Does High School Alcohol Abuse increase the chance of Hard Drug Usage?

Evan Knox, Joe Mathews, Alessandro Zito

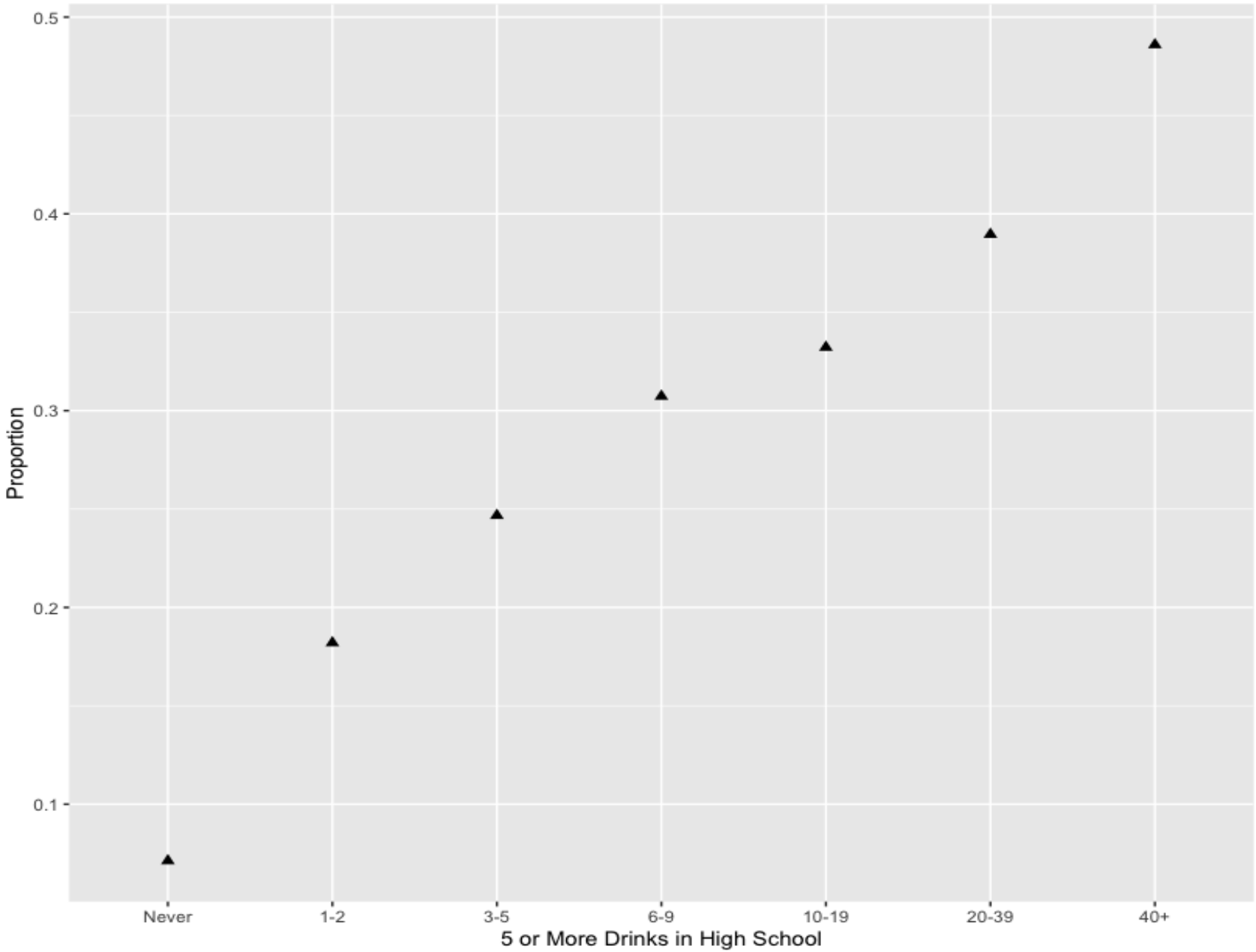
Duke Department of Statistical Science

February 18, 2020

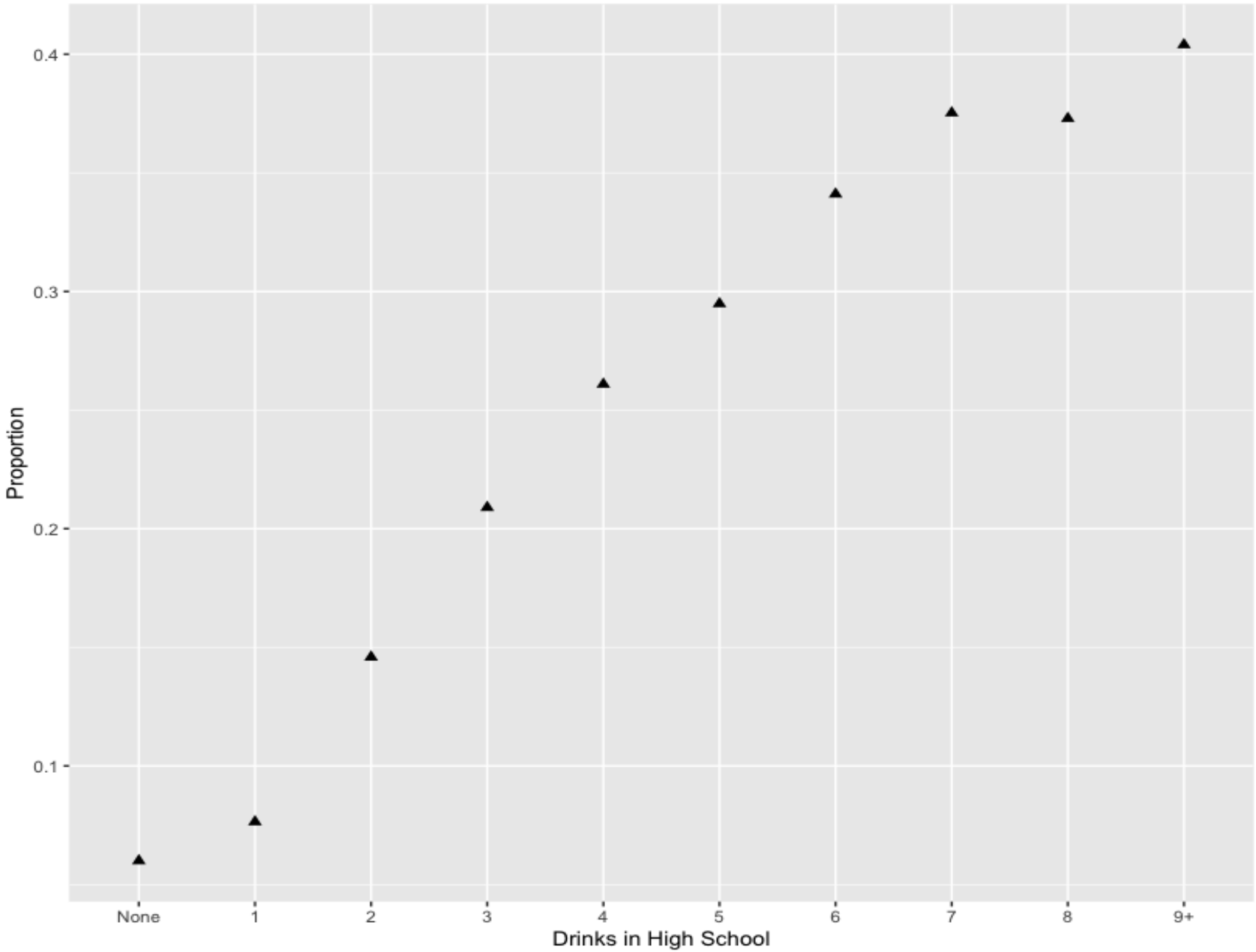
Main Question and Data Cleaning

- How does high school drinking affect college students' use of hard drugs? How does this effect relate to college drinking? How can we best screen for students who might have used hard drugs?
- Defined *drug use* variable as having ever tried a "hard drug" (excludes tobacco related products, marijuana, and performance enhancing drugs).
- Hand selected questions directly relevant to high school and college alcohol abuse, as well as other control variables
- Examples of data cleaning approaches:
  - Questions differed in answer choices across years (including the ordering of the responses).
  - Some responses to nested questions were contradictory, e.g. having a drink within a two week period but also not having a drink in the last thirty days.
  - Important variables were not included in all years. For example,

