# Measuring the impact of assortment change in retail stores using the synthetic control method

**Nidhin VC**
**Rahul Prakash**
**Safil PM**
Growth Analytics Center
Anheuser-Busch InBev, Bangalore, India

**Keywords**
Impact measurement, Synthetic control, Assortment Optimization, Retail Industry, heterogeneous treatment effects, hypothesis testing

**Abstract**

*This study offers an innovative method for evaluating the effects of assortment changes on retail store revenue, employing the Synthetic Control Method (SCM) with Lasso regression. In response to the limitations of randomized control trials for large-scale interventions, we propose a unique approach using Lasso regression to build synthetic control groups from comparable stores in competitor retailers. This innovative technique overcomes the obstacle of absent control data within the same retailer. Our analysis reveals a remarkable 3.7% increase in retailer revenue after implementing the assortment changes. To confirm the validity of our findings, we conducted placebo studies, solidifying the positive impact of these adjustments. These results advocate embracing the SCM with Lasso regression as a reliable tool for measuring the effects of interventions in the business world, especially when controlled experiments are not feasible. This method empowers retailers to assess the effectiveness of assortment optimization strategies and make data-driven decisions about future changes.*

*Corresponding author: Nidhin VC*
*Email addresses for the corresponding author: nidhin.chandrasekhar@gmail.com*
*The first submission received: 27th October 2023*
*Revised submission received: 11th February 2024*
*Accepted: 25th of March 2024*

## Introduction

Determining the true effect of an intervention can be quite challenging. One method that has worked well is using randomized control trials, which involve randomly choosing groups of units to receive different treatments and comparing the results. However, this approach isn't always possible, especially when the intervention in question needs to be applied to all the units leaving no group untouched for comparison. Additionally, for randomization to be effective, you often need a lot of participants to make sure the groups are similar, but this method isn't very effective when you only have a small number of units involved.

A possible solution to overcome the limitations of a classical comparative study using randomized controlled trials could be the approach proposed by Abadie et al. to use "synthetic control method" where control units without intervention are created synthetically.

In this paper, we are detailing an approach in which we are customizing the synthetic control method for measuring the impact of assortment change in retail stores' revenues. Many retailers make changes in their assortment once or many times during a year in their attempt to increase their revenue and to serve their customers better. It is extremely important to measure the impact of these assortment changes on the retailers' toplines. But in many cases, the retailer changes the assortment of all their stores leaving no stores without intervention for a comparative study. So, we developed an approach where the control units are synthetically created using the Lasso regression. Here the revenue of similar stores from competitors which didn't undergo any intervention is used as independent features. We use syndicated data sources for our analysis. For each treated store, only those stores that are comparable in terms of store type, revenue, volume, and region are considered for creating the control set.

'Synthetic controls' have often been used in measuring the impact of policy interventions in the field of social sciences. However, we could not find any papers discussing the use of synthetic control methods in the retail industry. Through this paper, we are aiming to fill this gap by offering a comprehensive guide to synthetic control methods for the retail sector.

**Prior Literature**

Historically, randomized control trials (RCT) have been used in measuring the impact of interventions. In scenarios where we cannot use the RCT, the common alternatives are DiD and Synthetic Control Measurement. Has estimation been used in evaluating the impact of interventions, many recent studies (example: Ryan et al. (2014)) raise concern over the plausibility of parallel trends assumption in this approach. These studies point towards the importance of finding alternatives to DiD for measuring the impact of interventions.

The current study utilizes the synthetic control approach, initially introduced by Abadie and Gardeazabal in 2003 and further refined by Abadie, Diamond, and Hain Mueller in 2010. Recently, this method has seen diverse applications, including exploring the impact of right-to-carry statutes (Donohue, Aneja, and Weber, 2019), the legalization of prostitution (Cunningham and Shah, 2018), changes in immigration laws (Bohn, Lofstrom, and Raphael, 2014), the evaluation of California's anti-smoking initiatives (Alberto Abadie, Alexis Diamond & Jens Hainmueller, 2010), the influence of corporate political ties (Acemoglu et al., 2016), tax policies (Kleven, Landais, and Saez, 2013), criminal organizations (Pinotti, 2015), among other significant policy topics. Our research contributes to the body of knowledge on constrained regression analysis (Doudchenko and Imbens, 2017; Abadie and L'Hour, 2018; Minard and Waddell, 2018) and refined propensity score matching (Hainmueller, 2011; Graham et al., 2012; Zubizarreta, 2015; Tan, 2017; Wang and Zubizarreta, 2018). In our approach, we have adapted the conventional penalized regression framework by incorporating Lasso regression to better assess the connections between test and control groups.

