Understanding the Relationship Between College Drinking Behavior and Sexual Assault in 1997

Phuc Nguyen, Ezinne Nwankwo, Frances Hung

2/19/2020

Abstract

In this paper, we study the factors that influence college drinking behavior and its effects on other harmful behaviors, in particular sexual assault. We answer three main questions of interest: understanding student-specific factors that influence college drinking behavior, determining if current school policies and alcohol eduction have any influence on college drinking, and understanding the relationship between college drinking behavior and sexual assault while accounting for other confounding factors (i.e. family drinking behaviour). Through the use of mediation analysis and structural equation models, we are able to isolate the direct effect of other factors on sexual assault and their indirect effect through college drinking behavior. We conclude that in 1997, high school drinking behavior had the largest effects, both directly and indirectly, on college drinking behavior and on incidences of sexual assault.

Introduction

Drinking in college is akin to a rite of passage; while policies around college campus seem to vary in terms of strictness, enforcement is often less stringent than at bars and clubs. Rather than focusing on drinking, many schools care more about the dangerous situations that drinking can lead to, like sexual assault and drunk driving. For this study, we focus on understanding sexual assault, as it is a pervasive and pertinent issue on college campuses. We are interested in studying three questions in the hopes that we can help colleges take preventative measures: (i.) which factors influence college drinking behavior (ii.) how school policies influence drinking behavior and (iii.) how those behaviors can lead to sexual assault. Our data comes from the 1997 Havard School of Public Health College Alcohol Study, a survey of students from more than 100 colleges across the US which sought to collect information on drinking habits, school policy, and personal activities/beliefs (Wechsler and Nelson 2008).

Materials and Methods

We use Structural Equation Modeling (SEM) for modeling relationships between information collected via the survey and our dependent variables `drinkcat` and `sexassault_ind`. SEM is a way of implementing dimension reduction (by using latent factors) while retaining coarse information on all variables of interest. This is different from other dimension reduction strategies which get rid of some variables outright (Hox and Bechger 1998). The SEM package we use in R, `lavaan`, uses partial least squares regression to find a linear model relating the independent and dependent variables through latent factors (Rosseel 2012). We standardize all of the variables we use in order to make our coefficients interpretable and comparable in the final model.

We create the following latent factors from existing survey question variables (Fig. **??**). Since these factors must be correlated, interpretable, and continuous (due to our use of the `lavaan` package), we ensure that the survey question variables included in each factor are uniformly ordinal in an interpretable way. Our two factors related to students' pre-college lives are Family Education/Drinking (G14-17) and High School Drinking Behavior (G9-11). The external-policy related factors are College Alcohol Policy (B1, B2, B9), College Alcohol Education (B7, B8), and City Alcohol Policy (B11-13). The rest of the latent factors have to do with students' college experiences and opinions: Academic Wellbeing (F1, F4), Personal Wellbeing (F6's), Support for Stricter Policies (D3's), Opinions on the Number of Appropriate Drinks (D1's), Time Spent on Activities (F5's), and General Interests for College Experience (split into activities involving and not involving partying/Greek life, A10's). All of the latent factors which draw from one survey question actually consist of several sub-questions within each question (i.e. D3's consist of student rankings of agreement with

many different school policies). We fit an ordered logistic regression to model the relationships between these latent factors and `drinkcat`. The graphical model is represented in Fig. **??**.

In order to understand the relationships between drinking behaviors and sexual assault, we fit another latent factor model with mediation on the indicator of being a victim of sexual assault given family education/drinking background, high school drinking, `drinkcat`, and participation in parties. Mediation analysis allows us to estimate how much of the association between a risk factor, such as participation in parties, and the risk of sexual assault is mediated by the victim's drinking behavior (MacKinnon, Fairchild, and Fritz 2007). We model the relationships for two outcomes: sexual assault incidents when the victim was drinking and when the assaulter was drinking. The equations for the mediation model are written below (also see Fig. **??**):

$$\text{drinkcat} \sim \text{communities} + \text{family} + \text{highschool}$$
$$\text{sexualassault} \sim \text{communities} + \text{family} + \text{highschool} + \text{drinkcat}$$

Results

Data Cleaning

Due to the structure of the survey (students skip certain questions based on their answers), there are a lot of missing values induced in the data for certain variables. We impute zero values for NAs if the variables are ordinal and if NAs are naturally lowest in the ordinal structure. Otherwise, we create indicator variables for whether a person answered NA or not for each variable in question. We additionally drop true missing values.

