# Sample Report for

# Project TIER Reproducibility Exercise:

# Animal House in Alcohol-Free Dorms?

Sample Solution

December 4, 2018

## I. Interpretation

**1)** The data were taken from Wechsler's (2001) survey of students enrolled in 119 colleges and universities in the US.

In the original dataset for the study (*0429-0001-Data.dta*), and in the processed dataset we used to generate the figures for this exercise (*analysis.dta*), the unit of observation is a student. In other words, each record (row) in the dataset represents one student who completed the survey.

The original dataset included records for 10,094 students. To construct the analysis dataset we used to generate the figures, we modified the sample in several ways, in the following order:

*i*) We deleted all students for whom the campus residence variable (*A6*) was equal to "Off campus house/apt" or "other," or for whom the value of this variable was missing. That reduced the sample by 6,379 observations.

*ii*) We deleted all students for whom the variable that reported the number of times the student had been drunk in the past thirty days (*C13*) had a missing value. That reduced the sample by an additional 1,569 observations.

*iii*) We deleted all students for whom the variable that reported the number of times the student had consumed five or more drinks in a row during senior year of high school (*G11*) had a missing value. That reduced the sample by an additional 110 observations.

*iv*) We then deleted 15 additional students who reported they did not live in alcohol-free housing (*B8*=0), but contradicted that by reporting they lived in alcohol-free housing because they requested it (*B9*=1), because they had been assigned to it (*B9*=2), or because all campus housing was alcohol-free (*B9*=3).

*v*) We then deleted 114 additional students who reported they lived in alcohol-free housing (*B8*=1), but the reason was ***not*** that they requested it (*B9*=1), nor that they had been assigned to it (B9=2), nor that all campus housing was alcohol-free

(*B9*=3). Equivalently, we deleted 114 students for whom B8 was equal to 1 and *B9* was missing or equal to 4, 5, 6 or 7.

***NOTE***

You should easily be able to see how, by following the instructions for the assignment, you would delete the observations described in items *i*, *ii* and *iii*.

For items *iv* and *v*, it may be harder to see how following the instructions in the assignment would lead you to delete the observations that are described. But, in the *processing.do* file included in this sample exercise solution, the command

```
drop if housing_category==.
```

deletes the 15 observations described in item *iv* as well as the 114 observations described in item *v*. It would be a good exercise to see if you can figure out why that command deletes precisely those observations. (It of course has to do with precisely how *housing_category* was defined.)

After all those deletions are made, the processed data file *analysis.dta* consists of 2,717 observations.

**2)** Let's use the term "heavy drinkers" to refer to students who report having gotten drunk at least three times in the last thirty days. Then figure 1 shows that the proportion of heavy drinkers is very similar among students living in alcohol-free housing and those not living in alcohol-free housing—for both groups, the proportion is just over 30 percent. Although we have not done any test of statistical significance, we can tell by looking at the bar graph that the difference between the proportions for the two groups is small.

It would have seemed reasonable to expect that the proportion of heavy drinkers would be lower among students living in alcohol-free housing, but in figure 1 we find that is not the case.

**3)** Figure 2 shows that the proportion of heavy drinkers was lower among students who lived in alcohol-free housing because they requested it than among students who lived in alcohol-free housing because they were assigned to it. Compared to the overall rate

of just over 30 percent observed in figure 1 for all students in alcohol-free housing, students who requested alcohol-free housing had a lower proportion of heavy drinkers (around 16 percent) and students assigned to alcohol-free housing had a larger proportion of heavy drinkers (around 38 percent).

The implication is that although (as seen in figure 1) the rates of heavy drinking are about the same in alcohol-free housing and non-alcohol-free housing, the rate is lower among students who chose to live in alcohol-free housing than among those who were assigned to live there.

**4)** All 2,717 students in the *analysis.dta* dataset are represented in figure 1, and 289 students are represented in figure 2.

The reason for this difference is that figure 2 used data only for students in alcohol-free housing. Consequently, the 1,909 individuals in the dataset who did not live in alcohol-free housing were excluded.

In addition, figure 2 used data only for individuals living in alcohol-free housing because they requested it or because they were assigned to it. Consequently, the 519 students who lived in alcohol-free housing because all campus housing was alcohol-free were excluded.