Inference and testing for SCM: In this paper, we are making use of the placebo-based testing approach with a uniform permutation proposed by Abadie et al. (2010, 2015) for inference.

**Details on our business context**

An important challenge that retailers have been grappling with is to have the right assortment range for the overall beer category in their stores which would help them maximize their revenue. There is often a plethora of SKUs from various manufacturers occupying retailer shelves that don't contribute significantly to their incremental revenue. This is a huge challenge and an equally significant opportunity, where AB InBev is helping retailers with an unbiased, transparent, and sophisticated assortment recommendation product.

Ensuring the best beer products are always available in retailers' outlets is key to their success. To do that, AB InBev has developed a solution to optimize the SKU range. This has been covered in detail in another paper we wrote [How Anheuser-Busch InBev is tackling the top challenges of Category Management for modern offline retail channel]. In this paper, we intend to cover aspects around the measurement of impact for the retailer due to our solution.

Most retailers review and revise their assortment periodically to take care of changing consumer preferences, introduce new or seasonal products, and drive incremental efficiencies aimed toward increasing profitability. The assortment change typically happens twice or at least once a year for the fashion, food, and beverage industries. An optimized assortment not only helps the retailer and the manufacturers to maximize the revenue but also enhances the in-store consumer experience.

When a retailer makes an assortment change, it is very important to measure how it affects the revenue of the retailer. However, in many cases, it is difficult to conduct a randomized control trial for measuring the impact of assortment optimization. This is because in most cases the assortment gets changed in all the stores and there aren't any control stores to compare against. Even when the assortment change is only partial, finding equivalent and comparable samples in test and control groups could be difficult and, in many cases, the sample sizes are not statistically significant. In this paper, we apply the synthetic control method to study the impact of assortment changes on the revenues of retailers.

**Commonly used approaches for measuring the impact of interventions**

Assortment recommendations aim for revenue maximization for the retailer and increasing the overall category market share by enabling the shoppers to make easy decisions and thereby improving overall customer satisfaction. For measuring its impact, as mentioned above, a randomized approach may not always be possible as we may not always have control sets from the same retailer. Therefore, we need to rely on a non-randomized approach with the observational data.

We have tried some of the commonly used non-randomized approaches for measuring the impact. Some of the methods we used are mentioned below:

**Propensity score matching (PSM)**

In PSM, researchers use statistical techniques to create one control unit for each treated unit based only on the observed characteristics. In the one-to-one propensity score matching technique, pairs of treated and untreated units are formed based on similar values of the propensity score. Once the pairs have been formed, the treatment effect can be estimated by comparing the outcomes of treated and untreated units.

**Regression Discontinuity Designs (RDD)**

In RDD, a variable with continuous distribution is selected to define test and control units. A threshold value is defined on this variable so that the units at either side of this threshold value are subjected to the treatment. By comparing the units lying closely on either side of the threshold, it is possible to estimate the treatment effect.

**Difference in difference (DiD)**

DiD approach measure the impact by contrasting the change in outcome variable pre and post treatment for the treatment and control units. If the treated and control units had same outcome variable trend in the pre-intervention period, and if the difference or gap between the treated and control unit changes (enlarges or shrinks) in the post-intervention period, this can be attributed to the treatment effect.

In DiD, sometimes it is difficult to evaluate if the control unit is accurate for the treated unit and if the assumption of parallel trends between treated and control unit in pre-intervention period is met.

