Mediation Analysis Synthetic Control

Giovanni Mellace*, and Alessandra Pasquini**

*University of Southern Denmark; **Sapienza University of Rome

February 2, 2018

Abstract: The Synthetic Control Method (SCM) allows estimating the casual effect of an intervention even when data on only one treated and few control units are available. This is possible by using information on the pre-intervention period to construct a "synthetic control" which mimic what would have happened to the treated unit in the post-intervention period in the absence of the intervention. In mediation analysis, the total effect of a treatment is decomposed into the direct effect of the treatment and its indirect effects which go through intermediate variables that lie in the casual pathway between the treatment and the outcome of interest (mediators). We introduce Mediation Analysis Synthetic Control (MASC) a generalization of SCM that allows decomposing the total effect of the intervention into its indirect effects which go through observed mediators and its direct effect. An extensive simulation study and an empirical application are currently in progress.

Keywords: MASC, Synthetic Control Method, Mediation Analysis, Causal channels, Causal Mechanisms, Direct and indirect effects.

JEL classification: C21, C23, C31, C33.

We have benefited from comments by... . Addresses for correspondence: Giovanni Mellace (giome@sam.sdu.dk), and Alessandra Pasquini (alessandra.pasquini@uniroma1.it).

1 Introduction

The Synthetic Control Method (SCM) introduced by Abadie and Gardeazabal (2003) and further developed in Abadie et al. (2010) and Abadie et al. (2015) is becoming very popular in program evaluation. SCM is attractive as it allows estimating the casual effect of an intervention even when data on only one treated and few control units are available. This is possible by using information on the pre-intervention period to construct a "synthetic control" which mimic what would have happened to the treated unit in the post-intervention period in the absence of the intervention. Gobillon and Magnac (2016) compare SMC to other interactive fixed-effects models based find that SCM performs very well as soon as in the post-intervention periods the counterfactual outcome of the treated unit lies in the convex hull of the outcomes of the control units. In a recent paper, Xu (2017) further exploits the connection between SCM and interactive fixed-effect models and proposed a new method which combines both approaches. Doudchenko and Imbens (2016) prose a modification of SCM where the weights are not contained to be positive and do not necessarily add up to 1. Athey et al. (2017) propose a new method which includes synthetic control and other panel data method as a special case.

Although all those methods are very well suited to estimate the total effect of an intervention, they do not take into account the fact that might exist some intermediate variables (hereafter referred to as "mediators") that lie in the causal path between the treatment and the outcome of interest, such that the total effect might be viewed as a sum of the different channels through which the treatment affects the outcome and may differ from the effects of interest. In the presence of one or more mediators, the total effect can be generally decomposed into a direct or net effect of the treatment on the outcome of interest and some indirect effects which are generated through the mediators. Policy conclusions that ignore the presence of such intermediate outcomes, might be misleading. For example, one may find a zero or even negative total effect of the intervention, even though its direct effect is positive. Moreover it is often important to quantify the indirect effects to better target the intervention. Consider the huge decrease in tobacco consumption after the introduction California anti-tobacco law, Proposition 99, estimated in Abadie et al. (2010). Proposition 99 not only increased tobacco price but also introduced several anti-tobacco informational campaigns. It would be extremely relevant for a policy maker to know how much of the decrease in tobacco consumption triggered by Proposition 99 is due to the increase in price and how much due to investment in informational campaigns. Both effect can be thought as indirect effects. This is the core of Mediation Analysis and several papers show how to identify and estimates direct and indirect effect under sequential conditional independence (see Pearl (2001), Robins (2003), Imai et al. (2010), Imai and Yamamoto (2013), Vansteelandt and VanderWeele (2012), Huber (2013), Vansteelandt and VanderWeele (2012), Huber et al. (2016), Huber et al. (2017)) both to the best of our knowledge none of the existing method are specifically designed for panel data (one can still control on pre-intervention outcomes in those methods) and cannot directly be applied in setting with one or few treated and few control units.