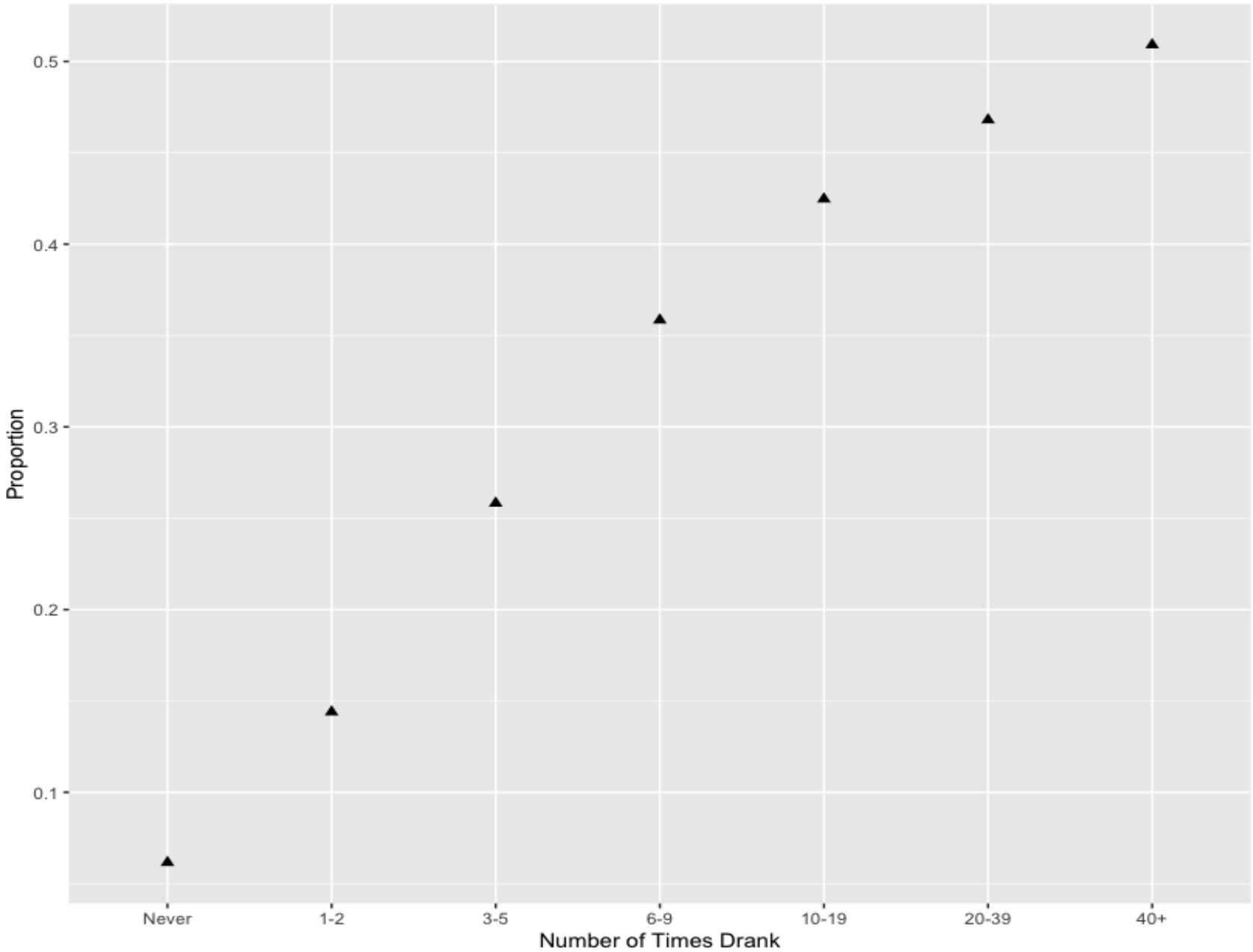
EDA


0.5
0.4
0.3
0.2
0.1
Proportion
Never
1-2
3-5
6-9
10-19
20-39
40+
5 or More Drinks in High School