Exploratory Data Analysis

For our dependent variable measuring drinking behavior, we choose to use `drinkcat`, a variable created by researchers which classifies drinking behavior into four ordinal risk categories, where `drinkcat = 2` and `drinkcat = 3` are categorized as binge drinkers. The sexual assault index variable draws from three questions in the survey related to occurences of sexual assault when alcohol is involved: D5_H (unwanted sexual advance, intoxicated perpetrator), D5_I (date rape/sexual assault, intoxicated perpetrator), E15 (intoxicated, unable to consent). If any one of these questions is marked as occuring at least once by the respondent, then the sexual assault variable is equal to one and is zero otherwise.

Main Results

Table **??** shows the estimates and 95% confidence intervals for the coefficents of latent variables in the drinking behavior model. We accept the model as its RMSEA 95% CI of [0.053, 0.054] is around the conventional threshold for a good fit (Hox and Bechger 1998). Since the survey answers generally do not have physical units, we standardize the variables to compare the size of the estimated coefficients. The dependent variable, i.e. the rating of levels of drinking, also does not have physical units. Thus, we will interpret these coefficents in terms of the relative magnitude and direction of their associations with more binge drinking. High school drinking history is most correlated with more drinking in college. The latent factor Parties summarizing the importance of fraternity/sorority, parties and athletics to students, with a lower value corresponding to more importance, has the second largest association. In other words, students who are more involved in these groups tend to drink more. The latent factor Communities measures the importance of other activities such as volunteering, religion, arts, activism and academic, has the third largest association. Students who care less about these activities tend to drink more. Fourthly, students who perceive a larger number of drinks as appropriate tend to drink more. The loadings of these four latent variables are shown in Table **??**. Other latent variables with significant but smaller associations with more drinking in college include Personal Wellbeing, Family Education/Drinking and Support for More Lenient Policy. That is, students who are generally more happy, whose families approve of drinking, or who support more lenient drinking policies on campus tend to drink more. There is not enough evidence to establish a relationship between stricter college drinking policy and drinking behaviors. We find a small association between more exposure to alcohol/drinking education and heavier drinkers. We hypothesize that students who already have drinking

problems or schools with more prevalent drinking culture might consequently have more educational programs on the issue.

We accept both of the mediation models for when the victim was drinking and when the assaulter was drinking (RMSEA 95% CI's of [0.029, 0.035] and [0.029, 0.034] respectively). More acceptance of alcohol use from family, drinking history in high school, more participation in parties/Greek life/athletics and more binge drinking are all significantly associated with higher risk of sexual assault in both models, as seen in Table **??** and Table **??**. Table **??** and **??** show that the indirect effects of family, high school drinking history and parties, mediated through a student's drinking behavior, are statistically signficant at 5% level. Interestingly, the indirect effects are larger when the victims were drinking compared to when the assaulters were drinking at the time of the incidents.

Sensitivity Analysis

We can compare our SEM model for college drinking severity to an elastic net model created using the `glmnet` package. Both of these models carry out dimension reduction using different methods, so it is interesting and reassuring to note that the results of both seem to corroborate one another. From the elastic net model, the most important predictors for heavy drinkers are A10_F(parties, -), B6_D (times they were part of drinking group which was asked to be quieter/less disruptive, +), D1_B (opinion on appropriate amount for off-campus bar drinking, +), D9_B (percentage of friends who are binge drinkers, +), and G11 (HS number of binges, +). This supports the results of our latent factor model, which finds strong correlations of college drinking with drinking attitudes, emphasis on partying, and high school drinking.

Discussion and Future Work

Using data from Harvard's survey on college drinking behavior, we model factors associated with heavy drinking behaviors and the risk of sexual assault. We find that high school drinking history is most correlated with more drinking in college. Importance of parties/Greek-life/athletics, little involvement in other communal activities, and permissive attitude on the amount of alcohol generally appropriate to consume are also highly correlated with heavier drinking. These results align with current knowledge of patterns of alcohol use. As for our seecond question on the relationship between college alcohol policies and drinking behaviors, we found no evidence of a relationship between strictness of college or city drinking policies and drinking behaviors. We also saw that schools that had less educational material on the effects of drinking were associated with less drinking among students. This result is counterintuitive but we believe that this may be due to the fact that colleges with more educational material are more likely to have high levels of drinking already on their campuses. Universities are creating the educational material to combat drinking problems that already exist on their campuses. Finally, when in came to studying sexual assault activity, we found that high school drinking habits alone and through mediator contributes the most to incidences of sexual assault. Other risk factors for sexual assault are also mediated by the victim's drinking habits. The mediated effect is larger if the victim was drinking when the incident occurred. A pitfall of our models is that we assume linearity and normality. In the future, we would like to extend to non-parametric models with SEM.