**5)** The left-most bar (labeled “Not Alcohol-Free”) in figure 3 is identical to the left-most bar (labeled “Not in Alcohol-Free Housing”) in Figure 1: both show that just over 30 percent of students not living in alcohol-free housing were heavy drinkers.

The bar (labeled “Assigned to alcohol-free housing”) that is second to the left in figure 3 is identical to the left-most bar (labeled “Not Voluntarily”) in Figure 2: both show that about 38 percent of students not living in alcohol-free housing because they were assigned to it were heavy drinkers.

The right-most bar (labeled “Requested alcohol-free housing”) in figure 3 is identical to the right-most bar (labeled “Voluntarily”) in Figure 2: both show about 16 percent of students not living in alcohol-free housing were heavy drinkers.

The only bar in figure 3 that is not represented in figure 1 or 2 is the one labeled “All campus housing alcohol-free.” It shows that the proportion of heavy drinkers in this group was near to 30 percent.

An interesting thing to note in figure 3 is that the proportion of heavy drinkers among students assigned to live in alcohol-free housing is greater than the proportion of heavy drinkers among students who do not live in alcohol-free housing. That suggests that forcing students to live in alcohol-free housing leads them to drink more than if they were just allowed to live in housing that was not alcohol-free—maybe there is something here related to some form of rebelliousness. That sounds like a reasonable hypothesis, but we don’t really know whether this causal story is correct. There may be something else going on; for instance, maybe a lot of the students who were assigned to live in alcohol-free housing had previously lived in non-alcohol-free housing and gotten into trouble for excessive drinking, and they were assigned to alcohol-free housing as a disciplinary measure. In that case, if the discipline were ineffective, we might end up with a lot of heavy drinkers assigned to live in alcohol-free housing.

**6)** Let’s use the term “high school drinkers” to refer to students who report having had five or more drinks in a row at least three times during senior year of high school.

Like figure 1, figure 4 shows the proportions of heavy drinkers in alcohol-free and non-alcohol-free housing, but figure 4 further disaggregates the data into high school drinkers and high school non-drinkers. We observe that, whether we consider high school drinkers or high-school non-drinkers, the proportion of heavy drinkers is about the same for students in alcohol-free housing and students not in alcohol-free housing. But these proportions differ for high school drinkers and high school non-drinkers: among high school non-drinkers, the proportions of drinkers (for residents of alcohol-free and non-alcohol-free housing) were a little under 20 percent; among high school drinkers, the proportions of drinkers (for residents of alcohol-free and non-alcohol-free housing) were about 53 percent.

**7)** Like figure 2, figure 5 shows the proportions of heavy drinkers among students living in alcohol-free housing by request and those assigned to alcohol-free housing, but figure 5 further disaggregates the data into high school drinkers and high school non-

drinkers. We observe that, whether we consider high school drinkers or high-school non-drinkers, the proportion of heavy drinkers is lower for students who requested to live in alcohol-free housing than for students assigned to alcohol-free housing. But these proportions differ for high school drinkers and high school non-drinkers: among high school non-drinkers, just over 20 percent of students assigned to alcohol-free housing were heavy drinkers and just over 15 percent of students who requested alcohol-free housing were heavy drinkers; among high school drinkers, almost 60 percent of students assigned to alcohol-free housing were heavy drinkers and about 25 percent of students who requested alcohol-free housing were heavy drinkers

**8)** Like figure 3, figure 6 shows the proportions of heavy drinkers among students living in each of the four types of housing defined by *housing_category.* The pattern we see when we compare figures 3 and six is analogous to the pattern we saw when we compared figures 2 and 5: whether we consider high school drinkers or high school non-drinkers, the proportions of heavy drinkers in the four categories ***relative to one another*** are similar to the relative proportions among the categories in figure 3; but for every housing category, the proportion of heavy drinkers is larger among high school drinkers than among high school non-drinkers.

**9)** In general, our sample does not provide any evidence of an association between whether or not students live in alcohol-free housing and the proportion of them who are heavy drinkers. Whether you look at our sample as a whole or disaggregate it by how much students drank in high school or by the reason for which they live in alcohol-free housing, the proportion of heavy drinkers is very similar for students who live in alcohol-free housing and those who do not live in alcohol-free housing.

