A central tool in empirical operations management research is treatment effect estimation from observational data (when units are not assigned to the treatment or control group randomly). For instance, in retail, companies might be interested in knowing the effect of natural disasters on their supply chain or sales; healthcare organizations might be interested in studying effects of operational decisions (e.g., adopting a new technology) or clinical decisions (e.g., prescribing a new medication). These are all examples of causal inference problems in which for measuring the effect of a treatment, the counterfactual outcomes need to be estimated and then be compared to the realized outcomes. However, such analysis from observational data is challenging due to potential confounders that can affect both the treatment and the outcome, leading to biased estimates for the treatment effect (also known as the “omitted variable bias” or OVB).

One common approach to reduce OVB is to use panel data models where a subset of units is exposed to a binary treatment during a subset of time-periods and the goal is estimating the counterfactual (untreated) outcomes for all treated units/time-period combinations. For addressing this task, the existing literature has mainly focused on two different types of estimators. The program evaluation literature (see [1] for a comprehensive review) uses unconfoundedness assumption in order to build estimators based on the stable patterns of outcomes over time which are universal among the units. In other words, this approach requires all units to evolve with the same structure over time. Program evaluation techniques usually rely on the existence of a large number of units relative to the time-periods.

On the other hand, in the synthetic control literature (see e.g. [2]) a different approach is taken. The idea here is to express outcomes of the treated units as weighted combinations of the outcomes of the control units. Weights are chosen to ensure that in the pre-treatment periods, weighted outcomes
for control units match approximately with the outcomes of treated units. These methods work well when we have access to a substantial time-periods of outcome for each unit. The key assumption underlying the synthetic control methods is the existence of a stable relationship between the outcomes of the treated and control units over time.

Methods in each of these two literatures exploit only a fraction of plausible patterns in the data. In particular, the program evaluation techniques seek for the consistent patterns over time, while the synthetic control techniques seek for the consistent patterns across the units. Following the recent advances in the machine learning and statistics literature on matrix completion \cite{3,4,5}, we consider the low-rank models which allow for richer models and can incorporate both patterns within units and time-periods. We introduce and study a class of estimators that minimize the distance between the estimated matrix and the original (incomplete) matrix, while penalizing models with higher complexity. Incorporating this penalty term which favors matrices with smaller ranks, leads us to choose a matrix which trades off between the accuracy and the generalizability well. Indeed, we consider a general class of matrix norms, the Schatten norms, that penalize the sum of powers of the singular values. In order to prove the consistency of our method, we generalize the existing results in the matrix completion literature by allowing the patterns of missing data to include a dependency structure. In particular, we assume that the missing portion of each row forms a block of entries ending on the right side of the row. The motivation behind this assumption is that, for each treated unit, the missing entries correspond to the time periods after an adoption time.

In order to illustrate the performance of our method, we work with a publicly available dataset\(^1\), consisting of sales of 111 potentially weather-sensitive items in 45 Walmart stores across the US on a daily basis for almost three years. We have also been provided with daily weather reports of the closest weather station to each store. The treatment here is any weather event that occurred during the interval

\(^1\) The dataset can be downloaded here: https://www.kaggle.com/c/walmart-recruiting-sales-in-stormy-weather
of study. The goal is to estimate the average treatment effect or in other words approximate the average change in the sale of items that is caused by the occurrence of weather events. Knowing the answer to this question leads to better inventory management and pricing decisions. We find that our method outperforms the program evaluation methods in terms of the accuracy of estimations. The synthetic control methods cannot be applied in this setting, due to a large number of treatments (weather events).

In a second numerical study with the data, we boost program evaluation and synthetic control methods by imputing the data using the simple mean-imputation approach, while our matrix completion method does not receive the same data imputation benefit. Despite having access to a smaller number of observed entries, our matrix completion method performs well and together with the augmented synthetic control methods achieves the best accuracy. As the final part of the simulation, we predict the average treatment effect using different methods and compare the results.

Additional simulations on other datasets are provided. A conclusion that can be inferred from these simulations is that the program evaluation methods work well in panels with large number of units compared to time-periods, while synthetic control methods work well in panels with large number of time-periods relative to units. The advantage of the matrix completion method is that it moves seamlessly between these two approaches irrespective of the relative number of time periods and units.

References: