe (workplace, school)

• The connections between people are based on questions such as
  – “Please identify up to 5 of your best friends in your school”
  – “Please identify up to 5 colleagues that are involved in your supervision in your workplace”

• Some researchers use more open methodologies (e.g. no specified universe or number of contacts), but must deal with additional problems, and must often rely on descriptive or qualitative analysis
  – Ex. New biography of Howard Hughes: Create a social network of Mr. Hughes’ main business partners, adversaries, and romantic partners.
Social Network Analysis

- Traditional SNA visualization (Social Capital / Power relations)
Social Network Analysis

• Examples of SNA variables
  – **Degree**: Number of contacts or ties to other people in the network
  – **Degree-Centrality**: Relative to the respondent’s own contacts, is the respondent in a central vs. peripheral position?
    • Being named by other as a friend, or leader, etc.
  – **Closeness**: Relative to the whole network, how close (or far) is the actor to others?
    • Local celebrity vs. Whole network celebrity
    • Group leader vs. Global leader
  – **Betweenness**: Number of contacts associated with the respondent but not with each other.
    • A ←→ B ←→ C
Social Network Analysis

• New applications of SNA:
  – Examine complex connections between large groups of people (see picture in a few slides), and empirically isolate coherent smaller groups (peers, gangs, workplace teams, etc.) from the larger network.

• Groups, and individuals in these groups, can be compared in follow-up analyses

• In other words, the SNA is a first step in a larger analysis
Fact 1: We have one technique that allows us to group people together based on different conceptualization of networks (e.g. friendship, business relationship) and to create empirical variables from the networks.

Fact 2: We have another technique that allows us to elegantly analyze data involving smaller units clustered into larger units, control clustering biases, and estimate contextual effects above and beyond individual effects.

Logical connection: Social network analysis can generate datasets in which individuals (smaller units) are grouped into their network clusters (larger units), ready to be analyzed with multilevel models.
Making the Connection

• Variance components (individual variance within groups vs. variance between groups) in multilevel analysis is of special interest in the social network context:

  – It can be interpreted as a measure of sociability of behaviors
  – The larger the between groups variance, the more social is the behavior (vs. individual behavior)

  – At one extreme, if 100% variance is within group, and 0% between groups, the behavior is purely individual... Groups don't matter!

  – At the other extreme, if 0% variance is within group, and 100% between groups, the behavior is purely social... Individuals behave in perfect conformity with their own group and all the variation is between groups!

  – In reality, there is often a division of the variance within and between groups, but different behaviors can be compared in regard to their level of sociability
## Making the connection

<table>
<thead>
<tr>
<th>Activity</th>
<th>Between groups variance (%)</th>
<th>Within groups individual variance (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cigarettes Smoking</td>
<td>15.1*</td>
<td>84.9</td>
</tr>
<tr>
<td>Gun carrying</td>
<td>12.6*</td>
<td>87.4</td>
</tr>
<tr>
<td>Threats</td>
<td>5.3*</td>
<td>94.7</td>
</tr>
<tr>
<td>Other weapon carrying</td>
<td>5*</td>
<td>95</td>
</tr>
<tr>
<td>Heavy drinking</td>
<td>3.5*</td>
<td>96.5</td>
</tr>
<tr>
<td>Theft over 50</td>
<td>2.6</td>
<td>97.4</td>
</tr>
<tr>
<td>Injury</td>
<td>2.1</td>
<td>97.9</td>
</tr>
<tr>
<td>Automobile theft</td>
<td>1.6</td>
<td>98.4</td>
</tr>
<tr>
<td>Theft under 50</td>
<td>1.2</td>
<td>98.8</td>
</tr>
<tr>
<td>Gun use</td>
<td>0.9</td>
<td>99.1</td>
</tr>
<tr>
<td>Marijuana use</td>
<td>0.8</td>
<td>99.2</td>
</tr>
<tr>
<td>Weapon use</td>
<td>0.7</td>
<td>99.3</td>
</tr>
<tr>
<td>Punching</td>
<td>0.4</td>
<td>99.6</td>
</tr>
<tr>
<td>Vandalism</td>
<td>0</td>
<td>100</td>
</tr>
</tbody>
</table>
Empirical example: Delinquency and Peer Networks

• It is well known that delinquents often have delinquent friends.