On the other hand, there is a strong association in our sample between high school and college drinking habits: in all of the ways we disaggregated the data, the proportion of heavy drinkers was greater among high school drinkers than among high school non-drinkers.

## Reference

Wechsler, Henry. Harvard School of Public Health College Alcohol Study, 2001. ICPSR04291-v2. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2008-02-05. http://doi.org/10.3886/ICPSR04291.v2

## II. Data Appendix

The data for this exercise were taken from a file called *04291—001-Data.dta*, which is part of ICPSR Study Number 4291. We modified that original data file to create the analysis data file *analysis.dta* that was used to generate the figures for this exercise.

Each observation in *analysis.dta* represents one student who responded to the alcohol survey. There are five variables in this data file: *drunk*, *hsdrunk*, *free*, *volfree*, and *housing*. Details on the definitions, coding and distributions of these five variables are provided below.

**Variable name:** *drunk*

**Definition:** Dummy variable indicating whether respondent reports having been drunk three or more times in the last thirty days.

**Possible values:** 0, 1.

**Coding:** 0: Respondent reports being drunk fewer than three times in the last thirty days.

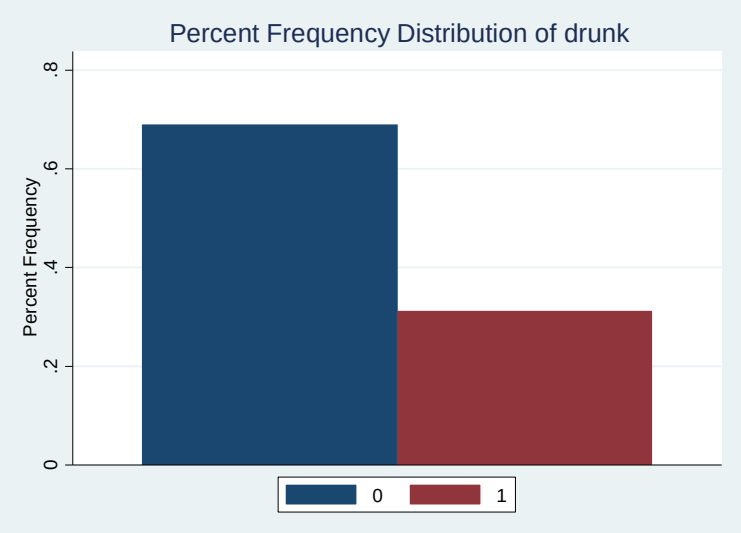
1: Respondent reports being drunk three or more times in the last thirty days.

**Missing observations:** 0/2717

**Frequency table:**

| *drunk* | Freq. | Percent |
|---|---|---|
| 0 | 1,871 | 68.86 |
| 1 | 846 | 31.14 |
| Total | 2,717 | 100.00 |

**Bar chart showing percent frequency distribution:**



**Variable name:** *hsdrunk*

**Definition:** Dummy variable indicating whether respondent reports having had five or more drinks in a row on three or more occasions during senior year of high school.

**Possible values:** 0, 1.

**Coding:** 0: Respondent reports having had five or more drinks in a row on fewer than three occasions during senior year of high school.

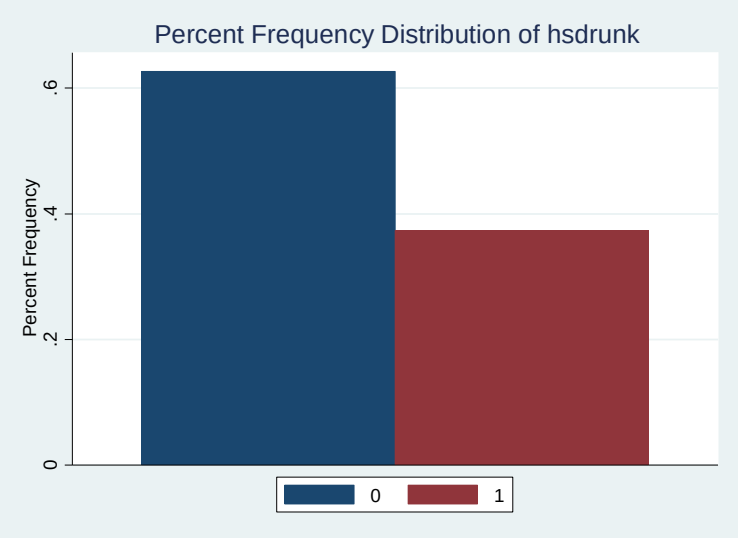
1: Respondent reports having had five or more drinks in a row on three or more occasions during senior year of high school.

