Homophily and Delinquency in Friendship Networks of Adolescents: association between similarity and tie formation in a network

Samantha Daoud

4/10/2021

Introduction

Social network analysis is important to examine the process of the social structure. In general, social networks are dynamic and recent researches have focused on these dynamics as they are important in friendship networks (Burk, Steglich, and Snijders, 2007) and using the stochastic actor oriented model will help to represent these dynamics driven by many tendencies (micro mechanisms). Friendship relationships are established based on forming ties between social actors who are in in this paper: students. These ties may develop or disappear representing the relational changes between actors within a network. In the case of friendship, these ties are expected to be the dyadic basics of the friendship relations between two actors. The influence of ties existence sometimes are related to determinants of similarity between individuals like having the same gender or ethnicity, sharing the same educational background or behaviour. Thus, the research question here is to what degree the ties formation in friendship networks is affected by the determinants of similarity?. This will help us to understand the different dynamics of friendship networks and show the different mechanisms of these networks.

In this paper, I also will examine the effect of delinquency behaviour on the tie formation in the friendship network of adolescents and also will examine demographic characteristics of the respondents like sex and ethnicity on friendship networks and on delinquency behaviour. Moreover, we will have better understanding to the similarity effect of demographic characteristics and behaviour by analysing data of four waves from a dutch school using Stochastic Actor-Oriented Models (SAOMs). Reasons behind hoosing a dutch secondary school as a population for my study are: the peer influence on the behaviour (Giordano, 2003) and the time that students are spending in class has a role in forming their identity and the define degree of socialising with others (Knecht, Snijders, Baerveldt, Steglich, & Raub, 2010).

Theoretical background

Earlier researches show the different mechanisms that prompt to develop the structure of social networks and evolve the ties formation. In friendship networks, reciprocity, transitivity and homophily are some of the main characteristics (Girvan, Newman, 2001; Knecht, A., 2008; Koskinen, Snijders, 2007) . And in our data sample, and based on the demographic characteristics, it is assumed that the networks are reciprocated i.e the recipient of tie is more likely to resend the tie to the same individual, transitive i.e the friend of a friend is more likely to be your friend and highly segregated on the level of sex/gender (gender homophily).

Homophily

Homophily can be introduced as the the tendency or preference to have friends or form ties with others who have the same demographic characteristics or behaviours (McPherson et al., 2001; Kandel, 1978). It

may be introduced as a key factor in tie formation and shaping friendship relationships. One simple kind of this homophily is the gender homophily where two adolescents of the same sex are more likely to be friends Moreover, homophily can be existed in many other attributes and friendship relationships. If two adolescents are are from the same school or educational background they are more likely to be friends. These different types of homophily may have an effect on the friends' behaviours, they are more likely to have the same risky behaviour if they are homophilous in one attribute or characteristic (Kandel, 1978). Peer influence and friend selection are two mechanisms that help us to have a better understanding on how homophily on delinquency behaviour can be observed within friendship relations of adolescents (Kandel, 1978).It is assumed that friendship networks are not formed randomly but based on selectivity. It is more possible that among friends who share the same demographic characteristics to be also similar in the behaviour (behavioural homophily).

Behaviour

Risky behaviour of adolescents has been the subject of research for many studies and this holds to many kinds like drugs usage (Kandel, 1996), alcohol (Sieving, Perry, & Williams, 2000), smoking (Ennett & Bauman, 1994) and delinquency (Baerveldt, Völker & Van Rossem, 2008, Haynie & Osgood, 2005). The friendship ties in the social networks affect and is being effected by such behaviours (Haas et al., 2010). Studies found that in general, adolescents are more likely to participate in a risky behaviour because their friends do. They may select friends who share the same risky behaviour or they may have a peer influence to act like them (Knecht, Snijders, Baerveldt, Steglich, & Raub, 2010). Researches assume that status characteristics may have an effect on the behaviours of adolescents as characteristics like sex or ethnicity play key role on the culture and habits of individuals. So students with similar demographic/ status characteristics are more likely to share the same interests and behaviours (Kao & Joyner, 2004).

In this paper, I focus on these two demographic or status characteristics: gender/ sex and ethnicity because these two have the most effect on the habits and behaviour of adolescents.

Hypotheses

Based on literature, I draw the following hypotheses:

H1. Friendship networks have low density as networks formation is not random so the out degree is expected to be negative. Also Friendship networks have high reciprocity and transitivity. Which means that it has more incoming and outgoing ties between actors.

H2. Students who share the same school background are more likely to reciprocate friendship than with others (school place homophily).However, this will create a segregated networks in term of school background.