References

Hox, JJ, and TM Bechger. 1998. "An Introduction to Structural Equation Modeling." *Utrecht University Repository.*

MacKinnon, David P., Amanda J. Fairchild, and Matthew S. Fritz. 2007. "Mediation Analysis." *Annual Review of Psychology* 58: 593–614.

Rosseel, Yves. 2012. "Lavaan: An R Package for Structural Equation Modeling." *Journal of Statistical Software* 48: 1–36.

Wechsler, Henry, and Toben Nelson. 2008. "What We Have Learned from the Havard School of Public Health College Alcohol Study." *Journal of Studies on Alcohol and Drugs* 69: 481–90.

Appendix

Phuc Nguyen, Ezinne Nwankwo, Frances Hung

2/19/2020

Appendix

Elastic Net Model

The most important predictors for heavy drinkers: A10_F(parties, neg), B6_D (drinking group which was asked to be quieter/less disruptive, positive), D1_B (opinion on appropriate amount for off-campus bar drinking, pos), D9_B (percentage of friends who are binge drinkers, pos), G11 (HS number of binges, pos)
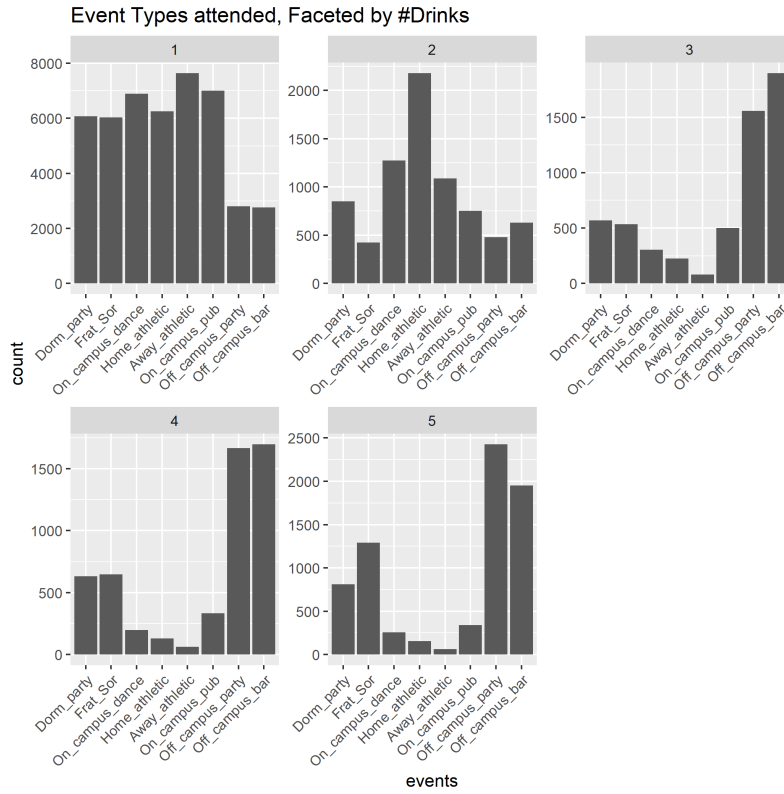