**Missing observations:** 0/2717

**Frequency table:**

| *hsdrunk* | Freq. | Percent |
|---|---|---|
| 0 | 1,703 | 62.68 |
| 1 | 1,014 | 37.32 |
| Total | 2,717 | 100.00 |

**Bar chart showing percent frequency distribution:**



**Variable name:** *free*

**Definition:** Dummy variable indicating whether respondent reports living in alcohol-free housing.

**Possible values:** 0, 1.

**Coding:** 0: Respondent reports not living in alcohol-free housing.

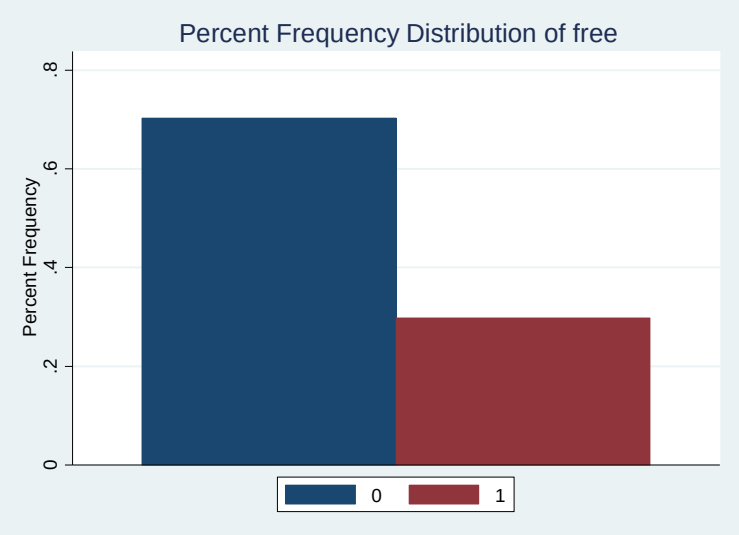
1: Respondent reports living in alcohol-free housing.

**Missing observations:** 0/2717

**Frequency table:**

| *free* | Freq. | Percent |
|---|---|---|
| 0 | 1,909 | 70.26 |
| 1 | 808 | 29.74 |
| Total | 2,717 | 100.00 |

**Bar chart showing percent frequency distribution:**



**Variable name:** *volfree*

**Definition:** Dummy variable defined only for individuals living in alcohol-free housing (*free*=1), indicating whether respondent reports requesting to live in alcohol-free housing or being assigned to live in alcohol-free housing.

**Possible values:** 0, 1.

**Coding:** 0: Respondent reports being assigned to live in alcohol-free housing.

1: Respondent reports requesting to live in alcohol-free housing.

**Missing observations:**

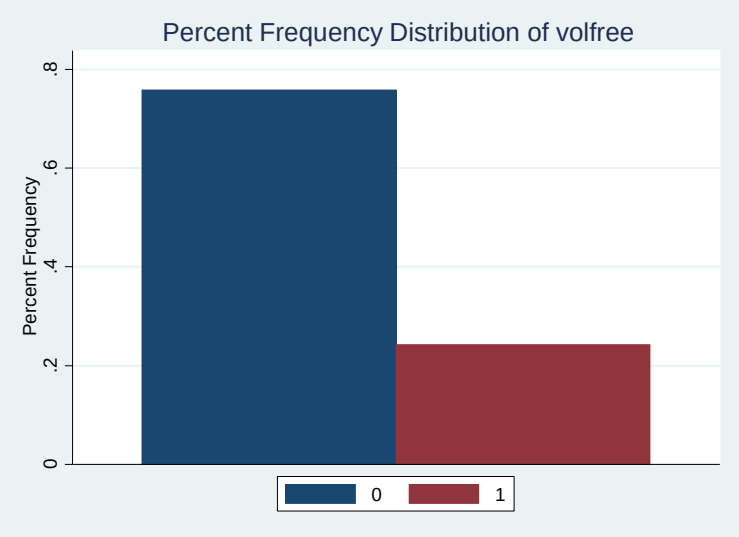
This variable is defined for the 289 observations for which *free*=1. There are no missing values of *volfree* among these observations.