We introduce Mediation Analysis Synthetic Control (MASC) a generalization of SCM that allows decomposing the total effect of the intervention into its indirect effects which go through observed mediators and its direct effect. For example, as we will discuss in more details MASC re-weights the control unit post-intervention outcomes by choosing weights that minimize the distance between treated and control in pre-intervention observable characteristics (including pre-intervention values of the outcome and the mediator) as well as post-intervention values of the mediator to mimic what would have happened to the treated in absence of the intervention if her mediator were set to her potential mediator under treatment. Following Abadie et al. (2010) we illustrate MASC with a simple dynamic factor model with interactive fixed-effects and show that both the direct and the indirect effects estimators are unbiased as the number of pre-intervention periods goes to infinity.

The rest of the paper is organized as follows: Section 2 introduces MASC; Section 3 propose possible inference procedures; Section 4 contain an extensive simulation study; Section 5 includes an empirical application to ...; and Section 6 concludes. All the technical proofs are relegated to the appendix.

2 The Mediation Analysis Synthetic Control Method

Assume we are interested in the effect of an intervention D, implemented from a time T, on an outcome Y. Suppose that part of the effect the effect of D on Y goes through a mediator M. Then, the total effect of the intervention can be decomposed into an indirect effect which goes through M and the residual effect, commonly called direct effect, which goes through other casual pathways. Although arguably extremely policy relevant identification of the direct and indirect effects of D is challenging. To see this let D_{it}, a binary indicator which is equal to one if unit i is affected by the intervention at time t. We will refer as "treated" to all the units that are exposed to the intervention at some point in time. For each unit i we can defined the potential mediator at time t as follow:

$$M_{it}(d) \text{ for d } \in \{0, 1\}.$$

$M_{it}(d)$ is the value that the mediator of unit i would take, at time t, if D_{it} were set to d. Assuming that there are no anticipation effects on the mediator in the pre-intervention period and that the standard Stable Unit Treatment Value Assumption (SUTVA) holds, the observed and the potential mediators are related through the following observation rule:

$$M_{it} = M_{it}(0)(1 - D_{it}) + M_{it}(1)D_{it}.$$

Note that M_{it} is always equal to $M_{it}(0)$ for both treated and control units in the pre-intervention ($t < T$) periods and we can only observe one of the two potential mediators in the post-intervention period ($t \geq T$). Similarly for each unit i at time t, we define the potential outcome as follow:

$$Y_{it}(d, M_{it}(d')) \equiv Y_{it}^{d,d'} \text{ for } d, d' \in \{0, 1\}.$$

$Y_{it}^{d,d'}$ is the value that the outcome of unit i would take at time t if we set $D_{it} = d$ and $M_{it} = M_{it}(d')$. The potential outcomes are functions of both the treatment and the potential mediator. Under SUTVA and assuming no anticipation effects in the pre-intervention periods, the observed and the potential outcomes are related by the following observation rule:

$$Y_{it} = Y_{it}^{0,0}(1 - D_{it}) + Y_{it}^{1,1} D_{it}.$$

Differently from the standard setting only $Y_{it}^{0,0}$ and $Y_{it}^{1,1}$ are observed (and still only one of them for each unit in each period), while $Y_{it}^{0,1}$, $Y_{it}^{1,0}$ are never observed. This make the identification of the direct and indirect effect more challenging.

Our parameters of interest are the total effect of the intervention as well as the direct and the indirect effects. Following the synthetic control literature, we will first define these parameters with respect to a single treated unit. This is in contrast with standard policy evaluation literature, where the total, the direct and the indirect effects are defined as averages, either with respect to the whole sample (Pearl (2001), Robins (2003), Imai et al. (2010), Imai and Yamamoto (2013), Vansteelandt and VanderWeele (2012), Huber (2013)) or with respect to the treated (Vansteelandt and VanderWeele (2012), Huber et al. (2017)). If more then one unit is exposed to the intervention (see Gobillon and Magnac (2016) and Adhikari (2015)) our method can be also used to decompose the Average Treatment Effect on the Treated (Vansteelandt and VanderWeele (2012), Huber et al. (2017)).