H3. Students who are similar in the status or demographic characteristics like sex or ethnicity are more likely to reciprocate friendship ties between each other. However, this will create a segregated networks in term of sex or ethnicity.

H4. Students who are delinquent ask less for friendship (less outgoing friendship ties) but expected be asked more for friendship (attract more incoming friendship ties). While who have behavioural similarity are more likely to ask each other for friendship.

H5. There is a significant effect for the status characteristics like sex and ethnicity on delinquency behaviour.

Data and methods

Data description

For this analysis, I used dataset from RSiena website and were collected between 2003 and 2004 as part of study in Knecht (2008). The data includes four waves of panel data for 26 students of which 17 girls and 9 boys in class followed over their first year at a secondary school to assess the friendship networks at four time points at intervals of three months. Students aged 11-13 at the beginning of school. The assessment of the friendship networks data were done by asking students nominate up to 12 classmates who consider them as good friends. Jaccard index that measured the change amount between each two consecutive waves shows values by .4 which indicates as safe and represent a status of stability. Hamming distance index shows the minimum number of mini steps between each two networks and indicates a sufficient number of observed changes.

Measures

This dataset contains many variables that based on the students' responses to survey questions and ranging on scales . For instance, network variable "friendship" was coded 0 = "no" friendship relation and 1 = "yes", there is a friendship relation, 9 = missing and 10 = not a member of the classroom (structural zero). Change actor variable "delinquency", rounded average of four types of delinquent activity (stealing, vandalizing, fighting, graffiti) measured in each of the four waves and coded as (1 = never, 2 = once, 3 = 2-4 times, 4 = 5-10 times, 5 = more than 10 times, 0 = missing). Constant actor variables: like sex coded as (1 = girl, 2 = boy), Ethnicity coded as (1 = Dutch, 2 = other, 0 = missing) and other variables like age and religion. "School advice" variable which indicates the school capacities of the student , coded as (4 = low, 8 = high, 0 = missing). Changing dyadic variable "primary" which indicating having gone to the same primary school, coded as (0 = no, 1 = yes).

The following figures show the graphs of the four networks, each network consists of one component, all networks are directed.
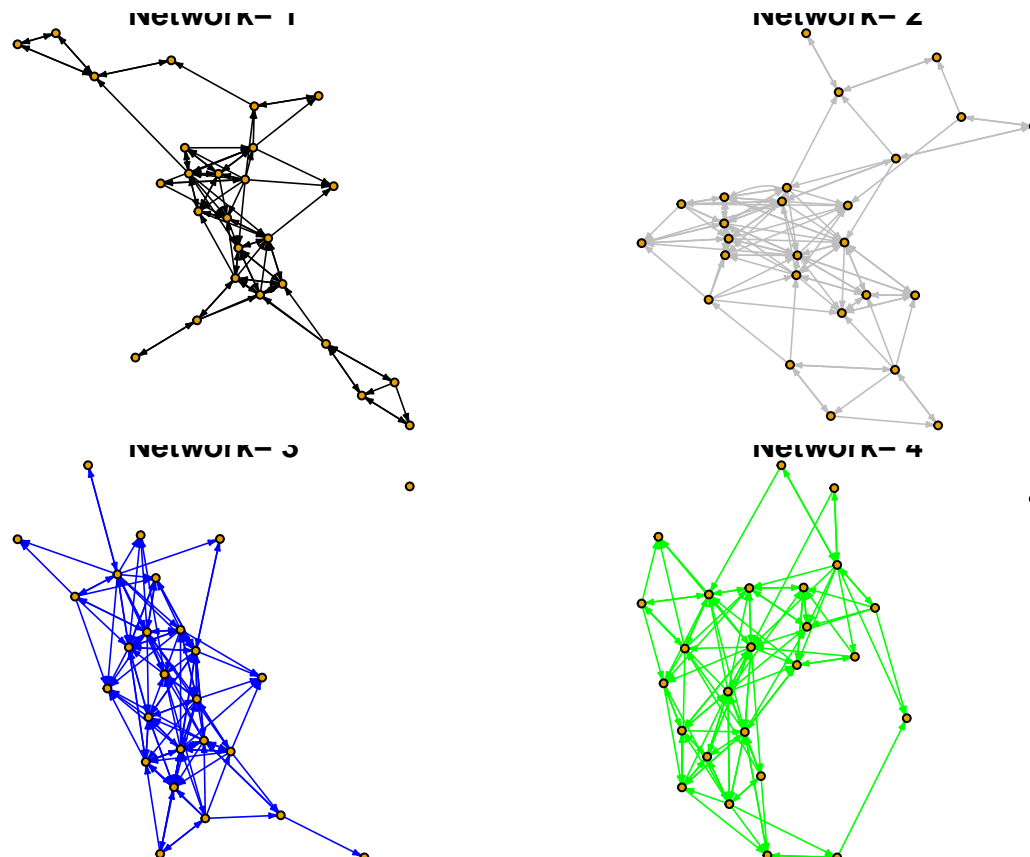
The below tables show the descriptive statistics of the four networks and the key variables.