For the 2428 observations for which *free*=0, this variable is not defined.

**Frequency table:**

| *volfree* | Freq. | Percent |
|---|---|---|
| 0 | 219 | 75.78 |
| 1 | 70 | 24.22 |
| Total | 289 | 100.00 |

**Bar chart showing percent frequency distribution:**



**Variable name:** *housing*

**Definition:** Categorical variable indicating type of housing student reports living in.

**Possible values:** 1, 2, 3, 4

**Coding:** 1: Respondent reports not living in alcohol-free housing.

2: Respondent reports being assigned to live in alcohol-free housing.

3: Respondent reports all campus housing is alcohol-free.

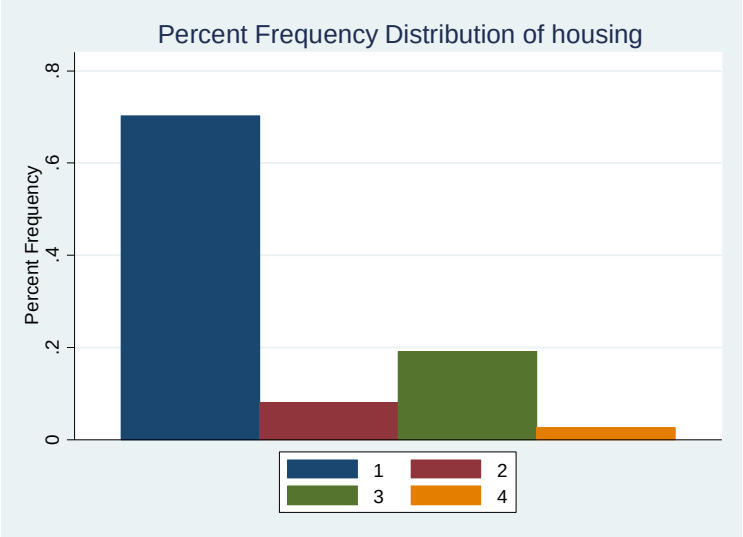
4: Respondent reports requesting to live in alcohol-free housing.

**Missing observations:** 0/2717

**Frequency table:**

| *housing* | Freq. | Percent |
|---|---|---|
| 1 | 1,909 | 70.26 |
| 2 | 219 | 8.06 |
| 3 | 519 | 19.10 |
| 4 | 70 | 2.58 |
| Total | 2,717 | 100.00 |

**Bar chart showing percent frequency distribution:**



## III. Figures

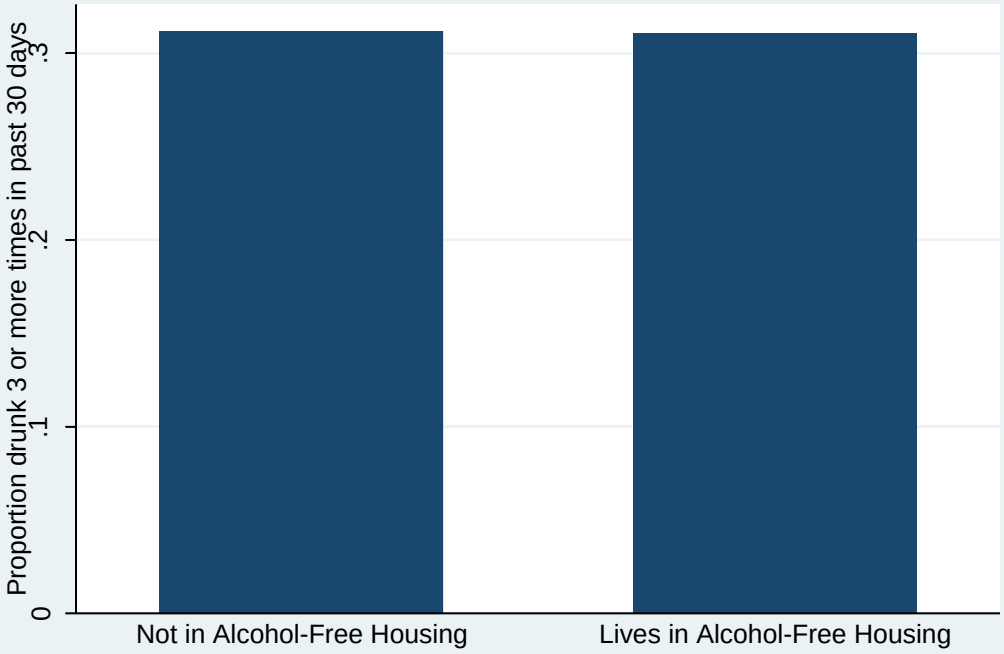



Figure 1: Rates of Heavy Drinking in Alcohol-Free Versus Not Alcohol-Free Housing