**Synthetic Control Method (SCM)**

The synthetic control method (SCM) is about creating a weighted combination of control units. We compare this with the treatment group. This method helps in balancing out any outside factors that might change as time goes on. It does this by adjusting the control group to resemble the treatment group more closely before any action is taken. Another advantage of SCM is that it provides a structured way for researchers to choose which control groups to use for their comparisons.

**Why Synthetic control suits our requirement**

Out of the methods mentioned in the previous section, we find Synthetic control methods to be most useful for our business context. This is because DiD, PSM, and RDD have characteristics that make them unsuitable for measuring the impact of assortment optimization. RDD separates the population into test and control using the threshold value of a single variable. In our case, the decision to select test stores is solely at the discretion of the retailers and is not based on any single variable such as revenue or volume. In many cases, they opt for 100% implementation.

When the assortment recommendations are delivered to a retailer, the expectation is that they would implement them in all their stores. This would make it difficult to estimate meaningful propensity scores using (PSM) due to the unavailability of control stores from the same retailer.

One of the challenges in effectively using DiD for measuring the impact of assortment optimization is finding certain groups of stores from the same retailer that are exposed to treatment and certain others that are not. For many retailers across the geographies that we have worked with, the assortment optimization got implemented in all their stores at the same time. This made it impossible to identify test and control sets where both the groups were not subjected to assortment treatment in the first time period and only one of the groups got exposed to assortment optimization treatment in the second period.

When pre-treatment trends are not parallel or if an event occurs that impacts only one of the units or groups, relying solely on the difference in difference (DiD) method can lead to biased estimates of the treatment effect. In such cases, the deviation between the groups' trends will reflect the difference in trends rather than the true effect of the intervention. This bias can distort the estimated treatment effect and hinder accurate analysis. To overcome the challenges mentioned, we turned to the synthetic control method, which proved advantageous in identifying suitable control stores by utilizing data from other retailers.

A significant benefit of synthetic control over DiD method is that, unlike DiD , synthetic control doesn't assume parallel pre-implementation trends. Parallel pre-implementation trends are not applicable in this case because multiple factors affect the revenue when the retailers have stores across the country. The synthetic control method also avoids extrapolation biases and allows a more focused description and analysis of the similarities and differences between the case of interest and the comparison unit (Alberto Abadie et. al 2010.)

**Customizing the Synthetic Control Method for measuring the impact**

SCM assumes that the pre-intervention characteristics of the treated unit can often be much more accurately approximated by a combination of untreated units than by just one untreated unit.

In our case, since the assortment change can happen in all the stores of a particular retailer, the control stores need to be from other retailers. As a data-driven statistical approach, the goal of SCM is to create a group of control units for the treated unit, that has similar pre-intervention characteristics. It allows the construction of a counterfactual by calculating a weighted average of the outcome variable from the control groups. The weights generated through techniques such as Lasso regression are assigned to each control unit to estimate the synthetic value. This weighted average approach ensures that the control units closely resemble the treated unit, thereby allowing for a more accurate estimation of the treatment effect.

Because comparison units are meant to approximate how the treated units would have behaved without the intervention, it is important to restrict the donor pool to units with outcomes that are impacted by the same external shocks and factors that the treated group was subjected to. For example, the group of stores selected to compare the revenue of a particular UK supermarket chain should also be subjected to similar lockdowns and other pandemic-related restrictions, economic factors like inflation, etc. The selection process of similar stores also ensures that the stores are of similar type, belong to the same region, and exhibit a similar magnitude of volume and sales.

<u>Creating Synthetic Control groups</u>

Consider a scenario with J+1 retailer stores, each indexed by j. Among these, unit j=1 is the focal store where an intervention occurred, and units j=2 to j=J+1 serve as potential comparisons, forming a "donor pool." The creation of a synthetic control group involves deriving a weighted average from the donor pool units. This entails representing the synthetic control as a vector of weights, W = (w2,wJ+1), with each weight wj satisfying $0 \leq wj \leq 1$ and the sum from w2 to wJ+1 equalling 1 for j=2,J.

The selection of a specific W value corresponds to choosing a synthetic control. The aim is to determine W in a way that aligns the characteristics of the treated unit with those of the synthetic control.