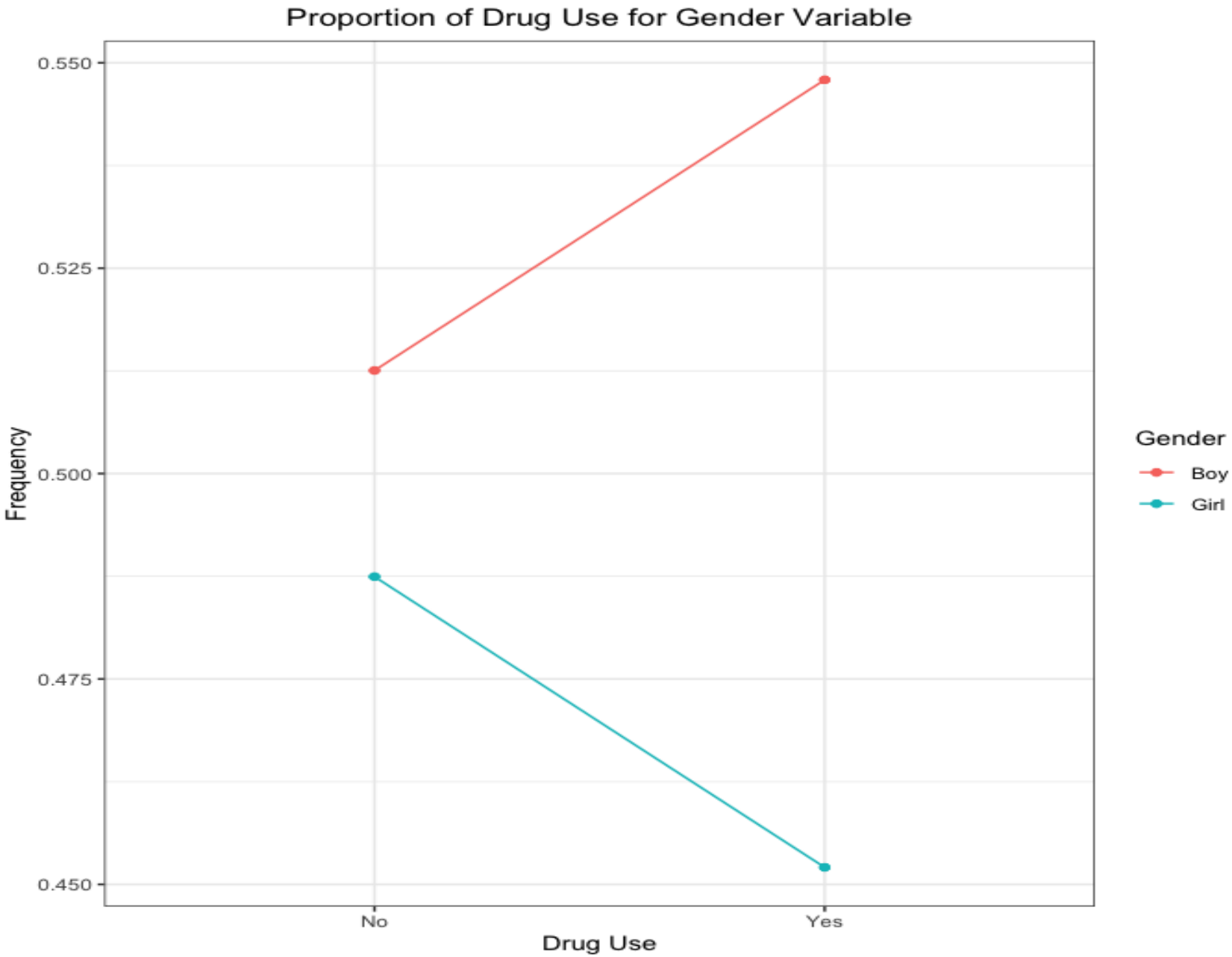
EDA



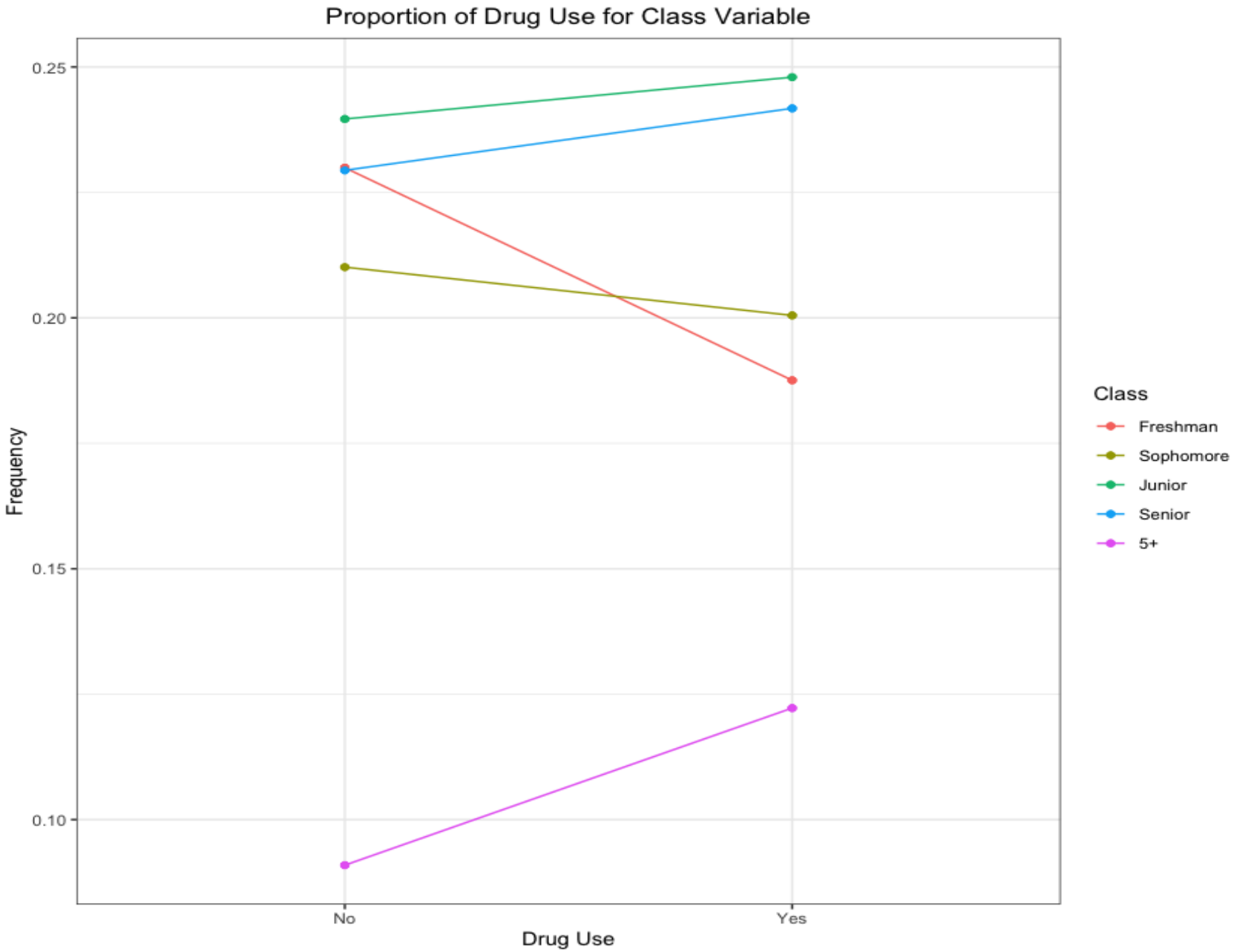
EDA


0.5
0.4
0.3
0.2
0.1
Proportion
Never
1-2
3-5
6-9
10-19
20-39
40+
Number of Times Drank