Figure 1: The figure shows the number of students attending certain events, faceted by number of drinks had at an event.
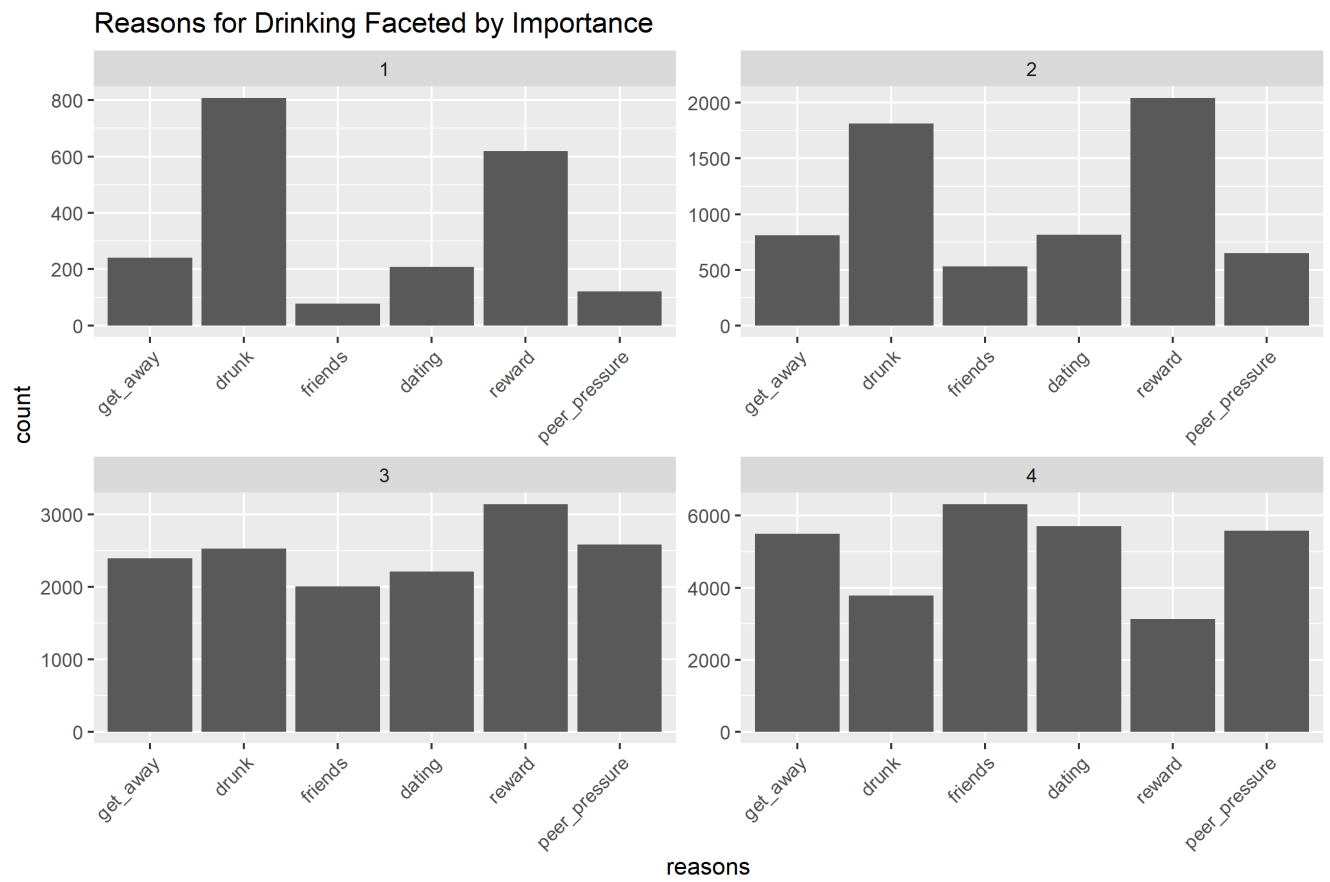
Figure 2: The figure shows the reasons why students say they drink, faceted by importance (very important=1).
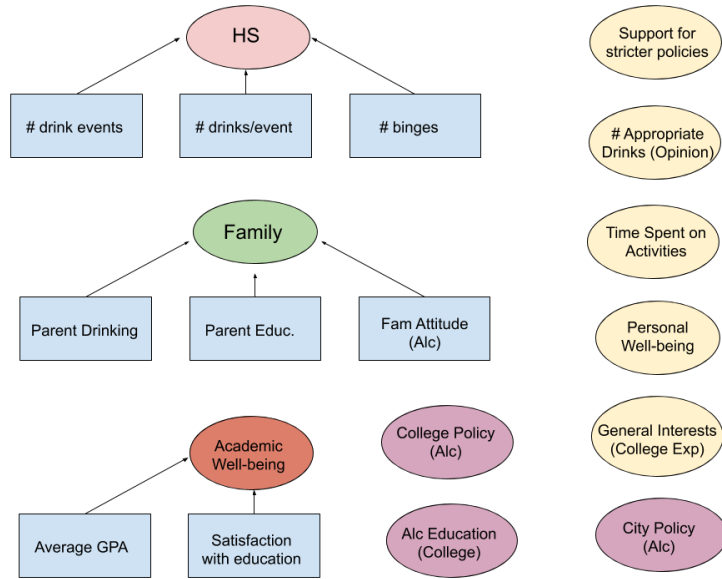
Figure 3: Figure shows latent factors (ovals) in the SEM for drinking behaviors. Squares show summary of actual questions in the survey.