Without loss of generality, assume that we observe J units ordered such that units 1 to n are treated while units $n+1$ to J are control units. Consider the first treated, unit 1, the effects of interest for unit 1 are the total effect α_{1t}, the direct effects $\theta_{1t}(M_{1t}(d))$, and the indirect effects $\delta_{1t}(d)$ in the post-intervention period ($t \geq T$) that are defined as:

$$\begin{aligned}
\alpha_{1t} &= Y_{1t}^{1,1} - Y_{1t}^{0,0}, \\
\theta_{1t}(M_{1t}(d)) &= Y_{1t}^{1,d} - Y_{1t}^{0,d}, \\
\delta_{1t}(d) &= Y_{1t}^{d,1} - Y_{1t}^{d,0}.
\end{aligned}$$

Notice that

$$\begin{aligned}
\alpha_{1t} &= Y_{1t}^{1,1} - Y_{1t}^{0,0}, \\
&= Y_{1t}^{1,1} - Y_{1t}^{1,0} + Y_{1t}^{1,0} - Y_{1t}^{0,0}, \\
&= \delta_{1t}(1) + \theta_{1t}(M_{1t}(0)),
\end{aligned}$$

$$\begin{aligned}
\alpha_{1t} &= Y_{1t}^{1,1} - Y_{1t}^{0,0}, \\
&= Y_{1t}^{1,1} - Y_{1t}^{0,1} + Y_{1t}^{0,1} - Y_{1t}^{0,0}, \\
&= \theta_{1t}(M_{1t}(1)) + \delta(0).
\end{aligned}$$

If α_{1t} is identified, identifying $\theta_{1t}(M_{1t}(1))$ and $\delta_{1t}(1)$ automatically implies identification of $\theta_{1t}(M_{1t}(0)) = \alpha_{1t} - \delta_{1t}(1)$ and $\delta_{it}(0) = \alpha_{1t} - \theta_{1t}(M_{1t}(1))$. Thus, we will focus only on those two parameters hereafter. Intuitively, the total effect α_{1t} can be identified with the synthetic control method as described in Abadie et al. (2010) and briefly summarized further below. The main contribution of this paper is to show that the same idea can be applied to identify the direct effects $\theta_{1t}(M_t(1))$ and in the presence of more treated units the indirect effects $\delta_{1t}(0)$. To this end we need to show that $Y_{1t}^{1,1}$, $Y_{1t}^{0,0}$, $Y_{1t}^{1,0}$ and $Y_{1t}^{0,1}$ are identified in the post-intervention period. First notice that $Y_{1t}^{1,1}$ is observed in the post-intervention period for unit 1 and $Y_{1t}^{0,0}$ can be estimated using the standard SCM. Our main challenge is that both $Y_{1t}^{1,0}$ and $Y_{1t}^{0,1}$ are never observed and cannot be estimated through a standard SCM.

The idea behind standard SCM is to use a linear combination of the control units to build a "synthetic control" that mimics what would have happened to the treated unit in the post intervention period in the absence of the intervention ($t \geq T$). This is done by re-weighting the control units post treatment outcomes with weights that are chosen to minimize the distance between treated and control units pre-intervention observable characteristics (including pre-intervention outcomes). In other words SCM creates and estimated value of $Y_{1t}^{0,0}$ in the post-intervention periods. The main assumption is that $Y_{1t}^{0,0}$ lies in the convex hull of the non-treated post-intervention outcomes, namely can be written as a liner combination of the latter.

We now introduce Mediation Analysis Synthetic Control (MASC) that generalizes uses SCM's idea to create "synthetic" values of $Y_{1t}^{1,0}$ and $Y_{1t}^{0,1}$ in the post intervention periods. For $Y_{1t}^{0,1}$ we propose to re-weight the control unit post-intervention outcomes by choosing weights that minimize the distance between treated and control pre-intervention observable characteristics as well as post-intervention values of the mediator. The intuition is that since $M_{1t} = M_{1t}(1)$ in the post-intervention period, including the distance between treated and control with respect to post treatment values of the mediator in choosing the weights will mimic what would have happened to the treated in absence of the intervention if her mediator were set to her potential mediator under treatment $M_{1t}(1)$. Finding a "synthetic" value of $Y_{1t}^{1,0}$ is more challenging and requires more than 1 treated unit. First we need estimate what value the mediator of unit 1 would have taken in the absence of the intervention ($M_{1t}(0)$). This could be done with a standard SCM using the mediator as an outcome. Second we propose to treat the reaming treated as a control in a SCM where we use also the distance between the first step estimate of $M_{1t}(0)$ and the other treated mediators in computing the weights.

In the spirit of Abadie et al. (2010) to further illustrate our approach we will introduce a factor model in which we assume that potential mediators of unit i are given by

$$
\begin{aligned}
M_{it}(0) &= \gamma_t + \beta_t Z_i + \vartheta_t \varrho_i + \nu_{it}, \\
M_{it}(1) &= \gamma_t + \beta_t Z_i + \vartheta_t \varrho_i + \psi_t D_{it} + \nu_{it},
\end{aligned}
$$

where γ_t is an unknown common factor with constant factor loadings across units. Z_i is a $(p \times 1)$ vector of observed covariates, β_t is a $(1 \times p)$ vector of unknown parameters, ϑ_t is a $(1 \times v)$ vector of unobserved common factors, ϱ_i is an $(v \times 1)$ vector of unknown factor loadings, ψ_{it} is an unknown parameter describing the impact of the treatment on the mediator, and ν_{it} are unobserved transitory shocks.
Similarly, we assume that the four potential outcomes are given by

$$\begin{aligned}
Y_{it}^{0,0} &= \zeta_t + \eta_t X_i + \lambda_t \mu_i + \varphi_t(0) M_{it}(0) + \epsilon_{it}, \\
Y_{it}^{0,1} &= \zeta_t + \eta_t X_i + \lambda_t \mu_i + \varphi_t(0) M_{it}(1) + \epsilon_{it}, \\
Y_{it}^{1,0} &= \zeta_t + \eta_t X_i + \lambda_t \mu_i + \varphi_t(1) M_{it}(0) + \rho_t(M_{it}(0)) D_{it} + \epsilon_{it}, \\
Y_{it}^{1,1} &= \zeta_t + \eta_t X_i + \lambda_t \mu_i + \varphi_t(1) M_{it}(1) + \rho_t(M_{it}(1)) D_{it} + \epsilon_{it},
\end{aligned}$$

where ζ_t is an unknown common factor with constant factor loadings across units, X_i is a $(r \times 1)$ vector of observed covariates which might be different from Z_i, η_t is a $(1 \times r)$ vector of unknown parameters, λ_t is a $(1 \times F)$ vector of unobserved common factors, μ_i is an $(F \times 1)$ vector of unknown factor loadings, ϵ_{it} are unobserved transitory shocks, and $\varphi_{it}(d)$ and $\rho_{it}(M_{it}(d))$ capture the impact on the potential outcomes of the potential mediator and the treatment respectively. In this model the total, direct and indirect effects of unit 1 are then given by