EDA


Proportion of Drug Use for Gender Variable
0.550
0.525
0.500
0.475
0.450
Frequency
No
Yes
Drug Use
Gender
Boy
Girl

EDA


Proportion of Drug Use for Class Variable
0.25
0.20
0.15
0.10
Frequency
No
Yes
Drug Use
Class
Freshman
Sophomore
Junior
Senior
5+

EDA Conclusion

- High school drinking variables appear colinear.
- Multi-colinearity issues for college and high school drinking behavior responses.
  - A more sophisticated model will be needed to account for this.
- Exploratory logistic model detects linear (possibly quadratic) patterns in ordered responses to high school drinking questions and drug use.
- Control variables (sex, class, etc.) should play an important role in our model.

The model

Our causal link of interest is

1. Binge drinking in **High School** $\Rightarrow$ Binge drinking in **College**
2. Binge drinking in **College** $\Rightarrow$ Use of **Heavy Drugs**

In other words, we want to check if, given a certain level of high school drinking, does college drinking act as a mediator in the choice of assuming drugs?

Possible link: Having earlier experiences with alcohol increases the desire to try heavier substances in college, and these drugs can be accessed in similar environments where college students consume alcohol. We

construct our aggregated measures for high school drinking and college drinking in two ways

1. **Crude aggregation**
2. **Confirmatory factor analysis** (CFA)

Model (I) - Crude aggregation

Two stage model

$$\mathbb{P}(Drug_i = 1) = \text{logit}(\alpha_1 + \beta_1^{HS} Binge_i^{HS} + \beta^C Binge_i^C + \mathbf{X}\gamma_1) \quad (1)$$

$$\mathbb{P}(Binge_i^C = 1) = \text{logit}(\alpha_2 + \beta_2^{HS} Binge_i^{HS} + \mathbf{X}\gamma_2) \quad (2)$$

where

- $Binge_i^{HS}$ equal to 1 if the student in the last year of HS has had either binged 5 or more drinks or has been drunk for more than 10 times, else 0
- $Binge_i^C$ equal to 1 if the student in the last two weeks has binged more than 5 drinks 4 times, or if in the last 30 days he had more than 10 drinks or he got drunk more than 10 times, else 0
- $\mathbf{X}$ contains the controls, such as age, gender, family feelings towards alcohol (as a polynomial contrast), race, education satisfaction, GPA, parents schooling, and "Drinking problem" variables.

Model (II) - Confirmatory factor analysis

Two stage model

$$\mathbb{P}(Drug_i = 1) = \text{logit}(\alpha_1 + \beta_1^{HS}\eta_i^{HS} + \beta^{C}\hat{\eta}_i^{C} + \mathbf{X}\boldsymbol{\gamma}_1) \quad (3)$$