In table 1, we can see that the parameters show that the mean of network 3 (10.23) is higher than other networks, while the means of network 2 and 4 are nearly the same (9.00, 9.15) while it is the least in network 1 (7.00).

Table 1: Descriptive statistics of longitudinal networks means

	Network 1	Network 2	Network 3	Network 4
Means	7	9	10.23	9.15

In table 2, we notice the changes' values indicate the ministeps between each two consecutive waves, we notice that for 1st period, changes are 86 while for the 2nd and 3rd periods are 98. Total of changes for the 3 periods are 282 which represent the no. of discrete choices that actors are taking.

Table 2: Descriptive statistics of changes between 3 waves

	Network 1-2	Network 2-3	Network 3-4
Changes (ministeps)	86	98	98

In table 3, we see that the means of advice is the highest (5.92), following delinquency mean is (1.74) while the means of constant variables ethnicity and gender are nearly the same (1.08, 1.35). Thus, based on these statistics, Hamming and Jaccard testing, networks data are suitable to be analysed using stochastic actor-oriented model.

Table 3: Descriptive statistics of key variables means

	delinquency	ethnicity	advice	gender
Means	1.74	1.08	5.92	1.35

Data recoding

The data contains missing values due to the absence of students and were treated using the ad-hoc model using imputation, (Huisman and Steglich, 2008) to avoid unbiased parameter estimates and get accurate standard error estimates. Students who left the classroom cause changes in the network composition and was treated by using the "structural zeros" i.e, all in-out tie variables of the left students were treated as zeros. Recoding the missing and structural zeros in all measures would improve the convergency of our models.

Stochastic Actor-Oriented Model

In order to test my research question and hypotheses, I used stochastic Actor-Oriented model (SAOM) estimated by RSiena software . I created data object for SAOM by joining all waves of the friendship network in addition to other variables. The data we got includes two dependent variables : (friendship and delinquency) where the type of friendship is one mode which indicates social ties, i.e people are connected to people, with four observations and with densities that goes up until the third network then we notice a slight decrease (0.14, 0.19, 0.22, 0.18). While the type of delinquency as behaviour with 4 observations and the race of actors are from 1 to 5. Constant covariates are: "sex" and "ethnicity". Dyadic covariate is "primary".

Then I specified a co-evolution empty model for the data by applying "get effects" function and fitted a simple model including only the the effects of transitivity and the interaction of reciprocity with transitivity and another model includes only the degrees related-effects for the friendship dynamics. Then I added interactions with gender, primary (same school) and the behaviour variable delinquency to test the homophily, sender and receiver effect of delinquency on friendship. Also added effects to the behavior objective function with interactions with gender and ethnicity.

