HW2

Son Luu (id), Xihan Qian (54285556) and Javier Mtz.-Rdz. (94785938)

Introduction

Quarto enables you to weave together content and executable code into a finished document. To learn more about Quarto see https://quarto.org.

Data Description and Summaries

When you click the **Render** button a document will be generated that includes both content and the output of embedded code. You can embed code like this:

```
1 + 1
```

```
[1] 2
```

You can add options to executable code like this

```
[1] 4
```

The `echo: false` option disables the printing of code (only output is displayed).

Exploratory Analysis

When you click the **Render** button a document will be generated that includes both content and the output of embedded code. You can embed code like this:

Formal analysis

Difference-in-Difference Estimation

DID, or Difference-in-Differences, is a quasi-experimental method that utilizes longitudinal data from intervention and control groups (Angrist and Pischke 2008). This approach is particularly effective in this design, as it aligns with the quasi-experimental nature of the study. This method aims to create a suitable counterfactual for estimating causal effects by comparing changes over time between the groups. The key assumption behind DID is that in the absence of the intervention, the average change in the outcome for the intervention group would have been the same as the average change for the control group.

Before applying the model, there are a few assumptions to be checked. Most importantly, there need to be parallel trends between the intervention and the control group. Moreover, at the baseline, the allocation of the intervention should not depend on the outcome. The model itself would be formulated as: $Y = \beta 0 + \beta 1 * [\text{ Time }] + \beta 2^* [\text{ Intervention }] + \beta 3^* [\text{ Time }^* \text{ Intervention }] + \varepsilon$. As there are different forms of interventions in the study, including discounts, calorie messaging, and combinations of the two, a linear model can be run first to determine which interaction terms (Time * Intervention) should be included.

In this proposed model, the key coefficient of interest for evaluating the impact of interventions is that of the interaction terms. Depending on the objectives, the analysis will focus on different coefficients of the interaction terms between time and the specific interventions. For example, to understand whether the interventions lead to increases in zero-calorie beverages, a positive and significant coefficient for zero-calorie beverages would indicate an effective increase in sales due to the intervention. A similar approach will be applied to investigate whether the combination of interventions lead to larger effects.

Conclusion

References

Angrist, Joshua D., and Jörn-Steffen Pischke. 2008. *Mostly Harmless Econometrics: An Empiricist's Companion.* Princeton University Press.