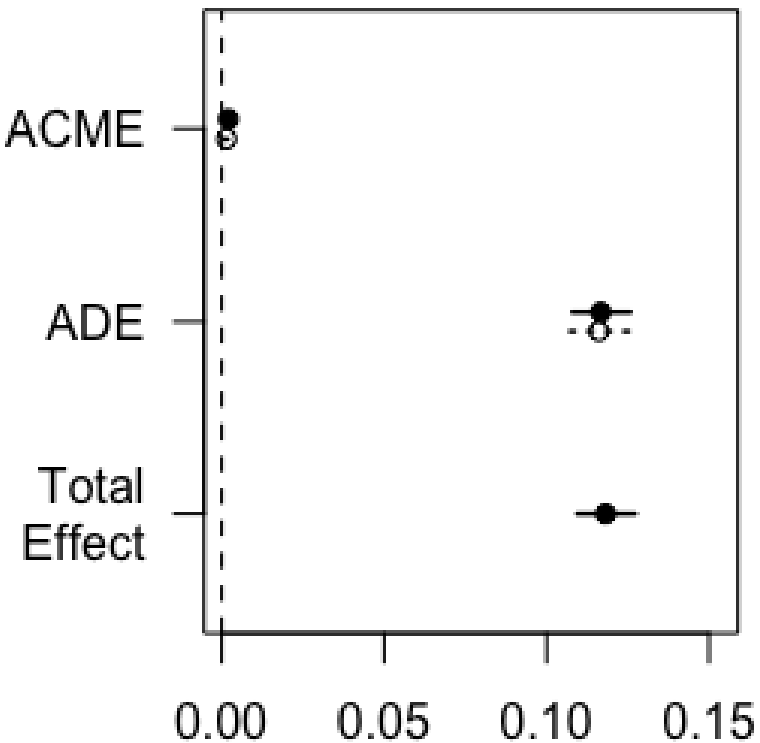
$$\eta_i^{C} = \alpha_2 + \beta_2^{HS}\eta_i^{HS} + \mathbf{X}\boldsymbol{\gamma}_2 + \epsilon_i \quad (4)$$

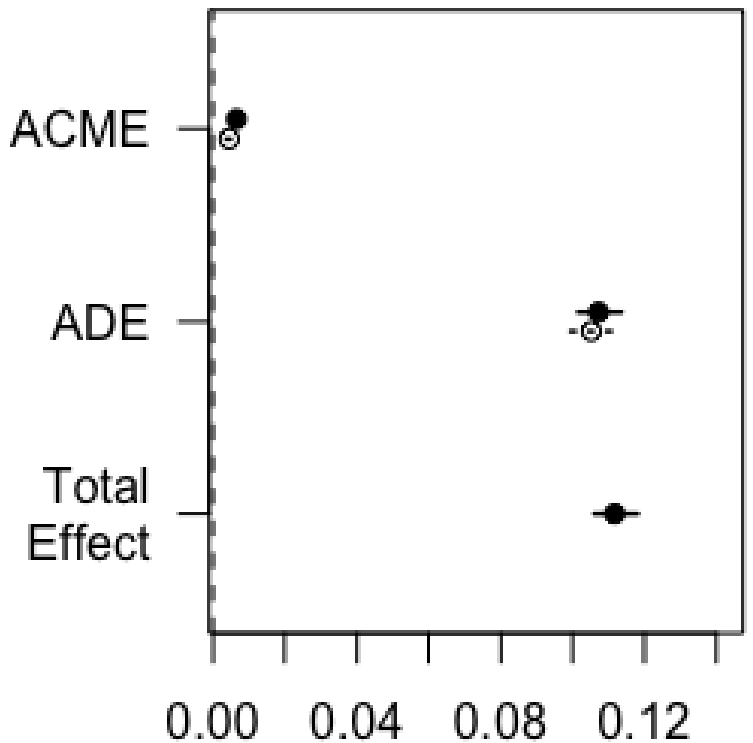
where

- η_i^{HS} is the latent factor behind drinking behaviour in high school (Binge5 times, how many times drunk and average drinks on occasion in last year of HS)
- η_i^{C} is the latent factor behind drink behaviour in college (alcohol related behavior in last 30 days, binge behavior in last two weeks.)
- $\mathbf{X}$ contains the same aggregated controls as before.

Each latent factor in computed separately with the `lavaan` package in `R`.