To assess the model convergence on the network dimensions, we can calculate the degree to which the model-simulated data are entered around the observed data. In our data, the convergence t-ratio for all parameters is very close to zero i.e below 0.1 and above -0.1, which means that the convergence has been achieved and it is a converged model so we can interpret the estimates. Also we see the overall maximum convergence ratio is 0.15 which is below .20 so it is considered as a good value.

```
## siena07 will create an output file Siena.txt .

## Estimates, standard errors and convergence t-ratios
##
##                                              Estimate   Standard   Convergence
##                                                          Error       t-ratio
## Network Dynamics
##     1. rate constant friendship rate (period 1)  7.1845  ( 1.6802  )   -0.0179
##     2. rate constant friendship rate (period 2)  7.7309  ( 1.2707  )   -0.0660
##     3. rate constant friendship rate (period 3)  7.2719  ( 1.2435  )   -0.0126
##     4. eval outdegree (density)                 -1.6157  ( 0.3332  )   -0.0455
##     5. eval reciprocity                          1.7982  ( 0.3155  )   -0.0953
##     6. eval transitive triplets                  0.5129  ( 0.0944  )   -0.0834
##     7. eval transitive recipr. triplets         -0.3743  ( 0.1266  )   -0.1256
##     8. eval indegree - popularity               -0.0525  ( 0.0357  )   -0.0426
##     9. eval outdegree - activity                 0.0140  ( 0.0199  )   -0.0690
##    10. eval indegree - activity                 -0.2085  ( 0.0817  )   -0.0681
##    11. eval primary                              0.9045  ( 0.2242  )    0.0240
##    12. eval same sex                             0.6046  ( 0.1574  )   -0.0401
##    13. eval delinquency alter                    0.3338  ( 0.2447  )   -0.0090
##    14. eval delinquency ego                     -0.1022  ( 0.1837  )    0.0023
##    15. eval delinquency similarity               3.1402  ( 1.8352  )   -0.0022
##
## Behavior Dynamics
##    16. rate rate delinquency (period 1)          1.1570  ( 0.4747  )   -0.0049
##    17. rate rate delinquency (period 2)          2.2628  ( 1.2471  )   -0.0680
##    18. rate rate delinquency (period 3)          2.3864  ( 1.1407  )    0.0245
##    19. eval delinquency linear shape            -0.1497  ( 0.2225  )   -0.0301
##    20. eval delinquency quadratic shape         -0.1061  ( 0.1368  )   -0.0417
##    21. eval delinquency: effect from sex        -0.0145  ( 0.3174  )   -0.0290
##    22. eval delinquency: effect from ethnicity   0.7438  ( 0.6012  )    0.0524
##
## Overall maximum convergence ratio:     0.2143
##
##
## Total of 3751 iteration steps.
```

Goodness of Fit

I assess the goodness of fit by checking the characteristics that are not explicitly including in the model specification. So first I calculate the fit with respect to the indegree distribution using SienaGOF. As we see

in the below graph, the red dots are the observed values, while the grey dots are the simulated data. We find a significant misfit and see a departure of the observed data outside the simulated data at point (17) in category 2 and point (2) in category 3. So it is a significant departure with p-value (0.03).

```
##   > Completed  1000  calculations
##
##   > Completed  1000  calculations
##
##   > Completed  1000  calculations
```

Goodness of Fit of IndegreeDistribution

p: 0.071
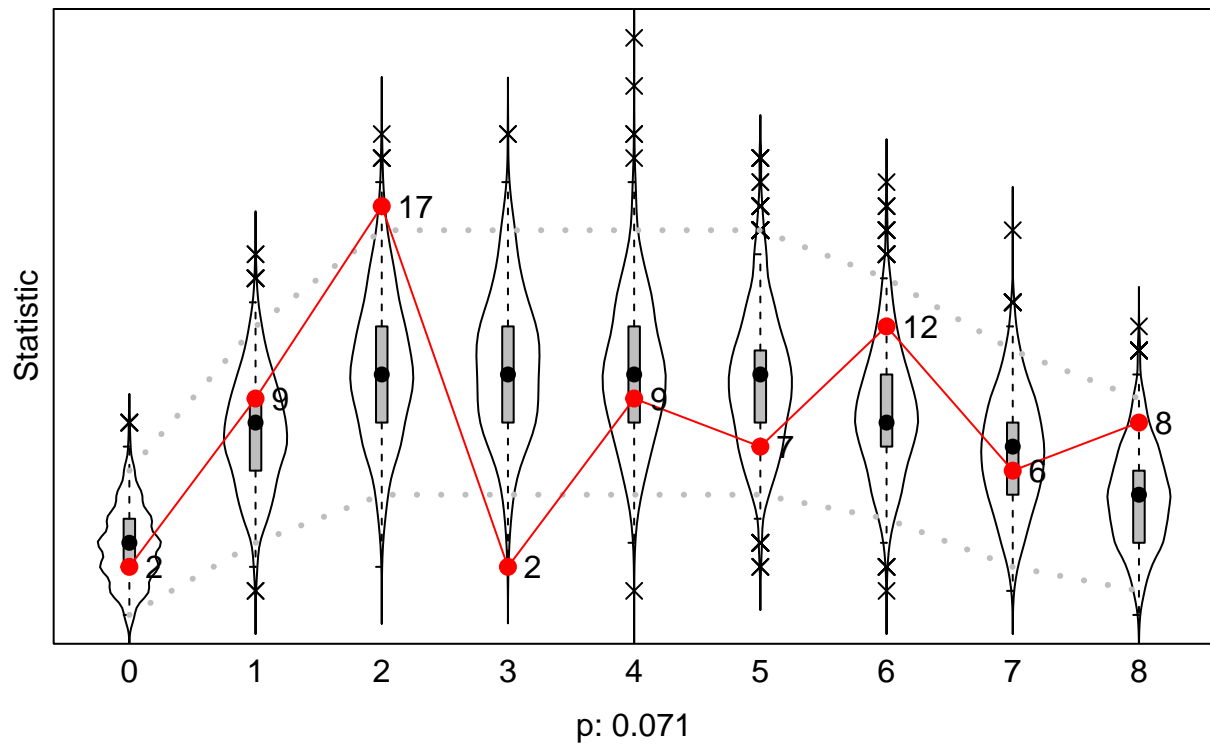
In order to obtain a better fit, I try to add more subgraphs, so now I include the out degree distribution to the model as shown in below graph. It looks better than the last graph, we find an insignificant fit where the observed data represented by the red line fits well in the simulated data but still there is a small departure at point (1) in category 7 and point 8 in category 8 with p- value (0.1).