Below are tables of the estimates for the following three models

$$\text{SEXASSAULT} \sim \text{communities} + \text{family\_drink} + \text{highschool\_drink}$$
$$\text{DRINKCAT} \sim \text{communities} + \text{family\_drink} + \text{highschool\_drink}$$
$$\text{SEXASSAULT} \sim \text{DRINKCAT}$$

Table 1: Estimates and 95 percent confident intervals of latent factors in the SEM for levels of drinking. High school drinking, support of more lenient drinking policy, importance of parties (parties), less exposure to alcohol education, family drinks, and personal wellbeing are associated with more drinking in college

latent_var	est	ci_lower	ci_upper	p_value
attitude_policy	0.10	0.05	0.14	0.00
communities	0.37	0.18	0.56	0.00
parties	-1.10	-1.28	-0.92	0.00
attitude_drink	0.37	0.34	0.41	0.00
college_policy	0.03	-0.03	0.09	0.29
college_edu	-0.09	-0.12	-0.05	0.00
city_policy	-0.17	-0.37	0.02	0.08
family_drink	0.09	0.04	0.13	0.00
highschool_drink	2.21	1.67	2.75	0.00
personal_wellbeing	-0.05	-0.09	0.00	0.04
academic_wellbeing	-0.04	-0.13	0.06	0.46
social_groups	0.19	-0.14	0.53	0.26

Table 2: Loadings of survey questions for latent factors.

latent_var	survey_question	est	ci_lower	ci_upper	p_value
parties	A10_A	1.00	1.00	1.00	NA
parties	A10_D	1.15	0.99	1.32	0
parties	A10_F	3.46	3.06	3.86	0
communities	A10_B	1.00	1.00	1.00	NA
communities	A10_C	1.29	1.08	1.50	0
communities	A10_E	1.16	0.98	1.34	0
communities	A10_G	2.50	2.16	2.84	0
communities	A10_H	2.86	2.43	3.29	0
attitude_drink	D1_A	1.00	1.00	1.00	NA
attitude_drink	D1_B	0.93	0.91	0.95	0
attitude_drink	D1_C	0.58	0.55	0.60	0
attitude_drink	D1_D	0.75	0.72	0.77	0
attitude_drink	D1_E	0.30	0.27	0.32	0
highschool_drink	G8	1.00	1.00	1.00	NA
highschool_drink	G9	7.21	5.43	8.98	0
highschool_drink	G10	8.63	6.50	10.77	0

Table 3: Direct and indirect effects of latent factors on risk of sexual assault when victim was drinking.

latent_var	effect_type	est	ci_lower	ci_upper	p_value
communities	indir	-0.27	-0.43	-0.11	0.00
family	indir	0.04	0.02	0.07	0.00
highschool	indir	0.88	0.30	1.47	0.00
communities	dir	-0.48	-0.74	-0.21	0.00
family	dir	0.00	-0.15	0.14	0.99
highschool	dir	0.66	-0.39	1.71	0.22

Table 4: Direct and indirect effects of latent factors on risk of sexual assault when assaulter was drinking.

latent_var	effect_type	est	ci_lower	ci_upper	p_value
communities	indir	-0.18	-0.27	-0.09	0.00
family	indir	0.03	0.01	0.04	0.00
highschool	indir	0.59	0.24	0.94	0.00
communities	dir	-0.83	-0.98	-0.67	0.00
family	dir	0.07	-0.01	0.16	0.09
highschool	dir	1.08	0.47	1.70	0.00

Table 5: Coefficient estimates of the three models. The response is indicator of sexual assault when the assaulter was drinking.

response	latent_var	est	ci_lower	ci_upper	p_value
SEXASSAULT	communities	-0.83	-0.98	-0.67	0.00
SEXASSAULT	family_drink	0.07	-0.01	0.16	0.09
SEXASSAULT	highschool_drink	1.08	0.47	1.70	0.00
DRINKCAT	communities	-1.51	-1.63	-1.39	0.00
DRINKCAT	family_drink	0.25	0.20	0.29	0.00
DRINKCAT	highschool_drink	4.96	3.55	6.36	0.00
SEXASSAULT	DRINKCAT	0.12	0.06	0.18	0.00

Table 6: Coefficient estimates of the three models. The response is indicator of sexual assault when the victim was drinking.