To achieve this, for each variable m (from 1 to k), X1m represents the mth variable value for the treated unit, and $X_0$m is a 1×J vector containing the mth variable values for the units in the donor pool. Abadie and Gardeazabal (2003), along with Abadie, Diamond, and Hainmueller (2010), opt for W which minimizes the following expression:

$\Sigma V_m(X_{1m} - X_{0m}W)^2$

were

$$X_{0m} = \begin{bmatrix} x_{2m} & x_{3m} & -- & x_{j+1m} \end{bmatrix}$$

$$\begin{bmatrix} w_2 \\ w_3 \\ -- \\ w_{j+1} \end{bmatrix}$$

and $V_m$ is a weight that reflects the relative importance that we assign to the $m^{th}$ variable when we measure the discrepancy between $X_1$ and $X_0W$.*

**Arriving at our solution**

As we already discussed above, in synthetic control we have a (J ×1) vector W of positive weights that sums up to one. That is, for vector $W = (w_2, w_{J+1})$ with $w_2 + \cdots + w_J+1 = 1$ and $w_j \geq 0$ for j = 2, J + 1. Each value of $(w_2, w_{J+1})$ represents a weighted average of the similar control stores and, therefore, a synthetic control.

In our approach, we identify similar stores for each treated store from the set of control stores. We used stores from five other retailers of the same country as the control stores. In order to overcome the ambiguity prevalent in selecting the control sets, we followed a data-driven procedure while constructing the control stores for each test store, as in Abadie and Gardeazabal (2003). Here we use observed quantifiable characteristics such as region, store type, volume level, and revenue pattern to reduce discretion while constructing the control sets.

Normally Synthetic control method doesn't allow the extrapolation of the outcome variable as the weights are restricted to be non-negative and they must sum up to one. This requirement can however be relaxed at the cost of allowing extrapolation. Abadie, Diamond, and Hainmueller (2015) show that in the context of estimating the effect of a treatment, a regression estimator can be represented as a synthetic control with unrestricted weights. The regression estimator then represents the weights that are not restricted in the range of [0,1] thereby allowing extrapolation. This approach helps to overcome the problem where the outcome variable of all the units in the control group is either smaller or larger than that of the treated unit. To accommodate this adjustment, we are using Lasso regression.

**Lasso regression**

The Lasso method applies a linear regression framework that incorporates a shrinkage technique, which pulls data points toward a central value, typically the average. It promotes the creation of models that are both streamlined and sparse, meaning they have a reduced number of parameters. Lasso regression is particularly effective in scenarios with significant multicollinearity or when the goal is to streamline the process of choosing a model by automating the selection or reduction of variables and parameters.

In our study, we employed Lasso regression to estimate the weights in the synthetic control method. This decision was made due to the need for extrapolation in certain cases where all the control stores had revenue values that were either smaller or larger than the revenue of the test store.

In such scenarios, it becomes challenging to create a synthetic revenue value by using a convex combination of similar control stores. A convex combination assumes that the weights are non-zero and sum to 1.

***Sample Lasso Regression Equation***

Revenue of treated store number 1 = Revenue of control store number 2 x 0.4 + Revenue of control store number 4 x 0.2+ Revenue of control store number 11 x 0.2 + Revenue of control store number 8 x 0.2

Once the (J × 1) vector of weights W = $(w_2, w_{J+1})$ are estimated for j = 2, J + 1, J = 1 being the test store, the value of the outcome variable in post implementation period is calculated by

$$\sum_{j=2}^{J+1} w_j Y^I_{jt} = Y^N_{it}$$

were

$$W = \begin{bmatrix} w_2 & w_3 & \cdots & w_{j+1} \end{bmatrix}$$

$$Y^i_t = \begin{bmatrix} y_{2t} \\ y_{3t} \\ - \\ y_{(j+1)t} \end{bmatrix}$$

and $Y^N_{it}$ is the revenue that would be observed for store i at time t in the absence of the implementation, $Y^I_{it}$ is revenue that would be observed for store i at time t if the store i is exposed to the implementation. Then the effect of implementation is calculated as the difference between $Y^N_{1t}$ and $Y^I_{1t}$ for each time to post implementation.