Goodness of Fit of OutdegreeDistribution

p: 0.111
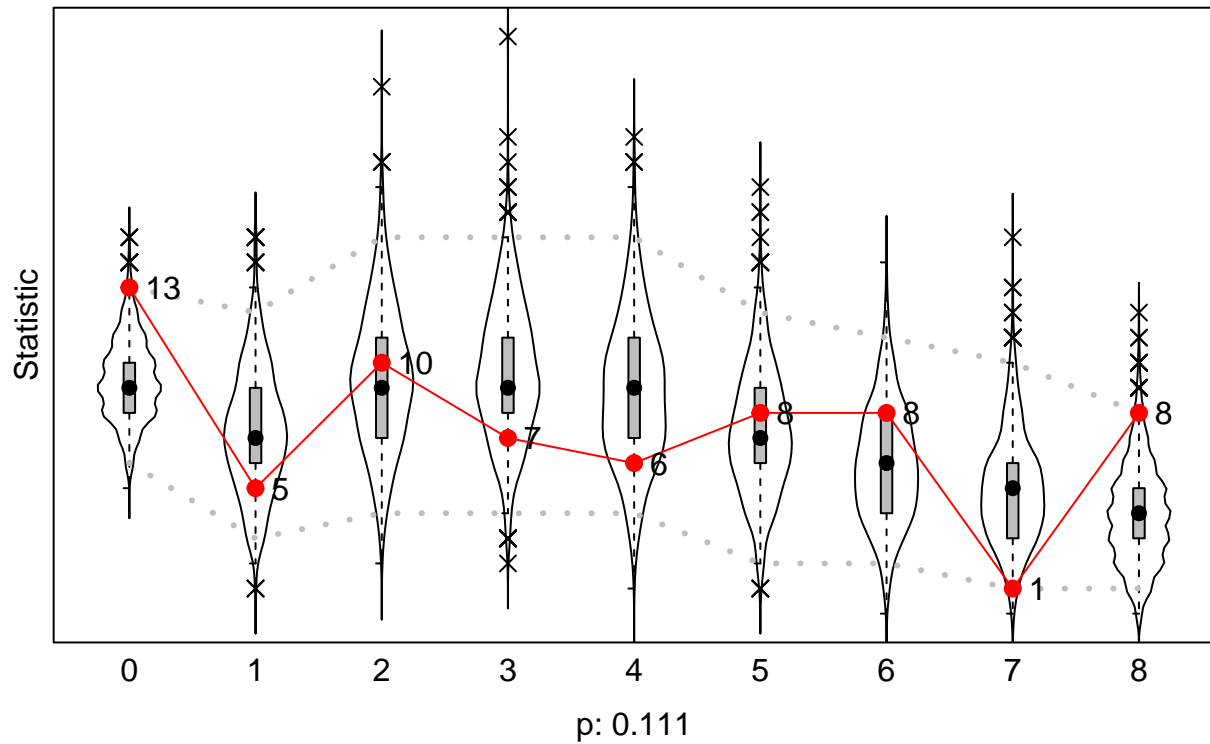
Now I add more subgraphs like "TriadCensus" as shown in the below graph. We obtain an insignificant better fit as the red lines lies in the middle of the simulated data with no departure and with p-value (0.99).