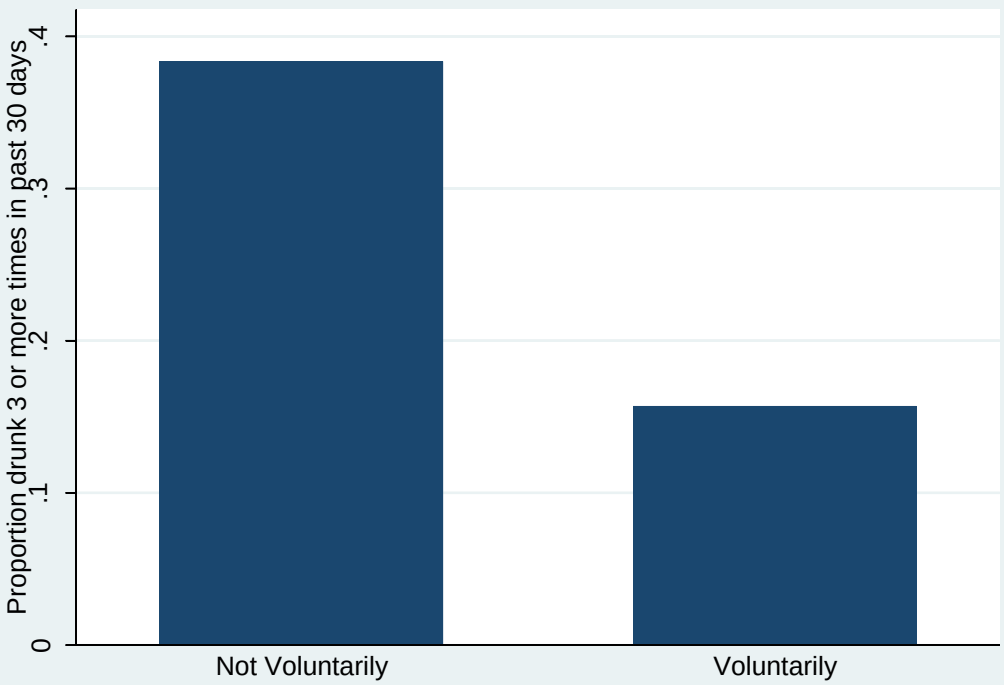



Figure 2: Students in Alcohol-Free Housing By Choice versus Assignment

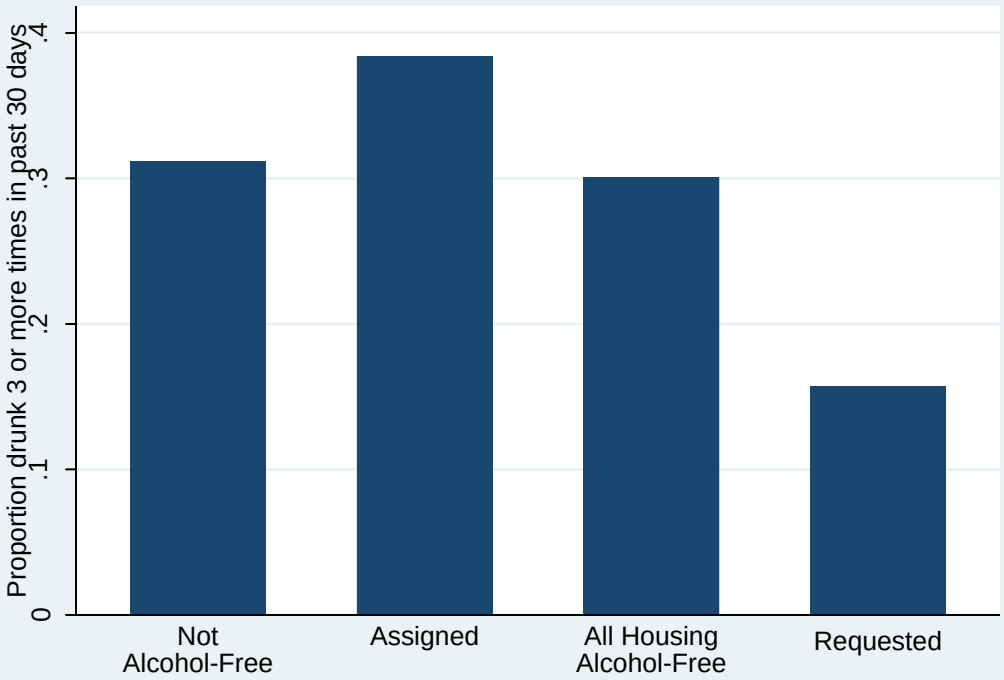



Figure 3: Drinking by Students in All Housing Types

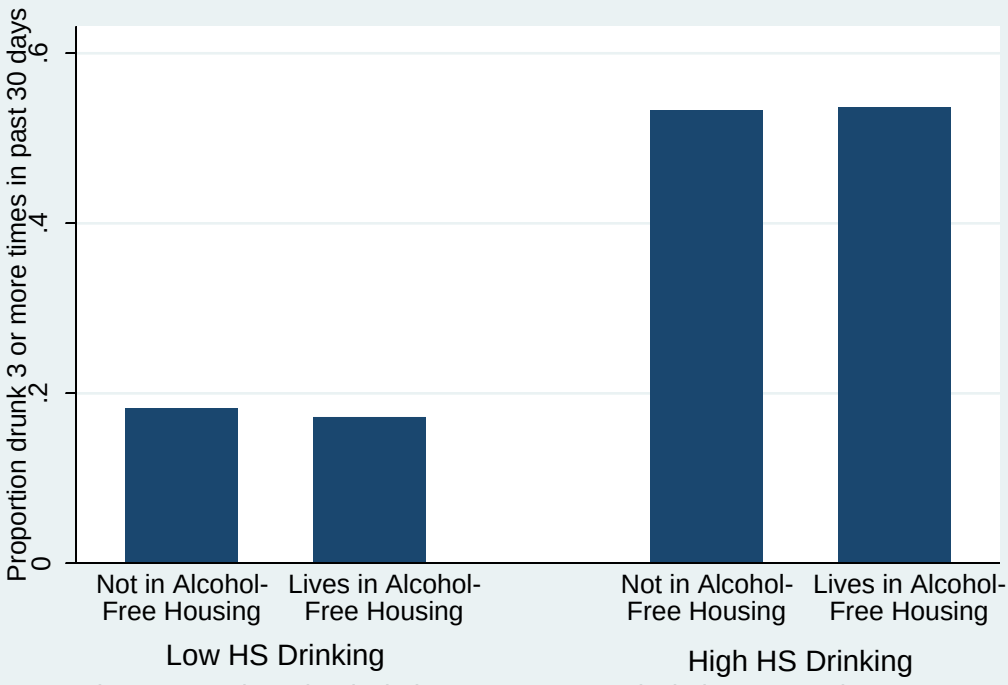



Figure 4: Students in Alcohol-Free versus not Alcohol-Free Housing Controlling for HS Drinking