• However, this pattern may reflect:
  – **Peer influences**: peer group-level factors increasing the risk of delinquency of respondents
  – **Self selection**: individual variations in risk factors for delinquency and seeking out like-minded peers – No peer group effects
  – **Hybrid model**: peer group factors interacting with individual risk factors for delinquency – Peer group influences conditional on individual differences
Empirical example: Delinquency and Peer Networks

• Data: 550 high school students, who could nominate up to 5 close friends
  – Data collected in 2006, from a mid-size high school in Kentucky

• The survey allowed respondents to nominate friends outside of schools, but they mostly nominated friends from school

• Analysis is based on friends from school, since friends must also take the survey
Empirical example: Delinquency and Peer Networks

• Part 1: Social Network Analysis
• Main concept: friendship (peer groups)
• SNA typically combines visualization techniques (mapping) with statistical tests (e.g. network cohesion, actor centrality, network subgroups)
• Based on both visualization and statistical tests, the researcher must make reasonable decisions about the final number of network clusters.
• UCINET software was used for the current study
Empirical example:
Delinquency and Peer Networks

- Raw data
Empirical example: Delinquency and Peer Networks

- Final results
Empirical example: Delinquency and Peer Networks

• The same data are used in the current study
• Based on the final results of the SNA, 56 peer groups (a.k.a. network clusters) were measured
• Final sample of respondents for the multilevel analysis is 514
  – 17 loners were deleted
  – Lost a few more respondents because they did not provide answers for key DVs
Empirical example: Delinquency and Peer Networks

• Part 2: Multilevel Analysis
• Level 1: Delinquency variables and individual risk factors
• Level 2: Group-average for the peer clusters on two key predictors: (1) peer-group average level of low self control; (2) peer-group average level of deviant attitudes
• For simplicity sake, preliminary results are estimated with the linear HLM regression for this example
  – Replicated them with the Poisson HLM regression and the patterns did not change
Empirical example: Delinquency and Peer Networks

• *Substantive question*: do peer-group average levels of low self control and deviant attitude affect respondents’ own level of delinquency, controlling for individual risk factors?
  – If the answer is *yes*: suggest peer influences
  – If the answer is *no*, but individual risk factors are significant: suggest self selection into peer groups

• I did not test the hybrid models (cross-level interactions) for this presentation

• Present results for the 5 delinquent behaviors with significant between-groups variance:
  – Cigarette smoking, Gun carrying, Threats, Other weapon carrying, and Heavy drinking
Empirical example: Delinquency and Peer Networks

- Multilevel analysis predicting Cigarette Smoking

<table>
<thead>
<tr>
<th></th>
<th>Coefficients</th>
<th>P-values</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Peer groups (Level 2)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average Low self control</td>
<td>0.049</td>
<td>0.169</td>
</tr>
<tr>
<td>Average Deviant attitudes</td>
<td>0.409</td>
<td>0.348</td>
</tr>
<tr>
<td><strong>Individual respondents (Level 1)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low self control</td>
<td>0.030</td>
<td>0.020</td>
</tr>
<tr>
<td>Deviant attitudes</td>
<td>0.330</td>
<td>0.124</td>
</tr>
</tbody>
</table>

**Controlling for**: Gender, Age, Minority status, SES, Attachment to parents; Attachment to school
Empirical example: Delinquency and Peer Networks

- Multilevel analysis predicting Gun Carrying

<table>
<thead>
<tr>
<th>Peer groups (Level 2)</th>
<th>Coefficients</th>
<th>P-values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average Low self control</td>
<td>-.002</td>
<td>.922</td>
</tr>
<tr>
<td>Average Deviant attitudes</td>
<td>.002</td>
<td>.992</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Individual respondents (Level 1)</th>
<th>Coefficients</th>
<th>P-values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low self control</td>
<td>.024</td>
<td>.007</td>
</tr>
<tr>
<td>Deviant attitudes</td>
<td>.306</td>
<td>.008</td>
</tr>
</tbody>
</table>

**Controlling for:** Gender, Age, Minority status, SES, Attachment to parents; Attachment to school
Empirical example: Delinquency and Peer Networks