Goodness of Fit of TriadCensus

p: 0.999
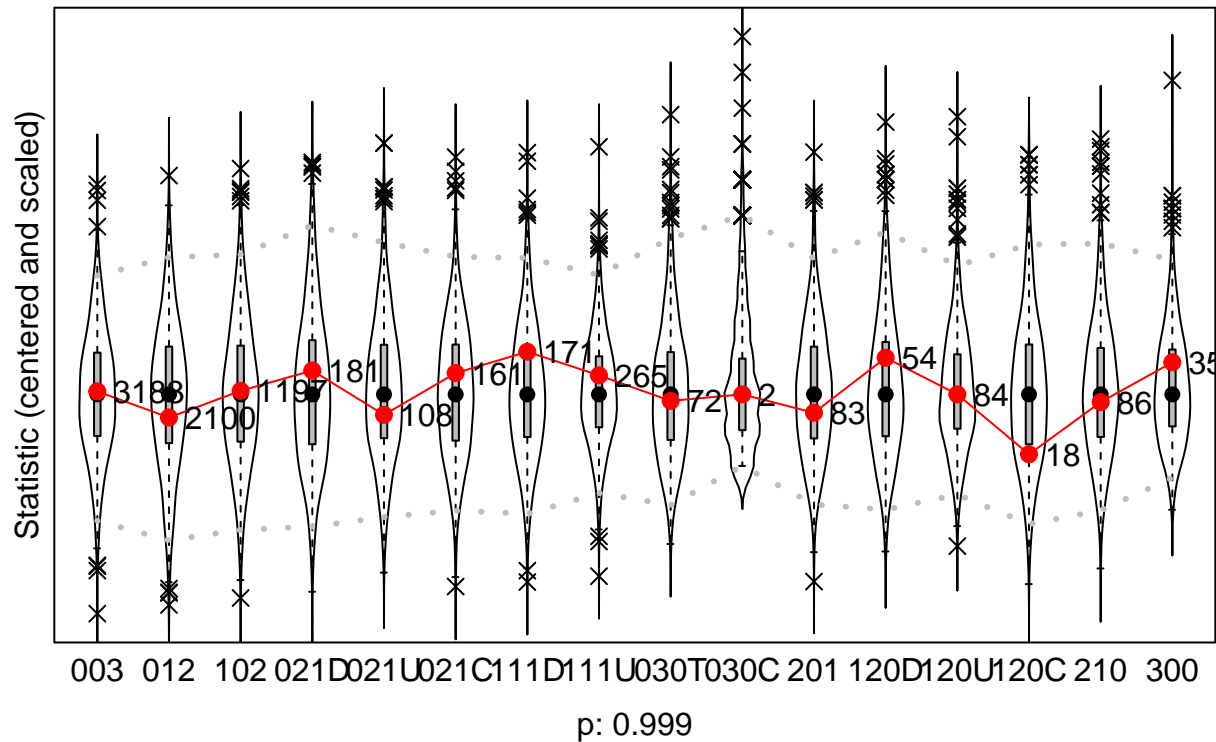
Calculating the goodness of fit for the overall behaviour distribution "delinquency" as shown in below graph, it looks good and we obtain a better fit but insignificant as the p-value is (0.96).

```
## Note: some statistics are not plotted because their variance is 0.
## This holds for the statistic: 5.
```