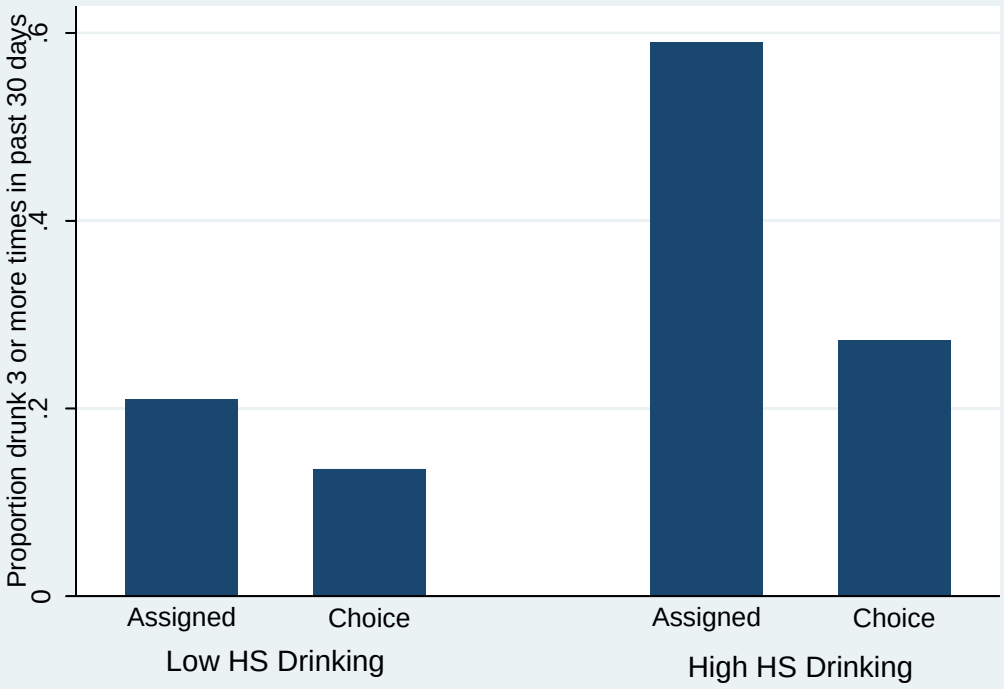



Figure 5: Students in Alcohol-Free Housing By Choice versus Assignment

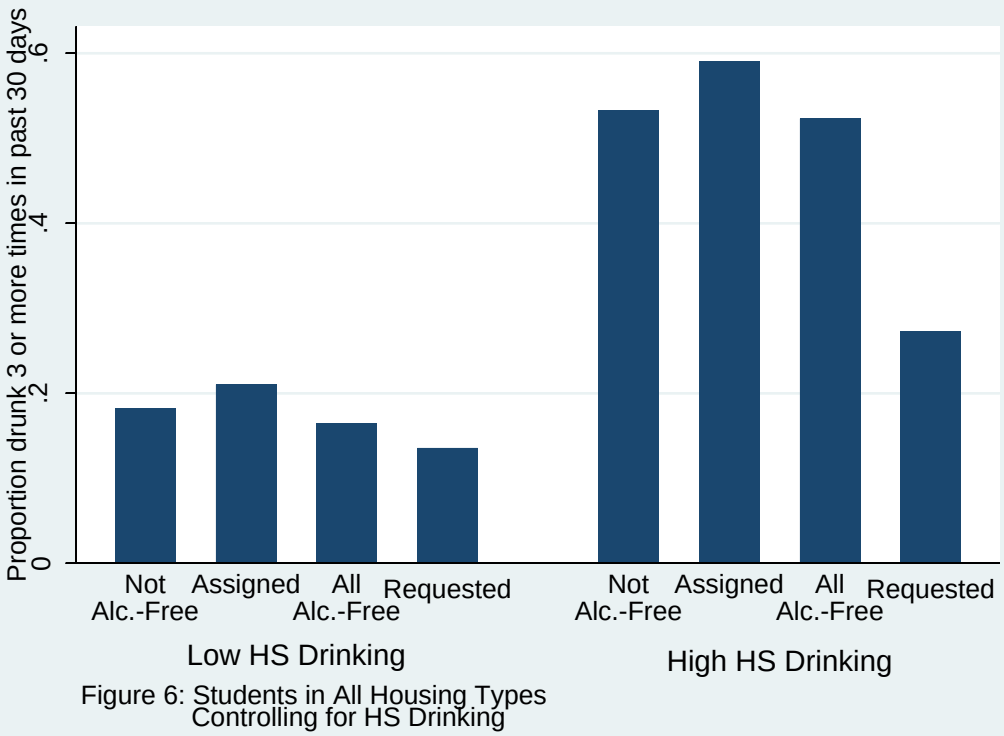



Figure 6: Students in All Housing Types Controlling for HS Drinking