Measured impact of implementation for store i at time t = $Y^I_{it} - Y^N_{it}$

**Data & Sample**

We used weekly sales data of all the stores for the retailer in which our assortment recommendations are implemented. We have used 104 weeks of sales data, from Oct 2019 to Oct 2021. Recommendations were implemented at the store level in April 2021, giving us 78 weeks of pre-implementation data for training and validating the model, and the remaining 26 weeks of post-implementation data for testing.

Synthetic revenue is constructed as a weighted average of potential control stores, with weights chosen so that the resulting synthetic revenue best reproduces the revenue of implemented stores in the absence of assortment change – that is in the pre-implementation period.

We have used statistical techniques to remove outliers from the analysis. We have removed both outlier stores and time periods in the pre-intervention period as including them in the analysis would bias the estimation of weights.

Our outcome variable of interest is the weekly revenue at the store level. To create synthetic values of weekly revenue, we considered stores from other five retailers where our recommendations were not implemented in the same period. We had a total of 592 stores from the implemented retailer and close to 2400 stores from the control retailers. Using the techniques described above, we created the synthetic revenue for the implemented stores using the weights associated with similar stores. We estimated the effect of assortment change as the difference between the actual and its synthetic versions in the weeks after implementation. We then conducted Placebo studies to confirm that the estimated effect of assortment change is not by chance. This has been explained subsequently in this paper.

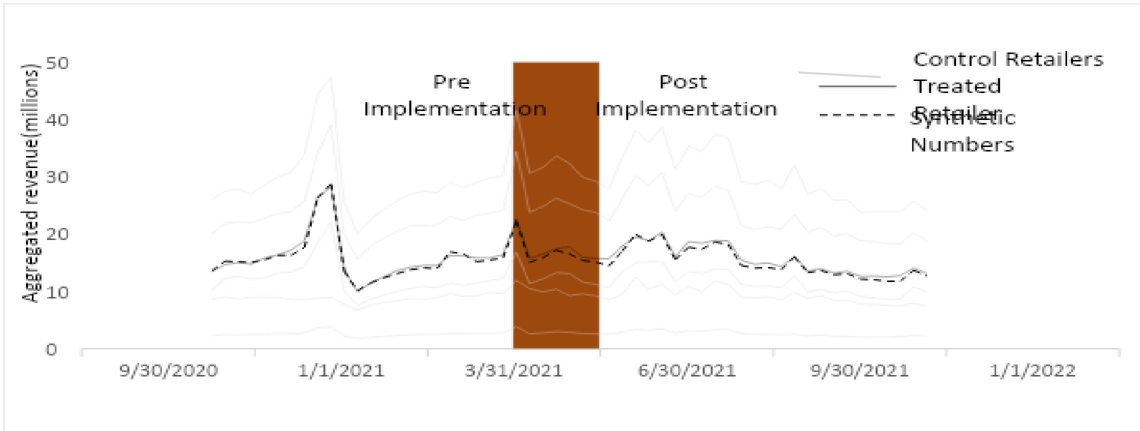


Figure 1: Calculating the aggregated synthetic revenue from the control retailers

**Results of the Synthetic Revenue Calculation**

As explained above, we construct the synthetic revenue for all the 592 stores of an implemented er from the weights of stores in the control set that most closely resembled each implemented store before implementation. Then we aggregate all the individual stores' synthetic revenue to get the retailer's synthetic revenue and compare it with the actual revenue of the retailer.