Goodness of Fit of BehaviorDistribution

p: 0.952
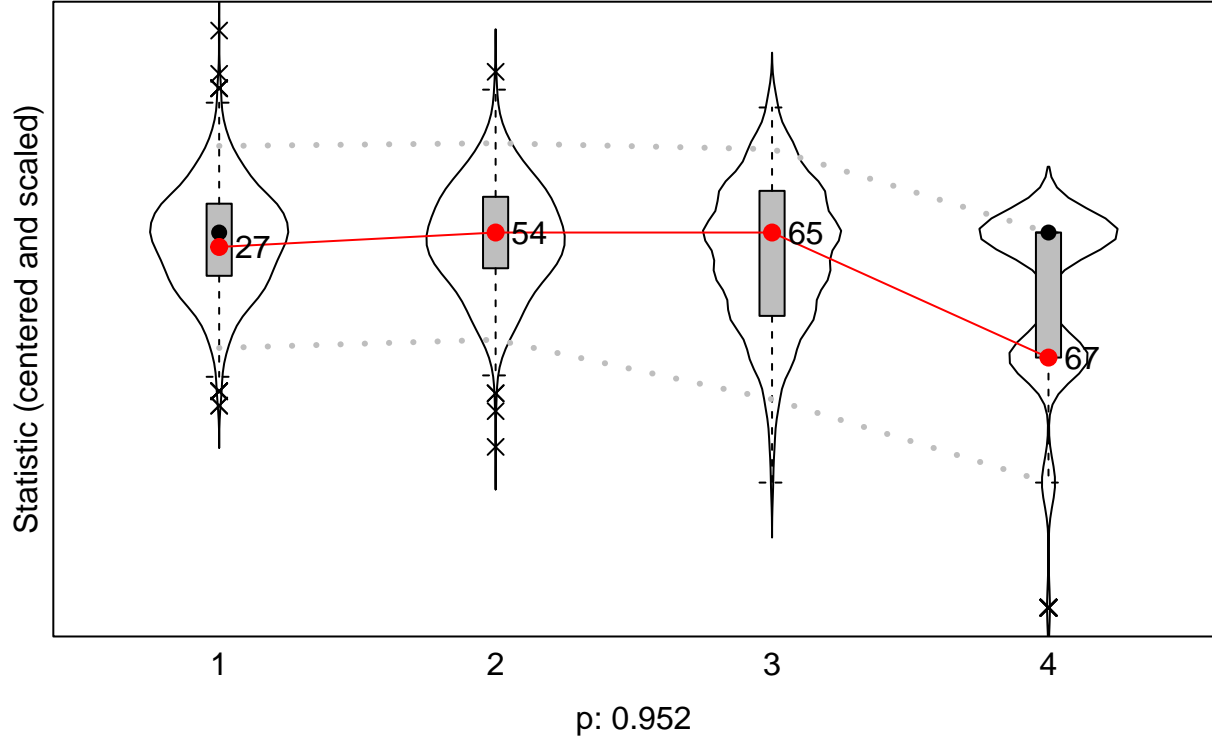
Results

Table 4 presents the results. We report the parameters' estimate, their standard errors and the p values. We notice that most of the coefficients are statistically significant. The first part of the table is about the network dynamics where we find the rate parameters which show the amount of friendship changes during the periods. We notice these changes slightly peak in period 2 (7.7) , they decrease (7.2) in period 3. All changes are significant at the level of (p < .001).

The outdegree (density) parameter is negative and significant (-1.6) which means that there is general tendency not to have ties to many others at a random. They avoid contact in general and prefer few new ties. Reciprocity parameter is positive (1.8) which means they prefer ties that are reciprocated and embedded in transitive triplets (0.5). While the interaction of reciprocity with transitivity is negative (-0.4). These parameters are significant at the level of (p < .001). This support hypothesis *(H1)*.

The parameter estimates of variables like primary- same school (0.9) and same sex (0.6) show that there is main effect of homophily on the level of same school and same sex, which means students who are from the same school or the same sex are more likely to be friends. The homophily index of same school and same sex is significant at the level of (p < .001). This support hypotheses *(H2 & H3)*.

The parameter estimate of delinquency alter, has positive effect (0.3), which means that students with bad behaviour i.e who are delinquent will be asked to be friends. While the parameter estimate of delinquency ego has negative effect (-0.09) which means that students who are delinquent will not ask other to be friends. Delinquency similarity has positive and high effect (3.1) which means that students who are similar with delinquency behaviour will ask each other to be friends. but these effects of the behaviour delinquency are not significant. This rejects hypothesis *(H4)*.

The second part is about the behaviour dynamics: where we find the rate parameters estimates of the changes of the delinquency variable during the three periods. We notice the changes increase gradually starting from

the 1st period (1.2) to 3rd period (2.3). The estimates of period 1 and period 2 are significant at the level of (p < .05) while period 2 is not significant.

We find positive effect for sex and ethnicity on delinquency but these effect are not significant. This rejects hypothesis *(H5)*.

Table 4: Descriptive statistics of the parameters and p-values

parameter	estimate	st.error	normal.variate	p.value
constant friendship rate (period 1)	7.185	1.680	4.28	0.0000
constant friendship rate (period 2)	7.731	1.271	6.08	0.0000
constant friendship rate (period 3)	7.272	1.243	5.85	0.0000
outdegree (density)	-1.616	0.333	-4.85	0.0000
reciprocity	1.798	0.316	5.70	0.0000
transitive triplets	0.513	0.094	5.43	0.0000
transitive recipr. triplets	-0.374	0.127	-2.96	0.0031
indegree - popularity	-0.053	0.036	-1.47	0.1410
outdegree - activity	0.014	0.020	0.70	0.4823
indegree - activity	-0.209	0.082	-2.55	0.0107
primary	0.904	0.224	4.03	0.0001
same sex	0.605	0.157	3.84	0.0001
delinquency alter	0.334	0.245	1.36	0.1725
delinquency ego	-0.102	0.184	-0.56	0.5781
delinquency similarity	3.140	1.835	1.71	0.0871
rate delinquency (period 1)	1.157	0.475	2.44	0.0148
rate delinquency (period 2)	2.263	1.247	1.81	0.0696
rate delinquency (period 3)	2.386	1.141	2.09	0.0364
delinquency linear shape	-0.150	0.223	-0.67	0.5013
delinquency quadratic shape	-0.106	0.137	-0.78	0.4379
delinquency: effect from sex	-0.014	0.317	-0.05	0.9636
delinquency: effect from ethnicity	0.744	0.601	1.24	0.2161

