Final Project

Daniel Pinckney

12/3/2020

```
## -- Attaching packages --------------------------------------------------- tidyverse 1.3.0 --

## v ggplot2 3.3.2      v purrr   0.3.4
## v tibble  3.0.3      v dplyr   1.0.2
## v tidyr   1.1.2      v stringr 1.4.0
## v readr   1.3.1      v forcats 0.5.0


## -- Conflicts ------------------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

Introduction

Does the type of building affect the level of LEED certification it receives? In this study, I plan to examine the extent to which new buildings receive higher LEED certification levels. I hypothesize that new buildings receive higher LEED certification than retrofits of existing buildings. When building new buildings, contractors and architects have more freedom to design within the LEED framework and are able to satisfy more of the criteria for higher certification levels. When retrofitting existing structures, builders must work with the existing shell of the building, potentially limiting their ability to reach the highest levels of certification. If I observe higher certification levels with new builds relative to existing structures, this would provide support for my hypothesis. If, on the other hand, I observe lower or similar levels of certification with new builds, this would provide evidence against my hypothesis. When I run my regression of the LEED certification level on the building type indicator variable, a positive, significant coefficient would indicate support for my hypothesis.

Data and Summarization

My sample is comprised of each of the LEED certified building projects in Montgomery County, Maryland from 2004 to 2015, excluding homes. This data is likely representative of most LEED projects since 2004, but the analysis could limited by the geographic restrictions. My unit of analysis is a class of building (e.g., a new construction). The explanatory variable of interest is whether the construction is a new building or a retrofit of an existing structure. The variable is coded =1 for new construction or core & shell and =0 for existing buildings and commercial interiors. An interesting covariate is the year built, as new building practices may also improve scores. My outcome variable is the average LEED certification level the building type receives. The variable is coded from 1 to 4, where 1 is a platinum level certification and 4 is certified. This variable is measured from US Green Building Councils LEED ratings that evaluate various green building strategies across several categories.

Table 1: Proportion of New Builds

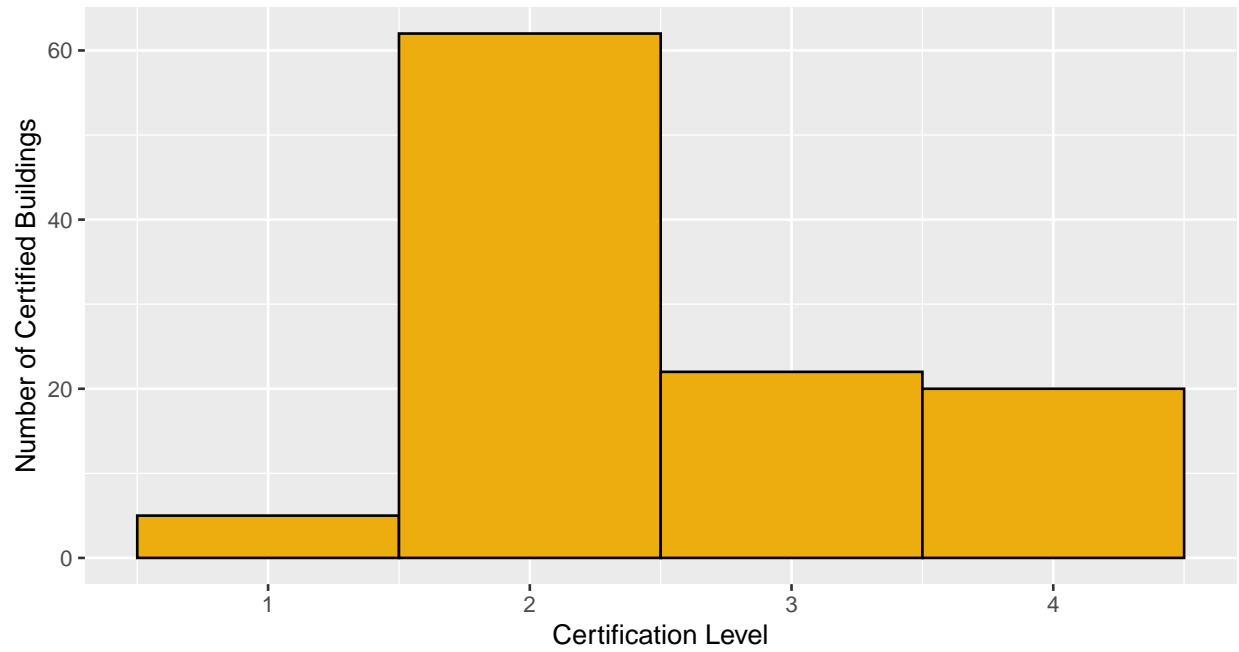| Proportion New Build |
|---|
| 0.6123596 |

Table 2: Average Certification Level

Mean Score
2.5

```
## `summarise()` ungrouping output (override with `.groups` argument)
```

Table 3: Average Treatment Effect of New Build

ATE	SE
0.0591677	0.1288733

Distribution of Building Certification Levels