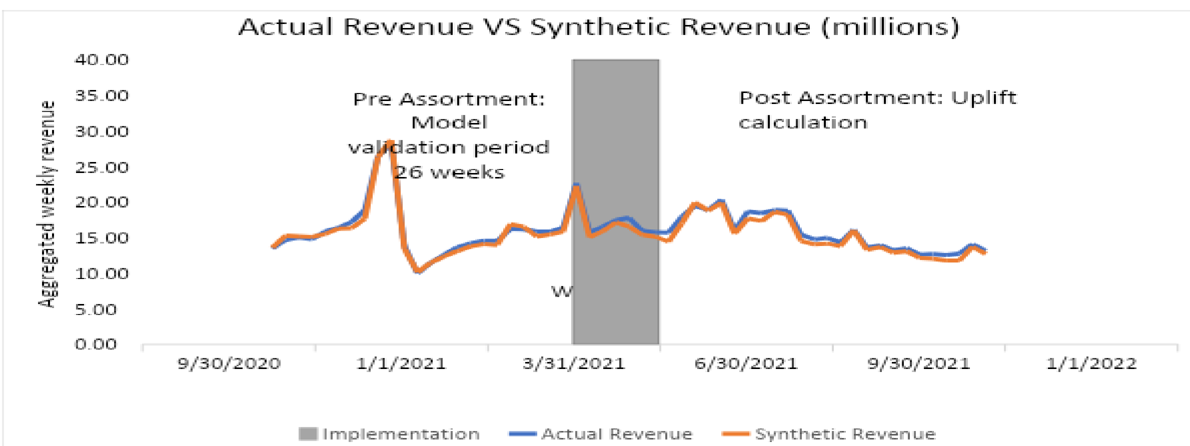


Figure 2: Comparison of Actual Revenue vs Synthetic revenue for the stores where the assortment changes are implemented.

We used 52 weeks of pre-implementation data for training the model (October 2019 to September 2020), and 26 weeks of data for validating and fine-tuning the model (October 2020 to March 2021). Post-implementation, we can see the effect of implementation. We track an uplift in revenue due to assortment implementation from April 2021 onwards.

*Table 1. Actual Revenue and Synthetic revenue averages

| Outcome | Average Revenue (Millions) Pre-implementation | Post-implementation |
|---|---|---|
| Actual Revenue | 16.19 | 15.78 |
| Synthetic Revenue | 15.98 | 15.22 |

* The revenue figures presented in this paper are for illustrative purposes and do not represent actual data

Here we are getting an accuracy of 99% during the pre-implementation period. In the post-implementation period, the estimated impact of our assortment recommendation is a 3.7% uplift.

**Placebo Studies**

To evaluate the credibility of our results, we conducted placebo studies where we considered one of the control retailers as treated retailer and used other control retailers as its control groups. In this way, we obtained the synthetic control numbers for the retailer which in reality had not undergone any treatment. This allowed us to compare the estimated impact of assortment implementation for the treated retailer (placebo retailer) to the corresponding synthetic revenue of other retailers. We will deem the impact of assortment implementation on the treated retailer significant if its estimated impact is significantly large relative to the placebo effect.

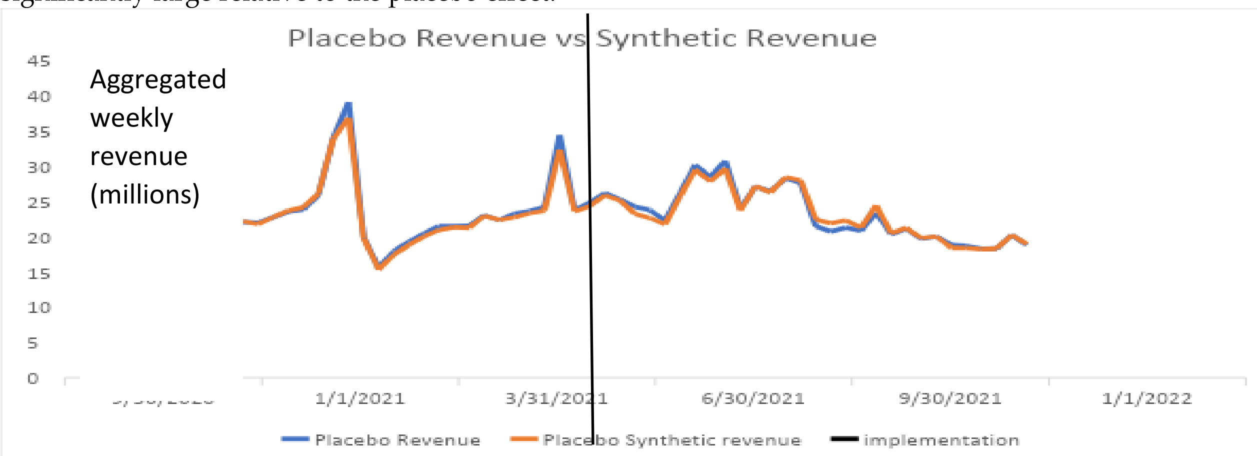


Figure 3: Placebo revenue vs its synthetic revenue.