Results


ACME
ADE
Total
Effect
0.00
0.05
0.10
0.15


ACME
ADE
Total
Effect
0.00
0.04
0.08
0.12

Classification Trees

- Common indicators of previous hard drug usage: drinking in high school (6+ times per month in last year), binge drinking in high school (6+ times in last year)
- Final indicator could be number of times drunk in last 30 days (10+), or else whether drinking has ever caused them to get behind in schoolwork. In case of 10+ times drunk in last 30 days, binge drinking in high school need only happen at least 1.
- Criteria hold (with higher HS binge threshold) for men; among women, HS binge threshold is slightly lower, and drug use seems more common among extreme ages (17- and 25+).
- AA meeting attendance/alcohol counseling is indicator of hard drug use in older/married at some point students
- Non-Greek members have similar criteria as genpop; Greek members predicted to have used hard drugs are (confusingly) former high school binge drinkers who binge moderately often now, haven't fallen behind on schoolwork, and last drank over a year ago.

Recent drinking-based tree

Figure: Tree containing variables about drinking behavior over last 30 days

More time-insensitive tree?

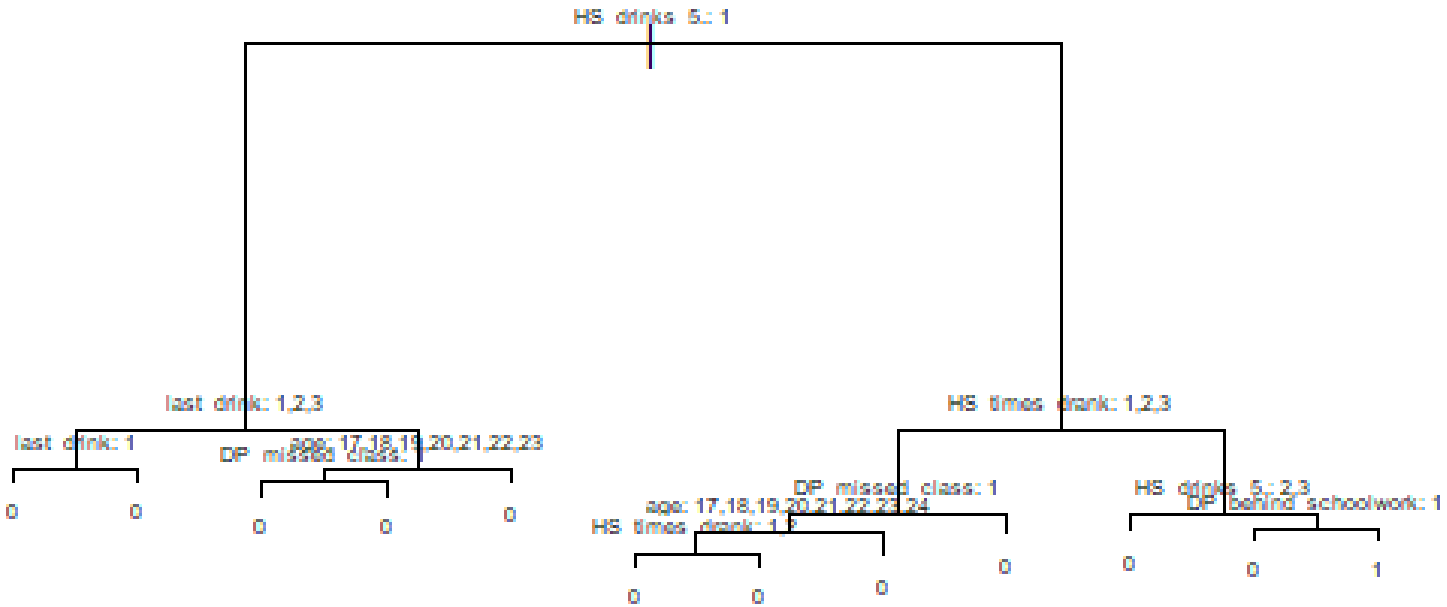


Figure: This tree is less based on recent drinking than the previous one, and can involve screening through academic channels as well.

Conclusions and Future Directions

- HS drinking, and binge drinking in particular, are influential predictors of hard drug usage
- College drinking doesn't appear to mediate much of HS drinking effect on hard drug usage
- Classification trees aren't good for rare outcomes; adjustment to regression tree for proportions or adoption of different screening method preferable