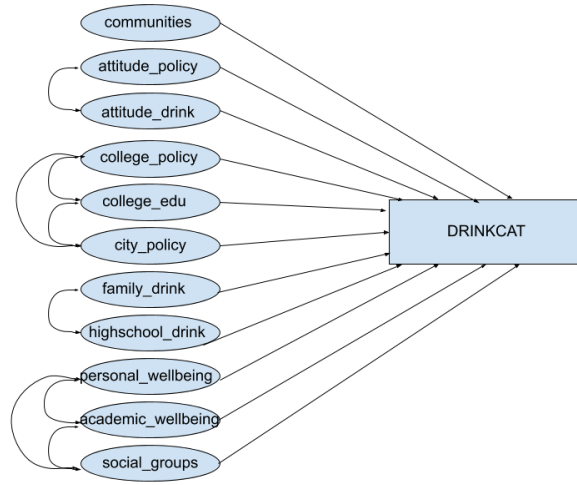| response	latent_var	est	ci_lower	ci_upper	p_value
SEXASSAULT	communities	-0.48	-0.74	-0.21	0.00
SEXASSAULT	family_drink	0.00	-0.15	0.14	0.99
SEXASSAULT	highschool_drink	0.66	-0.39	1.71	0.22
DRINKCAT	communities	-1.51	-1.63	-1.39	0.00
DRINKCAT	family_drink	0.25	0.20	0.29	0.00
DRINKCAT	highschool_drink	4.91	3.54	6.28	0.00
SEXASSAULT	DRINKCAT	0.18	0.07	0.29	0.00

Figure 4: Figure shows path diagram of SEM on risk factors for heavy drinking behaviors in college.
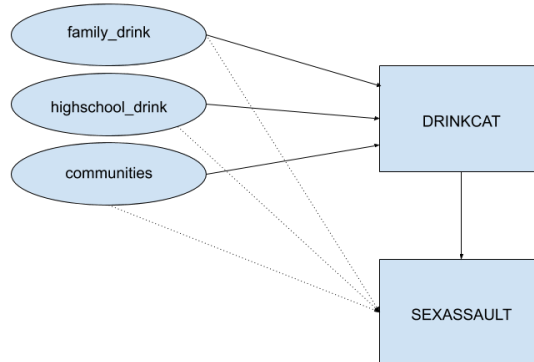

Figure 5: Figure shows path diagram of mediation model for the indirect effects of family, high school drinking behaviors and parties on risk of sexual assault.
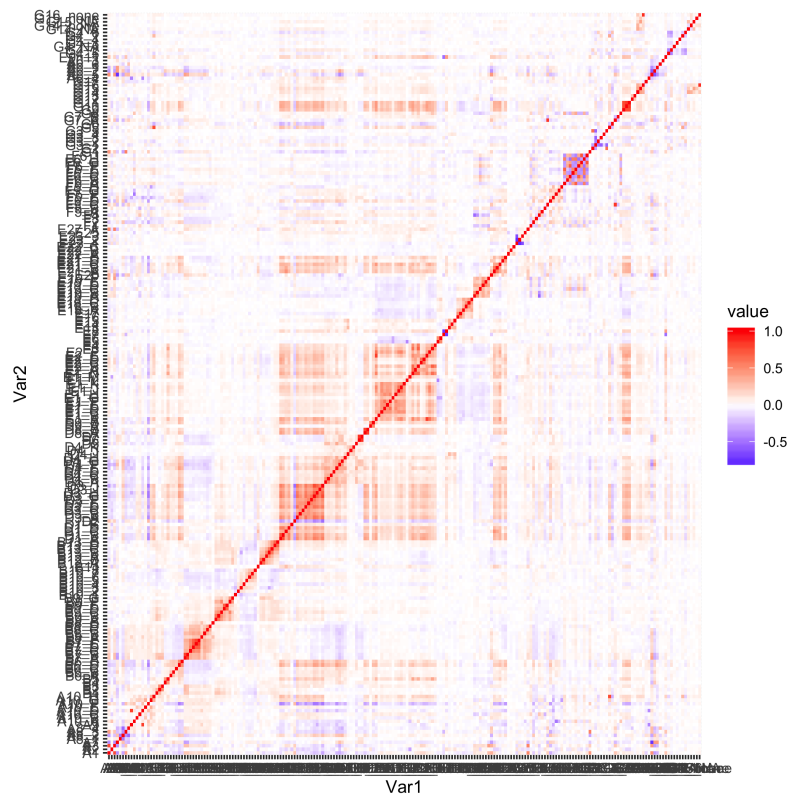
Figure 6: Correlations between Survey Questions.