Table 2. *Actual Revenue and Synthetic revenue averages for the placebo retailer

| Outcome | Average Revenue (Millions) Pre-implementation | Post-implementation |
|---|---|---|
| Actual Revenue | 23.50 | 23.28 |
| Synthetic Revenue | 23.18 | 23.21 |

* The revenue figures presented in this paper are for illustrative purposes and do not represent actual data

From the table above, we observe an impact of 0.3% for the placebo retailer which is significantly smaller than that of the implemented retailer (3.7%) as we saw before. This analysis suggests that the impact shown for the treated retailer in Figure 2 is the result of assortment implementation.

**Limitations**

The presented study on measuring the impact of assortment change in retail stores using the Synthetic Control Method (SCM) is comprehensive and detailed. However, like any research, there are certain limitations to consider:

1. Assumption of Similarity: The success of this method relies on the assumption that the control units used in the synthetic control group are sufficiently similar to the treated unit. In some cases,

finding similar stores may be challenging, and there could be unobservable differences that affect the outcomes.

2. Data Quality and Availability: For building the control group, data of similar stores from competitors that didn't undergo any intervention is used. There could be scenarios where the data from competitors is not available.

**Implications**

Theoretical implications

1. Addressing Challenges in SCM: The paper addresses some of the challenges associated with using the synthetic control method, such as the lack of control groups in observational data and the need for extrapolation in some cases. The use of Lasso regression helps overcome these challenges and improve the accuracy of the method.
2. Expanding the Applications of Synthetic Control: This paper demonstrates the successful application of the synthetic control method in a business context, specifically for measuring the impact of assortment changes on retailer revenues. This expands the potential applications of the method beyond its traditional use in policy evaluation.

**Practical Implications**

**Refined Impact Measurement for Retailers:**

Retailers can practically benefit from adopting the proposed SCM methodology to refine impact measurement. Accurately gauging the effects of assortment changes on revenue offers actionable insights crucial for strategic decision-making, assortment planning, and resource allocation.

**Discussion and Comparison with Current Literature**

Our study contributes to the ongoing research on measuring the impact of interventions in observational settings, particularly within the retail industry. Our utilization of the synthetic control method with Lasso regression addresses specific challenges faced by conventional approaches like randomized controlled trials and difference-in-differences, making it a valuable tool for retailers seeking to accurately evaluate the effectiveness of assortment optimization strategies.In this discussion section, we will compare and contrast our key findings with the existing literature, highlighting the contributions and implications of our research.

**Comparison with Existing Methods:**

Randomized Controlled Trials (RCTs): While considered the gold standard for causal inference, RCTs are often impractical in retail settings due to the difficulty of isolating control stores and ethical concerns associated with withholding interventions from certain customers. Our approach overcomes these limitations by leveraging data from other retailers to construct synthetic control groups, enabling robust impact estimation even in the absence of true randomization.

Difference-in-Differences (DiD): DiD relies on the assumption of parallel pre-trends between treatment and control groups. However, pre-existing differences in factors like seasonality or marketing campaigns can bias DiD estimates. Our study addresses this limitation by utilizing the synthetic control method, which can account for pre-intervention trends and adjust for discrepancies between the treated and control units.

Propensity Score Matching (PSM): PSM attempts to balance treatment and control groups based on observed characteristics. However, its effectiveness depends on the availability of relevant covariates and can suffer from model specification biases. Our Lasso regression-based approach offers an advantage by mitigating the need for extensive covariate selection and automatically selects the most relevant control units for each treated store.

**Key Findings and Alignment with Literature:**

Positive Impact of Assortment Changes: Our study finds a 3.7% uplift in retailer revenue post-assortment change, corroborating findings from existing research highlighting the potential benefits of optimized assortments for driving sales and profitability (Smith & Stephens, 2005; Gopalakrishnan & Krishnan, 2014).