- Multilevel analysis predicting Threats

<table>
<thead>
<tr>
<th>Peer groups (Level 2)</th>
<th>Coefficients</th>
<th>P-values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average Low self control</td>
<td>.022</td>
<td>.393</td>
</tr>
<tr>
<td>Average Deviant attitudes</td>
<td>-.165</td>
<td>.404</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Individual respondents (Level 1)</th>
<th>Coefficients</th>
<th>P-values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low self control</td>
<td>.054</td>
<td>.001</td>
</tr>
<tr>
<td>Deviant attitudes</td>
<td>.448</td>
<td>.001</td>
</tr>
</tbody>
</table>

Controlling for: Gender, Age, Minority status, SES, Attachment to parents; Attachment to school
**Empirical example:**

**Delinquency and Peer Networks**

- Multilevel analysis predicting Other Weapon Carrying

<table>
<thead>
<tr>
<th></th>
<th>Coefficients</th>
<th>P-values</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Peer groups (Level 2)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average Low self control</td>
<td>-.016</td>
<td>.286</td>
</tr>
<tr>
<td>Average Deviant attitudes</td>
<td>.082</td>
<td>.697</td>
</tr>
<tr>
<td><strong>Individual respondents (Level 1)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low self control</td>
<td>.016</td>
<td>.071</td>
</tr>
<tr>
<td>Deviant attitudes</td>
<td>.339</td>
<td>.003</td>
</tr>
</tbody>
</table>

**Controlling for:** Gender, Age, Minority status, SES, Attachment to parents; Attachment to school
Empirical example: Delinquency and Peer Networks

- Multilevel analysis predicting Heavy Drinking

<table>
<thead>
<tr>
<th></th>
<th>Coefficients</th>
<th>P-values</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Peer groups (Level 2)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average Low self control</td>
<td>-.028</td>
<td>.219</td>
</tr>
<tr>
<td>Average Deviant attitudes</td>
<td>.163</td>
<td>.545</td>
</tr>
<tr>
<td><strong>Individual respondents (Level 1)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low self control</td>
<td>.035</td>
<td>.002</td>
</tr>
<tr>
<td>Deviant attitudes</td>
<td>.328</td>
<td>.008</td>
</tr>
</tbody>
</table>

**Controlling for:** Gender, Age, Minority status, SES, Attachment to parents; Attachment to school
Empirical example: Delinquency and Peer Networks

• Preliminary Discussion:
  – Results from variance components show that some delinquent behaviors are more social (smoking, gun carrying), while other delinquent behaviors are more individual (vandalism, punching someone)
  – Results do not show support for group-level peer influences in term of average levels of low self control or deviant attitudes
  – Rather, they show fairly consistent support for the effects of individual levels of low self control and deviant attitudes
  – Thus, preliminary results suggest that self selection may play a greater role than peer influences in explaining why delinquents tend to have delinquent friends
Limitations

• Early decisions in the survey design and data collection have important implications at later stages
  – Researchers must think in terms of sampling of respondents but also sampling universe for the network
  – Results will vary based on sampling universe for the network: people at work, people in your neighborhood, people at school, people on Facebook, etc.
  – The number of possible contacts: identify your 3 best friends from work (vs. 5, vs. 10)
  – Wording:
    • “Identify your best friends” vs. “Identify people you trust the most”
    • “Identify people with the most power in your workplace” vs. “Identify people who directly exercise power over you on a daily basis in your workplace”
Limitations

• Creating network clusters is not an exact science. Decisions must be made by the researcher (include/exclude people; merge/divide network clusters)
• Loners typically become missing values for multilevel analysis. Large number of loners can become a sample bias.
• Network clusters are assumed to be fairly stable in multilevel models. If they are unstable or very dynamic in the real world, the chances of finding contextual effects become very small.
  – I.e. Level 2 network clusters could be statistical artifacts
• People may belong to different networks at different times of day, or days of the week, or months of the year
  – During business hours: workplace networks
  – In the evening: friends and family networks
  – On weekends: networks related to hobbies / entertainment
Limitations

• Social desirability bias: some respondents may report more connections with more popular, attractive, or powerful people than reality.
  – Ex. A teenage boy spends 80% of his free time playing online video games with other boys from school, but list the 5 most attractive girls at school as his best friends.
  – Creative use of survey questions and control variables may help to decrease this bias.
References


