Chapter 8

Groups: Online Appendix

In Chapter 8, we have reviewed studies that use data on marriages and friendship to learn more about intergroup relations. A key question that researchers pose, is to find out what causes the general tendency of group segregation. How do you know whether group segregation is due to opportunity and/or preference (homophily), if you only have data on *actual* choices and not on *preferences*?

Let's review in more detail how scholars attempt to detect the role of opportunities and preferences, and let's begin with marriages. The core idea, which scholars use to find out whether homophily can be an underlying mechanism of group segregation, is to compare *actual* behavior with what would have happened if people would make decisions entirely *random*, group blind. An intuitive measure that captures this idea is the *homophily index*:

$$\textit{Homophily index} = \frac{\textit{Group segregation index (observed)}}{\textit{Group segregation index (random)}} \tag{8.2}$$

If we see that the observed marriages are equally segregated as what would be expected under random choices, then the homophily index is 1. In that case, there is no evidence for homophily. If the observed network is more homogenous than expected, the homophily index exceeds 1, and this indicates that people *prefer* in-group ties.

To see how this works, consider the cross-classification of a fictitious population with 4700 Black and White marriages (Table 8A). Does the table show evidence for racial homophily, as expected by homophily theory?

Table 8A Marriages between Blacks and Whites: Observed

	Females		Total
Males	**Black**	**White**	
Black	500	1500	2000
White	200	2500	2700
Total	700	4000	4700

To answer this question, we first have to compute the *observed* group segregation, which is the number of group-bonding ties divided by the total number of ties (group-bonding and group-bridging). The number of endogamous marriages (group-bonding ties) is 500 + 2500 = 3000 in total. This means that the observed group segregation index is 3000 / 4700 = 0.64. Is this more than what you would expect by chance?

THINKING LIKE A SOCIOLOGIST

Can you come up with a way to compute the expected number of endogamous marriages in this example, assuming people marry 'randomly'?

If Blacks and Whites would *randomly* make their marriage decisions, then, intuitively, it seems that half of all marriages (4700) should be endogamous. But the two groups differ in size, and so you would actually expect a little bit more endogamous marriages than 50%. To see this, you need to compute the expected numbers for each cell, which you can do by multiplying the marginal distributions of the row and column, and then dividing by the total number of marriages. For example, you would expect to see that of the 700 Black females a total of ((700 × 2000) / 4700 =) 298 would marry with Black males. If you would do this computation for each cell, you get Table 8B. With this table, you can now see that, if marriage decisions in this fictitious population would be group-blind, then it would result in (298 + 2298=) 2596 endogamous marriages. The group segregation index under random choices would then be 2596/4700 = 0.55. And the homophily index is therefore 0.64/0.55 = 1.16. This indicates that there are more in-group marriages than expected by chance and structural forces (unequal group size).

Table 8B Marriages between Blacks and Whites: Expected

	Females		**Total**
Males	**Black**	**White**	
Black	298	1702	2000
White	402	2298	2700
Total	700	4000	4700

Another commonly used measure to detect homophily in marriage markets, is the so-called *odds ratio* (Kalmijn, 1998). To compute the odds ratio, you have to take three steps.

Step 1. Consider the odds that people marry someone from their own group as compared to someone of another other group. In our example (Table 8A), 500 Black males marry Black females, whereas 1500 Black males marry White females. This gives an odds of (500/1500=) 1/3.

Step 2. Subsequently, you compute the odds for another group. In this case: White males. Among this group, 200 marry Black females, and 2500 marry White females. That's an odds of (200/2500=) 1/12.5.

Step 3. In the third step, you consider the ratio of the odds of the two groups. Thus, the odds for Black males (1/3) is divided by the odds for White males (1/12.5), which gives you an odds ratio of 4.2.

What does this odds ratio tell us? It indicates that the odds are not the same for White and Black males. If they would be the same, then you would get an odds ratio of 1. To see this, suppose that White males would have equal odds as Black males, namely 1/3 (and not 1/12.5). This would mean that White and Black males have the *same* odds to marry Black women, as compared to marry White women. That would mean that there is no preference to marry in-group members. The odds ratio, in that case, is (1/3 divided by 1/3=) 1. If the odds ratio is higher than 1, it indicates in-group preferences. The higher the odds ratio, the more strongly such in-group preferences. In our example, the odds ratio is 4.2, which indicates quite strong racial homophily. Thus, in evaluating the odds ratio, you need to take a fourth step:

Step 4. Is the odds ratio 1? Is it, using the symbols from Table 8.1 (Chapter 8), the case that $(C_{BB} / C_{BW}) = (C_{WB} / C_{WW})$? If so, then there is no evidence for homophily in marriage choices. Is the odds ratio higher than 1? Is it that (C_{BB} / C_{BW}) is higher than (C_{WB} / C_{WW})? If so, then there is evidence for homophily.