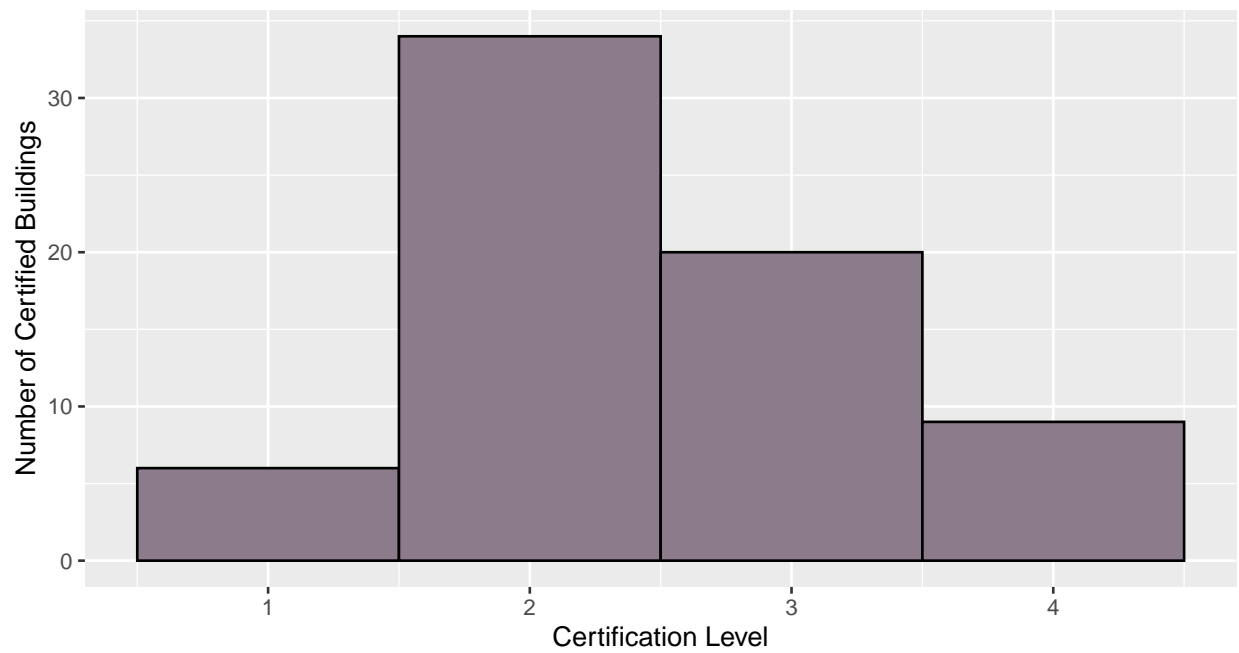
Distribution of New Building Certification Levels

Distribution of Existing Building Certification Levels

Results and Regression

	New Build Only	Add Date
New Build	0.059	0.074
	(0.129)	(0.129)
Certification Year		0.008
		(0.030)
Constant	2.464***	-14.082
	(0.101)	(60.913)
Num.Obs.	178	177
R2	0.001	0.002
R2 Adj.	-0.004	-0.009

* p < 0.1, ** p < 0.05, *** p < 0.01

Certification Levels Between New and Existing Buildings
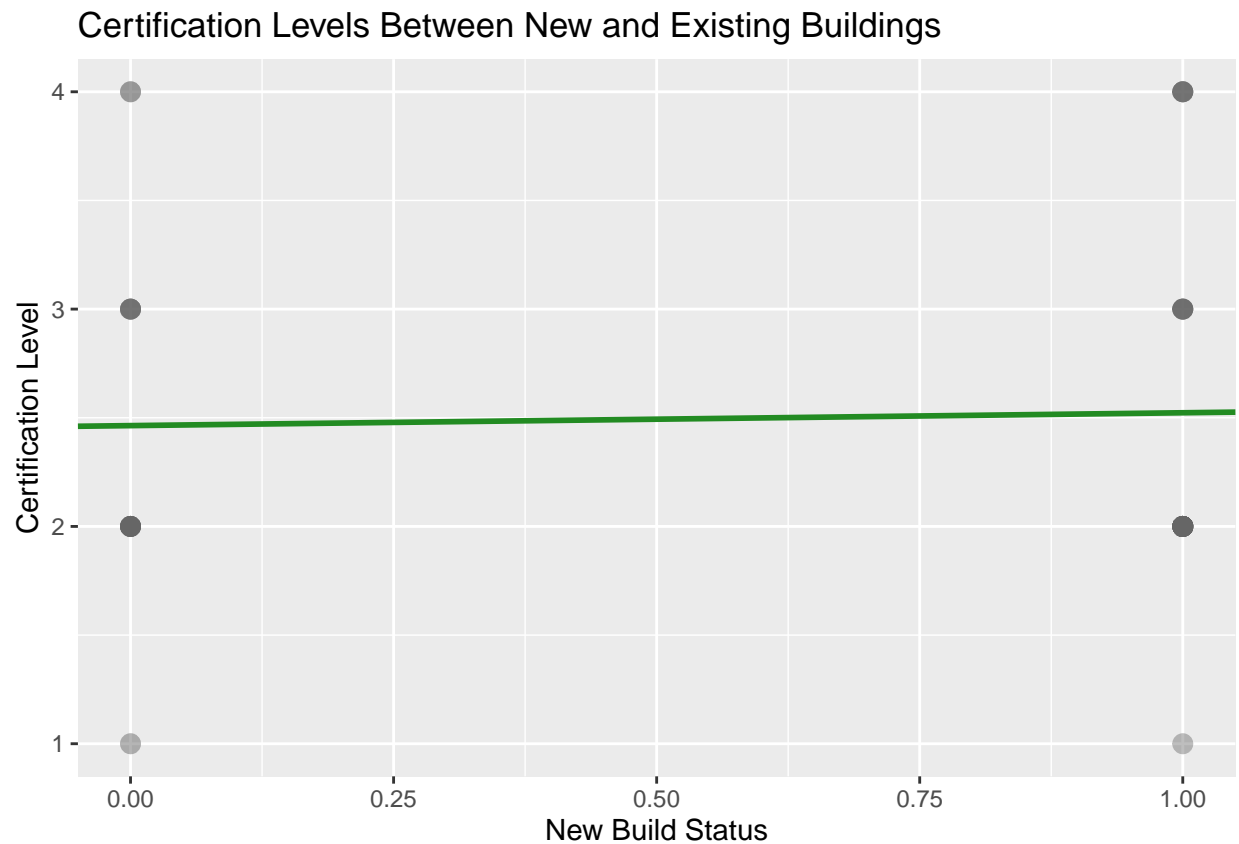
The main coefficient of interest, that on `new_build`, at first appears to support the hypothesis as it is positive. It takes a value of 0.05917, which means that for an one unit increase in `new_build`, moving from not being a new building to being one, there is an expected 0.05917 in its certification level. However, the coefficient does not appear to be statistically significant. The observed relationship could be due to chance error or some other confounder and the actual relationship could still follow the null hypothesis of no relationship. Thus, I cannot draw the conclusion that a causal effect exists and must fail to reject the null hypothesis.

Conclusion

a brief (one paragraph) concluding section that summarizes your results, assesses the extent to which you find support for your hypothesis, describes limitations of your analysis and threats to inference, and states how your analysis could be improved (e.g., improved data that would be useful to collect).