[1]

A MULTIVARIATE MATCHING STRATEGY FOR TIME SERIES CAUSALITY: OZONE EFFECTS ON MORTALITY

YOUNG LEE IBON TAMAYO-URIA MARIE-ABÈLE C BIND

Harvard University

Abstract. When drawing inferences on causal effects using observed data, it is desirable to duplicate a randomized experiment as close as possible in some sense by obtaining treated and control groups with resemling attributes. In this article we propose a multivariate matching strategy for time series with an application to obtain the causal effect of ozone level on mortality.

1. Introduction

Matching can be loosely thought of as any procedure that aims to unify the distribution of covariates in both the treated and control groups. This aim can be achieved by choosing *matched* samples of the treated and control groups with the intention of reducing bias due to presence of the covariates (Rosenbaum and Rubin, 1983). The intention of matching is, for every treated unit, to find one or more control units with similar attributes against whom the effect of the treatment can be assessed. By matching treated units to similar control units, matching enables a comparison of outcomes among treated and control units to estimate the effect due to the treatment. In recent years, there have been many literature works concerning matching and a good survey can be found Stuart (2010).

For many decades, researchers have tried to analyze time series through causal methods. Haavelmo (1944) and later Koopmans (1950) used means of a structural equation model to estimate causality. These methods have

[1]**Preliminary working draft - please do not quote or cite without authors permission.** This paper builds on the manuscript under a similar title authored by Ibon Tamayo-Uria, Young Lee, Stephane Shao, Luke Miratrix, Donald Rubin and Marie-Abèle C Bind.

Date: November 11, 2019.

been historically applied on econometric data, but the last years they have extended to other fields such as brain image analysis Zhang et al. (2017), climate change Guo et al. (2018) or air pollution Schwartz et al. (2017). But all these works do not avoid considering a relationship between factors as a causal effect. In this paper, we follow the path laid out in Bind and Rubin (2017) where the authors proposed four clear stages (conceptual stage, design phase, analysis phase and summary stages) to consider an observational dataset as a hypothetical randomized experiment. In this work, based on these phases, we propose a multivariate matching strategy for time-series data to answer scientific causality questions. Specially, we aim to develop an effective method to detect causal relations between multiple exposures and an outcome.

One of the potential fields where this method can be applied is epidemiology. This is why we tested our method in a epidemiological example. The harmful effects of air pollution on health could increase the probability to develop respiratory or cardiovascular diseases, cancer, even premature deaths. Air pollution is a mixture of pollutants and hence, it is difficult to specify the weight of each pollutant in each disease. In some cases the problem is created by the size of the molecules, while in others, the material properties. Once the importance of the pollutant is defined, protective and preventive solutions can be applied. When a pollutant is the origin of a disease, doctors can better understand the reaction pathway and search for protective solutions. On the other hand, identifying the specific concentrations where health problems become important, can be a very powerful tool for decision makers; they can use this information to validate legal limits and carry out a preventive strategy. Therefore, it is important to define a robust method that can help to find the causal factor of mortality.

Ozone is related with health problems, premature deaths, agricultural decrease production and several damages in urban infrastructures. Identifying the specific effect of a pollutant requires separation from the other copollutants' effects. Seasonal, weather and other factors could be strong confounders and it becomes more difficult to estimate the specific effect. Ozone is formed by the interaction between light and other pollutants, mainly those related to road traffic. So separating it can become a complex process. To carry out a new causality approach for time series, we focused on one city, to assess if the current legal threshold (70ppb) is enough to avoid effects on mortality. This approach could be complex, and providing practical tools allows for greater diffusion and can open opportunities for these techniques to be applied. For this

reason, in addition to developing the method, we have created a user-friendly Shiny application where anyone can perform the analysis.

The remainder of the paper is organized as follows. In Section 2, we describe the data under consideration that is obtained from the National Center for Health Statistics. In Section 3, we apply the methodology of with regards to applying the principles of Bind and Rubin (2017) to our setting. In particular, we highlight the Bayesian imputation in Section 3.3 where we model the potential outcomes by a count type distribution. Section 4 gives preliminary results on our matching and causal estimands and we conclude with remarks and discussion.

2. Data

In order to assess the proposed methodology, we carry out a case study analyzing a real data set which was based on the NMMAPS dataset for 19872000 obtained from the National Center for Health Statistics (iHAPSS, 2011). The dataset included information on air pollution, mortality and weather conditions for 108 large communities distributed across the USA. In order to achieve the present study aims, we selected data from Chicago.

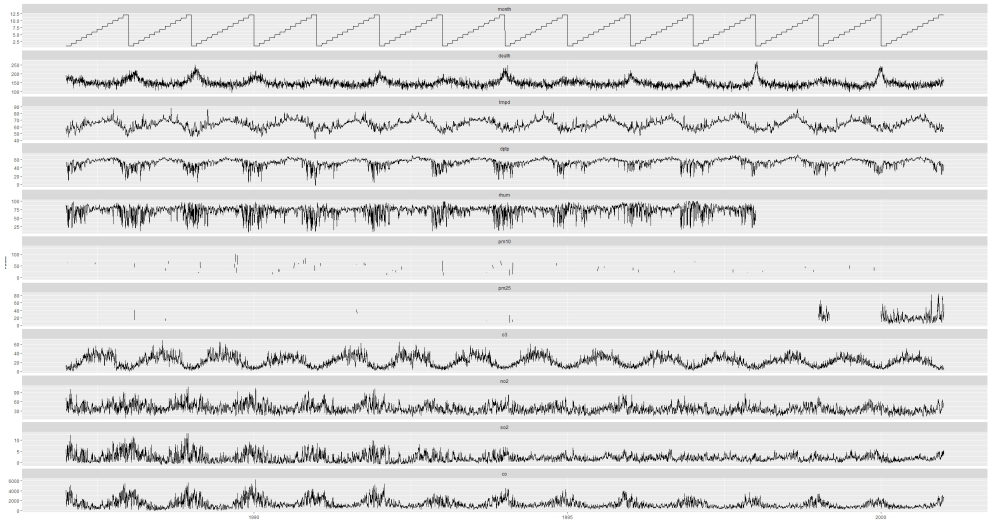

Figure 1. Times series of each of the original dataset

Air pollution data for the daily 8 hour maximum ozone concentration in each community was provided to the NMMAPS database by the USA Environmental Protection Agency (EPA) Aerometric Information Retrieval Service

(now called the Air Quality System database). We chose daily 8 h maximum concentration as the ozone concentration indicator because previous studies indicated that it has a stronger association with health outcomes compared with other metrics, and appears to be a more appropriate metric for investigating health effects of ambient ozone exposure. $\mathbf{PM}_{10}$, $\mathbf{PM}_{2.5}$, $\mathbf{O}_3$, $\mathbf{NO}_2$, $\mathbf{SO}_2$ and $\mathbf{CO}_2$ were the other covariates included in the study.

Daily meteorological data for each community was provided to the NMMAPS database by the National Climatic Data Center. Measurements from multiple weather stations were averaged to provide weather variables representing each community. Meteorological data included daily mean temperature ($\mathbf{TM}, \mathbf{C}$) and relative humidity (RH, %). Other information was obtained from the NMMAPS database. At the outset, there were 5114 days records in Chicago from 1987/01/01 to 200/12/31. The collected variables were $\mathbf{PM}_{10}$, $\mathbf{PM}_{2.5}$, $\mathbf{O}_3$, $\mathbf{NO}_2$, $\mathbf{SO}_2$ and $\mathbf{CO}_2$. $\mathbf{PM}_{2.5}$ and relative humidity were excluded because they had 85% and 21% missing values respectively. The average, first and third quartiles of the variables were: $\mathbf{O}_3$ 19.9 ppb $(11.8 - 26.4)$, $\mathbf{PM}_{10}$ 33.6 $\mu g/m^3$ (20.6-42.6), $\mathbf{PM}_{2.5}$ 17.4 $\mu g/m^3$ (10.8-22.6), $\mathbf{NO}_2$ 25.5 $\mu g/m^3$ (19.9-30.4), $\mathbf{SO}_2$ 5.2 $\mu g/m^3$ (2.9-6.8), $\mathbf{CO}_2$ 802 $\mu g/m^3$ (603-945), temperature 50 degrees F (35-67) and relative humidity 69.2 (58.1-80.4).

3. Methods

This section presents the four stages of the causal pipeline that we use to construct plausible hypothetical randomized to study the effects of ozone has on mortality. These four steps entail:

Variable	Description
date	the date in `YYYY-MM-DD` format
$\mathbf{O}_3$	daily ozone time series, measured in parts per billion (ppb)
death	daily mortality counts from all causes excluding accidents
$\mathbf{PM}_{10}$	daily $\mathbf{PM}_{10}$ time series, measured in $\mu g/m^3$
$\mathbf{PM}_{2.5}$	daily $\mathbf{PM}_{2.5}$ time series, measured in $\mu g/m^3$
$\mathbf{NO}_2$	daily $\mathbf{NO}_2$ time series, measured in $\mu g/m^3$
$\mathbf{SO}_2$	daily $\mathbf{SO}_2$ time series, measured in $\mu g/m^3$
$\mathbf{CO}_2$	daily $\mathbf{CO}_2$ time series, measured in $\mu g/m^3$
Temperature	daily average temperature (in degrees F)
Relative humidity	daily average humidity (in %)

Table 1. Definition of the variables of the dataset

3.1. **Stage 1: The conceptual stage.** The causal question of interest is whether an increase of ozone level cause an increase in the mortality in inhabitants population of study. In order to estimate the causality effect, we firstly need a matched dataset with treated and control days. We aim to develop a pipeline to match, exposed days to a certain exposure (which would be the treatment days) with non-exposed days (which would be the control days). In the matching phase, the outcome should be out of the process. Once days are matched, the outcome variable can be included and the effect of the exposure can be estimated. Let's consider the next structure: The observed data set has N ordered days defined by the date. It is complemented by a set of covariates, $X_i = X_{ij}, j = 1, ..., J$ where J is the number of covariates and $i = 1, ..., N$ is the position of the day. The dataset has also an outcome vector, Y_i. Inside the covariate matrix we distinguish the main variable (M_i) from the others. The variables month, year and day of the week are extracted from the *date* variable. To estimate the previous days effect, several lags are calculated for each variable of interest.

Once the structure was defined, we followed the next steps to describe the hypothetical intervention experiment. First, we select the location where we want to carry out the intervention experiment. The intervention entails an increment of the concentration of the main exposure to a value higher than a previously defined threshold, T. The day of the intervention is considered a treated day. and each of these days needs to be matched with a control day (concentration of the main exposure lower than T). To reduce any kind of bias, each pair has to happen in the same year, month and day of the week. Furthermore, the main exposure and the other covariates have to keep constant in the previous days. This number of days are defined previously as number of lags. In each month of each year, we select randomly how many interventions we will do: 0, 1 or 2. In the case of 0, nothing will be done. In the case of 1, we will select randomly the day of the week and two separate weeks of the month. Then, we will randomly decide if the treatment or the control will happen first. If the treatment is selected, the main exposure, M_i, will be raised higher than the predefined threshold, T. If the control is selected first, we will keep M_i constant. In the case of 2 interventions per month, we will follow the previous steps twice. We iterate the process from the first date until the end. Finally, we extract records from the dataset that coincide with the treated and controls days (about the outcome and the other variables). As we consider that the effect can influence also in the next days, outcome records from the following F days were also collected.

3.2. **Stage 2: The design stage.** In this stage, the aim is to obtain a balance subset of the observed data which the assignment to exposure is unconfounded, i.e., the exposure assignment is independnet of the potential outcomes given the pre-exposure covariates. Although the balance of time-independent covariates is important to consider, balance in background time-varying covariates is challenging to achieve in time series setting. We denote this task by 'historical time series matching'. To our experience, standard matched-sampling strategies (e.g., using estimated propensity score) are not successful at that task, nevertheless balancing all background covariates is essential to assess causality. For example, if one is interested in the causal effect of high (vs. low) O_3 exposure on a mental health outcome (e.g., number of psychiatric hospital admissions occurring the next day), one would desire to compare two similar days, similar with respect to other background covariates (e.g., temperature and $PM_{2.5}$), as shown in Figure 2.

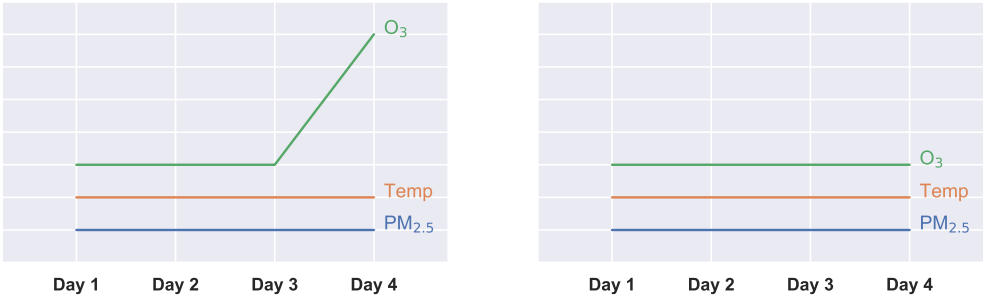

FIGURE 2. Example of historical matching (O_3: ozone, Temp: temperature).

We have developed a constrained pair matching algorithm using a maximum bipartite matching such that: (i) there is one node per unit (i.e., time series), partitioned into treated nodes and control nodes, and (ii) the edges are pairs of treated and control nodes with covariates $X_{i,t}$ and $X_{i',t'}$, and (iii) an edge exists if and only if $\Delta(X_{i,t}, X_{i',t'}) < \infty$. In order to ensure covariate balance, we allow a treated unit to be matched with a control unit if the component wise distances between their respective covaraite vectors are less than some η. For any pair of covariate vectors $X_{i,t}$ and $X_{i',t'}$, we define the L_1 difference between them as follows:

$$\Delta(X_{i,t}, X_{i',t'}) = 0 \text{ if } |X_{i,t}^{(k)} - X_{i',t'}^{(k)}| < \eta \quad \text{for some } \eta > 0. \tag{3.1}$$

uniformly over all k. By construction, using a maximum bipartite matching algorithm on this graph as implemented in the R package *igraph* produces

the largest set of matched pairs that satisfy the unit-specific proximity constraints set by our thresholds. Diagnostic tools can then ensure balance in time-dependent, and more importantly, time-varying, covariates between the treated and control time series. Given the available covariates, one can mimic a randomized intervention and provide transparent diagnostics about the success of creating comparable groups of polluted vs. less polluted days.

3.3. **Stage 3: Analysis of the hypothetical experiment.** We consider a dichotomous treatment, that is, for any given unit in a study population, we can either treat that unit or not. We assume that there is only one version of treatment, and that particular treatment of one unit does not impact another (Cox, 1958). If we treat the unit, we observe the outcome denoted by $Y_i(1)$. If not, we observe $Y_i(0)$. This pair of values, $(Y_i(0), Y_i(1))$, are the i^{th} unit potential outcomes. In the context of causal inference, we consider the causal effect of binary treatment W, with $W = 1$ indicating assignment to treatment and $W = 0$ indicating assignment to control. Following convention in the literature (confer e.g., Rubin (1978)), the causal effect for individual i, with $i = 1, \ldots, N$ is defined as a comparison of potential outcomes, $Y_i(1)$ and $Y_i(0)$, where these are the outcomes that would be observed under the $W = 1$ and $W = 0$, respectively. We can never observe both $Y_i(0)$ and $Y_i(1)$ for any unit i, because it is not possible to go back in time and expose the i^{th} unit to the other treatment, $Y_i^{obs} = Y_i(1) \cdot W_i + Y_i(0) \cdot (1 - W_i)$. This is why causal inference is a missing data problem (Holland, 1986). Because of this missing data problem, researchers typically rely on Neymanian (Splawa-Neyman et al., 1990) and Bayesian (Rubin, 1978) approaches to estimate the average of the unit-level treatment effect, i.e., the Average Treatment Effect (ATE):

$$\tau = \frac{1}{N} \sum_{i=1}^{N} \left(Y_i(1) - Y_i(0) \right). \tag{3.2}$$

Posterior calculations of the causal estimands. To draw inferences on causal estimands, we need to derive the posterior process for this quantity. Here we highlight the steps in a Bayesian setting that is needed to derive the distribution of the causal estimand: **Step 1** : Using the following identity,

$$Y_i(0) = Y_i^{\text{mis}} \mathbf{1}_{(W_i=1)} + Y_i^{\text{obs}} \mathbf{1}_{(W_i=0)}, \text{ and } Y_i(1) = Y_i^{\text{mis}} \mathbf{1}_{(W_i=0)} + Y_i^{\text{obs}} \mathbf{1}_{(W_i=1)} \tag{3.3}$$

where W_i is the assignment mechanism for unit i (Rubin, 1978), the following conditional distribution of the missing potential outcomes, given the observed potential outcomes is derived by marginalizing the parameters: i.e.

$Y^{\text{mis}}|Y^{\text{obs}}, \theta_k$, where θ_k here denotes the parameters entwined with kernel k. **Step 2** : This step involves an intermediate computation of the posterior distribution of θ_k. We then proceed to **Step 3** : combine steps of **1** and **2** to arrive at our goal, i.e. which is the posterior distribution of the estimand, thus completing the final stage, **Step 4**. We are now equipped with the tools to evaluate causal estimands of interest, for example, the quantity defined in equation (3.2).

Sampling procedure. Owing to the complexity of the model, it is may not be easy to sample from the conditional distribution $Y^{\text{mis}}|Y^{\text{obs}}, \theta_k$. Hence, we resort to standard MCMC inference. As we explicated above, our new Bayesian framework has become a crucial component of drawing inferences. It permits us to systematically reason about parameter uncertainty. Under our proposed model where the joint potential outcomes follow a Poisson law Ψ, with $\vartheta \equiv (\beta_c, \beta_t)$, the posterior predictive distribution of $\mathbf{Y}^{\text{mis}}$ takes the form

$$\mathbb{P}(\mathbf{Y}^{mis} \,|\, \mathbf{Y}^{obs}, \mathbf{W}, \beta_c, \beta_t) = \prod_{i=1}^{N} \Psi(Y_i^{mis} \,|\, e^{x_i^\top [W_i\beta_c + (1-W_i)\beta_t]})$$

which can also be interpreted as the conditional distribution of the missing potential outcomes given the observed values. In this step we combine the prior distribution of ϑ, i.e. $\mathbb{P}(\vartheta)$ with the distribution of the observed data given ϑ to compute the posterior distribution of ϑ, i.e. $\mathbb{P}(\vartheta \,|\, \mathbf{Y}^{\text{obs}}, \mathbf{W})$. Using the quantities of the likelihood function $\mathcal{L}(\vartheta \,|\, \cdot\,)$ with the prior distribution $\mathbb{P}(\vartheta)$, we get:

$$\mathbb{P}(\vartheta \,|\, \mathbf{Y}^{\text{obs}}, \mathbf{W}) = \frac{\mathbb{P}(\vartheta) \cdot \mathcal{L}(\vartheta \,|\, \mathbf{Y}^{\text{obs}}, \mathbf{W})}{\mathbb{P}(\mathbf{Y}^{\text{obs}}, \mathbf{W})}.$$

The repestive quantities of likelihood $\mathcal{L}(\vartheta \,|\, \cdot\,)$ together with the so-called marginal likelihood $\mathbb{P}(\mathbf{Y}^{\text{obs}}, \mathbf{W})$ can be computed seperately as follows:

$$\mathcal{L}(\vartheta \,|\, \mathbf{Y}^{\text{obs}}, \mathbf{W}) = \mathbb{P}(\mathbf{Y}^{\text{obs}}, \mathbf{W} \,|\, \vartheta),$$

as well as $\mathbb{P}(\mathbf{Y}^{\text{obs}}, \mathbf{W}) = \int_\vartheta \mathbb{P}(\vartheta) \cdot \mathcal{L}(\vartheta \,|\, \mathbf{Y}^{\text{obs}}, \mathbf{W}) d\vartheta$. Concretely, we then have the posterior for ϑ as follows

$$\mathbb{P}(\beta \,|\, \mathbf{Y}^{obs}, \mathbf{W}) = \mathcal{N}\left(\beta \,|\, 0, 10000\mathbf{I}_{2k}\right) \prod_{i=1}^{N} \Psi\left(Y_i^{obs} \,|\, e^{x_i^\top [(1-W_i)\beta_c + W_i\beta_t]}\right).$$

To sample β from its posterior $\mathbb{P}(\beta \,|\, \mathbf{Y}^{obs}, \mathbf{W})$, we use Metropolis-Hastings algorithm with the acceptance probability $\mathcal{AP} = \min\{1, \mathcal{A}(\beta^{new})\}$ where

$$\begin{aligned}
&\mathcal{A}(\beta^{new}) \\
&= \exp\Bigg\{ -\frac{1}{2}(\beta^{new})^\top \beta^{new} + \sum_{i=1}^{N} \Big[Y_i^{obs} \left(x_i^\top [(1-W_i)\beta_c^{new} + W_i \beta_t^{new}] \right) \\
&\qquad -e^{x_i^\top [(1-W_i)\beta_c^{new} + W_i \beta_t^{new}]} \Big] + \frac{1}{2}(\beta^{old})^\top \beta^{old} \\
&\quad - \sum_{i=1}^{N} \left[Y_i^{obs} \left(x_i^\top [(1-W_i)\beta_c^{old} + W_i \beta_t^{old}] \right) - e^{x_i^\top [(1-W_i)\beta_c^{old} + W_i \beta_t^{old}]} \right] \Bigg\}
\end{aligned}$$

The next step entails the computation of the distribution of $\mathbb{P}(Y_i^{mis} \,|\, \mathbf{Y}^{obs}, \mathbf{W})$. The expression is readily computed as follows:

$$\begin{aligned}
\mathbb{P}(Y_i^{mis} \,|\, \mathbf{Y}^{obs}, \mathbf{W}) &= \int \mathbb{P}(Y_i^{mis} \,|\, \mathbf{Y}^{obs}, \mathbf{W}, \beta)\, \mathbb{P}(\beta \,|\, \mathbf{Y}^{obs}, \mathbf{W})\, d\beta \\
&\approx L^{-1} \sum_{\ell=1}^{L} \Psi\left(Y_i^{obs} \,|\, e^{x_i^\top \{W_i \beta_c[\ell] + (1-W_i)\beta_t[\ell]\}} \right)
\end{aligned}$$

where $\beta[\ell] = [\beta_c[\ell]^\top, \beta_t[\ell]^\top]^\top$ are sampled from $\mathbb{P}(\beta \,|\, \mathbf{Y}^{obs}, \mathbf{W})$ as in Step 2. Above equation indicates that $\mathbb{P}(Y_i^{mis} \,|\, \mathbf{Y}^{obs}, \mathbf{W})$ can be approximated by a mixture of Poisson with the weights of mixtures are equals. In order to sample Y_i^{mis}, we first draw L samples of β and then pick one sample of β to plug in the above Poisson distribution to draw Y_i^{mis}, thus giving the causal effect:

$$\tau_{fs} = \frac{1}{N} \sum_{i=1}^{N} (Y_i(1) - Y_i(0)) = \frac{1}{N} \sum_{i=1}^{N} ((2W_i - 1)Y_i^{obs} + (1 - 2W_i)Y_i^{mis}). \quad (3.4)$$

4. Discussion and concluding remarks

We first assess the new potential outcome framework using Poisson random variable in a controlled setting. Our estimation framework is first tested on synthetic data generated by our Poisson model Ψ with $Y_i(0) = \Psi(\exp(0.2 + 1.7x_i))$ and $Y_i(1) = \Psi(\exp(c + 1.6x_i))$. For each of these pair, we record the number of events where these can be number of relapses or hospital visits, say. The value c is chosen such that the true average treatement effect is of value 20. Figure 3 shows two paths of the posterior mean of τ_{fs} converging to the true value and producing commensurate accuracy to the ground truth, which

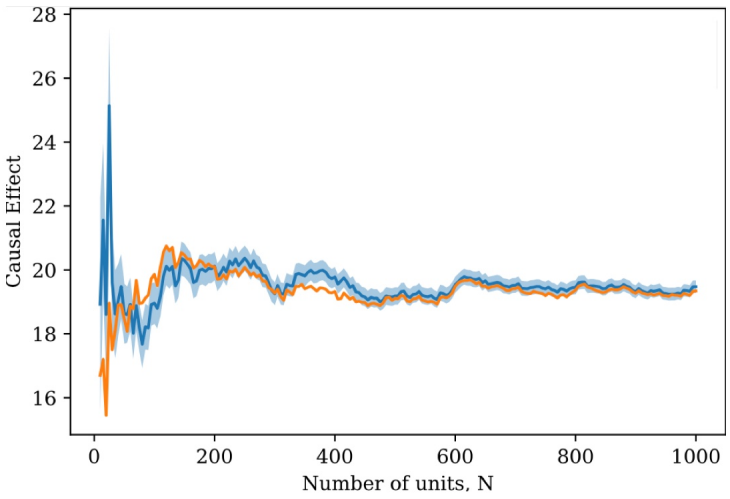

FIGURE 3. Evolution of posterior mean as N increases for model.

reassures that our inference procedures operate sensibly and recover the true parameters asymptotically when we increase the number of MCMC iterations.

With respect to the running example on the real data set, we have implemented the the matching strategy under on the US National Morbidity, Mortality, and Air Pollution Study (NMMAPS) conducted between 1987 and 2000 (iHAPSS, 2011). The dataset included daily concentrations on air pollutants, mortality and meteorological variables for 108 large communities distributed across the USA. The threshold to construct days with high and low O_3 concentration was *a priori* defined based on the current legal limit for ozone by the US EPA. We see that, if the covariates are not balanced, it is non-sensical to perform a causal inference mechanism with respect to our Poisson model.

References

Bind, M.-A. C. and Rubin, D. B. (2017). Bridging observational studies and randomized experiments by embedding the former in the latter. *Statistical methods in medical research*, page 0962280217740609.

Cox, D. R. (1958). Planning of experiments.

Guo, Y., Gasparrini, A., Li, S., Sera, F., Vicedo-Cabrera, A. M., de Sousa Zanotti Stagliorio Coelho, M., Saldiva, P. H. N., Lavigne, E., Tawatsupa, B., Punnasiri, K., Overcenco, A., Correa, P. M., Ortega, N. V., Kan, H., Osorio, S., Jaakkola, J. J. K., Ryti, N. R. I., Goodman, P. G., Zeka, A., Michelozzi, P., Scortichini, M., Hashizume, M., Honda, Y., Seposo, X., Kim, H., Tobias, A., Íñiguez, C., Forsberg, B., Åström, D. O., Guo, Y. L., Chen, B.-Y., Zanobetti, A., Schwartz, J., Dang, T. N., Van, D. D., Bell, M. L., Armstrong, B., Ebi, K. L., and Tong, S. (2018). Quantifying excess deaths related to heatwaves under climate change scenarios: A multicountry time series modelling study. *PLOS Medicine*, 15(7):e1002629.

Haavelmo, T. (1944). The probability approach in econometrics. *Econometrica*, (12):1115.

Holland, P. W. (1986). Statistics and causal inference. *Journal of the American Statistical Association*, 81(396):945–960.

iHAPSS (2011 (accessed on 1 December 2011)). *Internet-Based Health & Air Pollution Surveillance System (iHAPSS) Mortality, Air Pollution, and Meteorological Data for 108 U.S. Cities 19872000*. http://www.ihapss.jhsph.edu/data/NMMAPS/Data/data_download_gz.htm.

Koopmans, T. C. (1950). Statistical Inference in Dynamic Economic Models. *Cowles Commission*, (10).

Rosenbaum, P. R. and Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1):41–55.

Rubin, D. (1978). *Multiple Imputation for nonresponse in surveys*. Wiley.

Schwartz, J., Bind, M.-A., and Koutrakis, P. (2017). Estimating Causal Effects of Local Air Pollution on Daily Deaths: Effect of Low Levels. *Environmental Health Perspectives*, 125(1):23–29.

Splawa-Neyman, J., Dabrowska, D. M., and Speed, T. P. (1990). On the application of probability theory to agricultural experiments. essay on principles. section 9. *Statistical Science*, 5(4):465–472.

Stuart, E. A. (2010). Matching methods for causal inference: A review and a look forward. *Statist. Sci.*, 25(1):1–21.

Zhang, B.-G., Li, W., Shi, Y., Liu, X., and Chen, L. (2017). Detecting causality from short time-series data based on prediction of topologically equivalent attractors. *BMC systems biology*, 11(Suppl 7):128.