

Problem Set 1

ECON833, Prof. Jason DeBacker
Gyumin Kim

Q1. A brief summary of my interest in economics including in-text citations, a list of references, at least one equation, and at least figure.

My research interest lies in identifying causal relationships within the context of service operations. I was first inspired by my former advisor, Hyunseok Lee, whose work focuses on retail operations and causal inference. His influence motivated me to pursue empirical research, particularly in uncovering causal relationships.

To prepare myself, I took econometrics courses during my master's program and gained hands-on experience with econometric tools such as Stata. Among the causal inference methods, I focused most deeply on Difference-in-Differences (DID) and Instrumental Variables (IV), both of which are widely used in operations management.[3] While developing these technical skills, I also searched for a interesting research question. Because I had access to RFID fitting room data, I explored research topics related to fitting rooms. Around that time, I read a paper titled "Estimating the Stockout-Based Demand Spillover Effect in a Fashion Retail Setting" [5]. This paper applied the DID method to measure stockout-based size demand spillover effects in fashion retail settings and showed that incorporating these effects into transshipment policies could reduce substantial costs. Building on this idea, I assumed that such effects might be highly correlated with fitting room operations, and I began developing research question around them.

Inspired by this study, I explored the DID method in greater depth and eventually applied it in my master's thesis, where I used both DID and Difference-in-Difference-in-Differences (DDD) methods.).

Difference-in-Differences(DID)

According to [3]Teck et al. (2017), DID method is a powerful tool for estimating causal effect of the treatment, when the data are available for both the treatment and control groups in pre-and post-treatment periods. if there are observations on pre-and posttreatment periods for both the treatment and control group. It calculates the treatment effect by comparing the difference in the change in average outcomes over the two periods for the two groups.

In this context, what we call causal effect is the treatment effect for the treatment group. Table 1 illustrates the basic settings of DID. We have two groups: The treatment group, which is affected by the treatment and the control group, which is not. Each group is observed in two periods:pre-treatment and post-treatment, allowing us to capture the effect of the intervention. However, simple comparison of the treatment and control group is not enough to identify a true causal relationships. To do so,

we must rely on counterfactual. Therefore, causal effect is defined as the difference between observed outcomes of treatment group if they treated and counterfactual outcomes of the same group if they are not treated.(Joshua et al. 2015)[1]

Table 1: Matrix for understanding DID setting

	Pre-Treatment	Post-Treatment
Treatment Group	T_{pre}	$T_{post}(T_{counterfactual})$
Control Group	C_{pre}	C_{post}

According to the lecture [2], counterfactual outcomes for the treatment group can be written as $T_{pre} + (T_{counterfactual} - T_{pre})$. Here, T_{pre} represents the average outcome of the treatment group before treatment, and $T_{counterfactual}$ represents what the outcome would have been in the post-treatment period if the group had not received the treatment. Since, counterfactual outcomes are not observable in real life, we approximate $T_{counterfactual} - T_{pre}$ using the observed change in the control group, $C_{post} - C_{pre}$. If this assumption holds, then the causal effect can be estimated with observed outcomes.

$$\text{Causal Effect} = (T_{post} - T_{pre}) - (C_{post} - C_{pre})$$

[6]This is known as "parallel trend assumption". Without validating or verifying this assumption either quantitatively or qualitatively, we cannot guarantee that estimated effect reflect a true causal relationship. If the assumption does not hold, the result may be biased or contaminated by other factors such as differences in the characteristics of the two groups.

[3]Given this setting, we can obtain the DID estimate of the treatment effect on outcome Y through the coefficient of $W_{i,T}$ in the regression. β_{DID} , which we are interested in, is calculated as above.

$$Y_{i,T_i} = \alpha + \beta_{DID}W_{i,T_i} + \gamma G_i + \delta T_i + \epsilon_{i,T_i}$$

Fitting rooms and Difference-in-Differences-Differences(DDD)

As I mentioned at the start, Li et al.(2023) motivated me to study the relationship between the stockout-based size spillover effect and the fitting room operations.

According to Lee et al.(2021)[4], fitting rooms play significant role in consumer purchase decision, especially in searching information and evaluating alternatives. Compared to the other retail settings, customers who facing stockouts of their preferred size may choose different size of same clothing item. In such cases, customers may use the fitting rooms because they cannot try on the clothes in the middle of the stores. By allowing customers to confirm the fit of alternatives sizes, fitting room may increase confidence in the purchase decision. Therefore, the stockout-based size

spillover effect could be larger as the more customer bring the item into the fitting rooms.

To examine this, We applied similar approach to Li et al.(2023), who studied such effect using shoes. However, since customers typically do not bring shoes into fitting rooms, we focused on jeans, where fitting room usage is more relevant. This allows us to directly test the relationship between the stockout-based size spillover effect and the fitting room operations.

The treatment group is defined as jeans that are one size larger than size i when size- i jeans of product p are out of stock on day d in store s . The control group is defined as jeans that are two sizes larger under the same conditions. In addition, we created a new variable called “fitting rooms”, which captures the number of size- i jeans of product p that were brought into fitting rooms on day d in store s . Using these definitions, we applied a Difference-in-Difference-in-Differences (DDD) model, specified as follows:

$$Y_{j,s,d} = \alpha + \beta_{DDD}W_{ijsd} * fittingrooms_{j,s,d} + \gamma G_i + \delta T_i + \epsilon_{iT_i}$$

This model still needs further development, and additional robustness tests are required to strengthen the results.

In sum, through this research journey, I became deeply interested in econometrics, especially in causal inference methods within economics.

References

- [1] Joshua D. Angrist and Jörn-Steffen Pischke. *Mastering 'Metrics: The Path from Cause to Effect'*. Princeton University Press, 1st edition, 2015.
- [2] Data Science for Causal Inference. [week 5-1] (difference-in-differences) [video]. https://youtube.com/watch?v=SwPB3SocwLY&si=jV0_Z_c0YIjryrrh, 2023. Accessed: 2025-08-24.
- [3] Teck-Hua Ho, Noah Lim, Sadat Reza, and Xiaoyu Xia. Om forum—causal inference models in operations management. *Manufacturing & Service Operations Management*, 19(4):509–712, 2017.
- [4] Hyun Seok (Huck) Lee, Saravanan Kesavan, and Vinayak Deshpande. Managing the impact of fitting room traffic on retail sales: Using labor to reduce phantom stockouts. *Manufacturing & Service Operations Management*, 23(6):1580–1596, 2021.
- [5] Songtao Li, Lauren Xiaoyuan, Suan Feng Lu, and Simin Huang. Estimating the stockout-based demand spillover effect in a fashion retail setting. *Manufacturing & Service Operations Management*, 25(2):371–810,C2, 2023.
- [6] Ashesh Rambachan and Jonathan Roth. A more credible approach to parallel trends. *The Review of Economic Studies*, 90(5):2555–2591, 2023.