Role of Synthetic Control in Business Settings: This study extends the use of the synthetic control method beyond its traditional application in social sciences, demonstrating its effectiveness in measuring the impact of business interventions like assortment optimization. Previous research has explored the potential of SCM in business contexts, but primarily focused on evaluating marketing campaigns or pricing strategies (Xu, 2017; Cavallo & Zhang, 2020). Our work strengthens the case for adopting SCM as a reliable tool for impact measurement in the retail sector.

Importance of Accounting for Unobserved Heterogeneity: Our utilization of Lasso regression acknowledges the presence of unobserved factors that may influence retailer revenue beyond controllable variables. This aligns with recent studies emphasizing the need for robust estimation techniques that account for unobserved heterogeneity in causal inference (Abadie et al., 2015; Ferretti & Peretti, 2016).

**Conclusion**

In this paper, we have examined an approach to measure the impact of assortment change on retailer revenues. We have used a variant of the synthetic control methodology using Lasso Regression to identify the control stores with matching store type, revenue volume, region, etc. as independent variables. This allowed us to create a control group that would closely mimic the behaviour of the test stores in the pre-implementation period.

Before resorting to SCM, we tested our other popular approaches such as PSM, RDD and DiD. To ensure the credibility of our synthetic control model, we deployed placebo studies. As a result of our synthetic control experiment, we were able to observe a significant uplift in the retailer’s revenue.

Although our focus here has been on measuring the impact of assortment change on retailer revenues, the method proposed in this paper can be used to investigate the impact of various other interventions on a retailer’s revenue.

**References**


Abadie A, Gardeazabal J. (2003). The economic costs of conflict: A case study of the Basque Country. American Economic Review 93: 113–132

Abadie A. 2005. Semiparametric difference-in-differences estimators. The Review of Economic Studies 72: 1–19

Athey, S. and Imbens, G.W. (2006), “Identification and Inference in Nonlinear Difference-inDifferences Models”, Econometrica , vol. 74, no. 2, 431-497.

Begay, M., Traynor, M., and Glantz, S. (1993), “The Tobacco Industry, State Politics, and Tobacco Education in California”, American Journal of Public Health, vol. 83, no. 9, 1214-1221

Bertrand, M., Duflo, E., and Mullainathan, S. (2004), “How Much Should We Trust Differences-in-Differences Estimates?”, Quarterly Journal of Economics, vol. 119, no. 1, 249-275.

Abadie, A. and G. W. Imbens (2011). Bias-corrected matching estimators for average treatment effects. Journal of Business & Economic Statistics 29 (1), 1–11

Abadie, A. and J. L’Hour (2018). A penalized synthetic control estimator for disaggregated data

Ando, M. and F. S¨avje (2013). Hypothesis testing with the synthetic control method

Botosaru, I. and B. Ferman (2017). On the role of covariates in the synthetic control method

Mitze T, Kosfeld R, Rode J, et al.(2020) Face masks considerably reduce COVID-19 cases in Germany. Proc Natl Acad Sci U S A. 2020;117(51):32293–3230

Dimick JB, Ryan AM(2014). Methods for evaluating changes in health care policy: the difference-in-differences approach. JAMA. 2014;312(22):2401–2402.

Doudchenko N, Imbens GW.(2017) Balancing, regression, difference-in-differences and synthetic control methods: a synthesis [preprint]. arXiv. 2017. ttp://arxiv.org/abs/1610.07748

Bertrand M, Duflo E, Mullainathan(2004) S. How much should we trust differences-in-differences estimates? Q J Econ. 2004;119(1):249–275

Ferman B, Pinto C. (2019) Inference in differences-in-differences with few treated groups and heteroskedasticity. Rev Econ Stat. 2019;101(3):452–467

Becker M, Klößner S(2018). Fast and reliable computation of generalized synthetic controls. Econ Stat. 2018;5:1–19.

Cassel, C. M., C.-E. Sarndal, and J. H. Wretman (1976). Some results on generalized difference estimation and generalized regression estimation for finite population. Biometrika 63 (3), 615–620.