Conclusions and discussion

In this paper, I discussed the question of research on to what degree tie formation is affected by determinants of similarity and posed hypotheses and investigated them through SAOM models. I also examined the network dynamics on the micro level. The descriptive results and models show that friendship networks of adolescents in the dutch school are not randomly formed. It is based on gender preferences, selection preferences on the level of behaviour. Similarity in the level of sex and ethnicity affect significantly on tie formation within friendship networks. While there is insignificant effect for the similarity of behavior on the tie formation. Nonetheless, this study was only for time interval of three months; so for future research we can investigate similarity on a longer time periods and different behaviours.

References

Baerveldt, C., Völker, B., & Van Rossem, R. (2008). Revisiting selection and influence: An inquiry into the friendship networks of high school students and their association with delinquency. Revue Canadienne de

Criminologie et de Justice Pénale, Octobre, 559-587.

Brass, D.J., Galaskiewicz, J., Greve, H.R., Tsai, W., 2004. Taking stock of networks and organizations: a multilevel perspective. Academy of Management Journal 47, 795–817. Burk, W.J., Steglich, C.E.G., Snijders, T.A.B., 2007. Beyond dyadic interdependence: Actor-oriented models for co-evolving social networks and individual behaviors. International Journal of Behavioral Development 31, 397–404.

Ennett, S. T., & Bauman, K. E. (1994). The contribution of influence and selection to adolescent peer group homogeneity: The case of adolescent cigarette smoking. Journal of Personality and Social Psychology, 67, 653-666.

Haynie, D. L., & Osgood, W. D. (2005). Reconsidering peers and delinquency: How do peers matter? Social Forces, 84(2), 1109-1130.

Haas, S.A., Schaefer, D.R., Kornienko, O., 2010. Health and the structure of adolescent social networks. J. Health Soc. Behav. 51, 424–439.

Jin, E.M., Girvan, M., Newman, M.E.J., 2001. Structure of growing social networks. Physical review E, 64, 046132.

Kandel, D. B. (1996). The parental and peer contexts of adolescent deviance: An algebra of interpersonal influences. Journal of Drug Issues, 26(2), 289-315.

Kandel, D.B., 1978. Homophily, selection, and socialization in adolescent friendships. Am. J. Sociol. 84, 427–436.

Kao, G., & Joyner, K. (2004). Do race and ethnicity matter among friends? Activities among interracial, interethnic, and intraethnic adolescent friends. Sociological Quarterly, 45(3), 557- 573.

Knecht, A., 2008. Friendship Selection and Friends' Influence. Dynamics of Networks and Actor Attributes in Early Adolescence. PhD dissertation, University of Utrecht.

Knecht A, Snijders TAB, Baerveldt C, Steglich CEG, Raub W. Friendship and delinquency: Selection and influence processes in early adolescence. Social Development. 2010; 19:494–514.

Koskinen, J.H., Snijders, T.A.B., 2007. Bayesian inference for dynamic network data. Journal of Statistical Planning and Inference 13, 3930–3938.

M. Huisman, C. Steglich, 2008. Treatment of non-response in longitudinal network studies. Social Networks 30 (2008) 297–308.

McPherson, M., Smith-Lovin, L., Cook, J.M., 2001. Birds of a feather: homophily in social networks. Annu. Rev. Sociol. 27, 415–444.

Skyrms, B., Pemantle, R., 2000. A dynamic model of social network formation. PNAS 97, 9340–9346.

Sieving, R. E., Perry, C. L., & Williams, C. L. (2000). Do friendships change behaviors, or do behaviors change friendships? Examining paths of influence in young adolescents' alcohol use. Journal of Adolescent Health, 26(1), 27-35.

Snijders, T.A.B., 2001. The statistical evaluation of social network dynamics. Sociol.Methodol. 31, 361–395.

Snijders, T.A.B., 1996. Stochastic actor-oriented dynamic network analysis. Journal of Mathematical Sociology 21, 149–172.

Snijders, T.A.B., 2001. The statistical evaluation of social network dynamics. In: Sobel, M., Becker, M. (Eds.), Sociological Methodology, Basil Blackwell, Boston and London, pp. 361–395.

Snijders, T.A.B., Steglich, C.E.G., and van de Bunt, G.G. (2010), Introduction to actor-based models for network dynamics, Social Networks 32, 44-60.