Scholars also study whether *friendship* networks are driven by opportunities and/or homophily. Let's illustrate their approach with an example. Figure 8.9 presents a social network of nine members of two groups, say Muslims (black nodes) and Christians (white nodes).

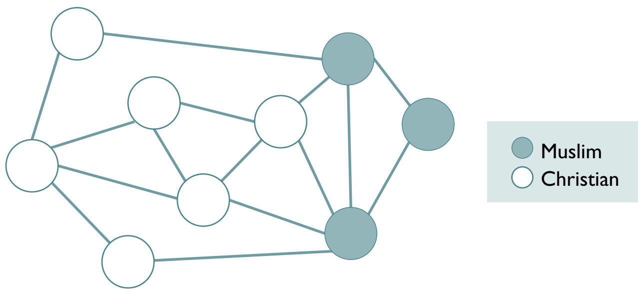

Figure 8A A social network of two groups: Muslims and Christians

The observed network is a 'sample' (snapshot) of a larger population (say the entire population of the country). Is there evidence for religious homophily? Again, the starting point to answer this question, is to compute the *observed* group segregation. There are 10 group-bonding ties and 15 ties in total, so the fraction of group-bonding ties is 10/15 = 0.67. The question is whether this observed group segregation is more than you would expect if Muslims and Christians would make friendships *randomly*. What would the network look like if they would make choices irrespective of religious identity?

THINKING LIKE A SOCIOLOGIST

Can you come up with a way to answer this question? What would be a way to study this?

Since we have two groups in the population, a fraction *m* of all individuals are Muslim and a fraction *c* of them are Christian. This means that to each node we can randomly assign with probability *m* the religion “Muslim”, and with probability *c* the religion “Christian”. There are 9 individuals in this network, of which 3 are Muslim (so $m = 1/3$) and 6 are Christian (so $c = 2/3$). This means that, if we consider a relation between two individuals in this network, the probability that both individuals are Muslim is 1/3 * 1/3 (= 1/9). The probability that a tie is established between two Christians is 2/3 * 2/3 (= 4/9). Taken together, we expect to see a group segregation of $m^2 + c^2 = (1/3)^2 + (2/3)^2 = (1/9 + 4/9 =)\ 0.56$, under conditions of *random choices*. Given that the *actual* friendship network contains a higher fraction (i.e., 0.67) of group-bonding ties than what would be expected by *random* assignment (i.e., 0.56), we could conclude that there is a tendency towards religious homophily.

If you do not study a *sample* from a larger population, but rather *all actors* in a specific setting –such as a school- then you study a so-called *complete network* (Kadushin, 2012). In that case, all pupils have equal opportunities to nominate pupils from their in- and out-group as their friends in school. Computationally, when studying complete networks (rather than a sample from a population) the probabilities for random tie-formation are different.

Let’s use Figure 8A again, but now consider it as a complete network of friendships in school. The chance that two randomly chosen actors in school are both Muslim is no longer 3/9 * 3/9 (=1/9). The reason is that pupils cannot nominate themselves as friends. This means that a Muslim pupil in school will not have a 3/9 chance to nominate another Muslim in school, but rather a chance of 2/8. Thus, if pupils randomly make friends in school, then we expect to see 3/9 * 2/8 = 1/12 =0.08 of all ties being between two Muslims. And the probability that two friendships are between Christian pupils is then: 6/9 * 5/8 = 0.42. We expect to see that 50% of the friendships are with in-group members and 50% with out-group members. Indeed, this is what you’d always expect to see for the complete social network when choices are made at random. Clearly, the observed 67% bonding ties are more than expected under random choice, and this gives evidence to suggest that religious homophily plays a role.

References

Kadushin, C. (2012). *Understanding Social Networks: Theories, Concepts, and Findings*. New York, NY: Oxford University Press.

Kalmijn, M. (1998). Intermarriage and Homogamy: Causes, Patterns, Trends. *Annual Review of Sociology, 24*, 395–421.