$$\begin{aligned}
\alpha_{1t} &= \varphi_t(1) M_{it}(1) - \varphi_t(0) M_{it}(0) + \rho_t(M_{it}(1)), \\
\theta_{1t}(M_{1t}(1)) &= \rho_t(M_{1t}(1)) + (\varphi_t(1) - \varphi_t(0)) * M_{1t}(1), \\
\theta_{1t}(M_{1t}(0)) &= \rho_t(M_{1t}(0)) + (\varphi_t(1) - \varphi_t(0)) * M_{1t}(0), \\
\delta_{1t}(1) &= \rho_t(M_{1t}(1)) - \rho_t(M_{1t}(0)) + \varphi_t(1) * (M_{1t}(1) - M_{1t}(0)), \\
\delta_{1t}(0) &= \varphi_t(0) * (M_{1t}(1) - M_{1t}(0)).
\end{aligned}$$

As mentioned above, for the total effect we can just use the standard SCM. In particular, we want to find the $(1 \times (J - n))$ vector of positive and adding up to 1 weights

$L^* = (l^*_{n+1}, ..., l^*_J)$ such that in the post intervention period

$$Y^{0,0}_{1t} = \sum_{i=n+1}^{J} l^*_i Y_{it}.$$

As we shown by Abadie et al. (2015) this is the case if $\forall\ t = 1, ..., T-1$, L^* also satisfies

$$\begin{aligned} \sum_{j=n+1}^{J} l^*_j Y_{jt} &= Y_{1t}, \\ \sum_{j=n+1}^{J} l^*_j M_{jt} &= M_{1t}, \\ \sum_{j=n+1}^{J} l^*_j Z_j &= Z_1. \end{aligned}$$

This justifies choosing the weights that minimize the distance between the observable characteristics of the treated and the one of the control units in the pretreatment period. More formally, let $\Omega^\alpha_1 = (X_1, Y_{11}, \ldots, Y_{1T})$ be a $((T+r)\times 1)$ vector, $\omega^\alpha_{0i} = (X_i, Y_{i1}, \ldots, Y_{iT})$ be a $(1 \times (T+r))$ vector, and $\Omega^\alpha_0 = (\omega^\alpha_{0,n+1}, \ldots, \omega^\alpha_{0J})'$, then

$$\begin{aligned} L^* &= \min_{l_{n+1},\ldots,l_J} ||\Omega^\alpha_1 - L\Omega^\alpha_0|| \\ & s.t. \quad l_{n+1} \leq 0, ..., l_J \leq 0, \sum_{i=n+1}^{J} l_i = 1, \end{aligned}$$

where $||\Omega^\alpha_1 - L\Omega^\alpha_0|| = \sqrt{(\Omega^\alpha_1 - L\Omega^\alpha_0)' (\Omega^\alpha_1 - L\Omega^\alpha_0)}$. It is also possible to give more weight to specific observable characteristics by using the alternative distance $||\Omega^\alpha_1 - L\Omega^\alpha_0||_V = \sqrt{(\Omega^\alpha_1 - L\Omega^\alpha_0)' V (\Omega^\alpha_1 - L\Omega^\alpha_0)}$ (see Abadie et al. (2010) for a data driven procedure to choose V). As the choice of V can only affect efficiency, we follow Gobillon and Magnac (2016) and take V as the identity matrix.

Let $E(\hat{Y}^{0,0}_{1t}) = \sum_{i=n+1}^{J} l^*_i Y_{it}$, Abadie et al. (2010) show that, if L^* exist,

$$E(\hat{Y}^{0,0}_{1t}) = Y^{0,0}_{1t} + o(T)$$

Consequently, estimating the total effect as $\hat{\alpha}_{1t} = Y_{1t} - \hat{Y}_{1t}^{0,0}$ is justified by the fact that

$$\lim_{T\to\infty} \hat{\alpha}_{1t} = \alpha_{1t} \quad \forall \quad t \geq T \tag{2.1}$$

The first step of MASC is estimation of $Y_{1t}^{0,1}$. This requires additional constraints. Indeed, we want to construct a "synthetic" unit which is identical to the treated, not affected by the intervention, and, at the same time, has the same value of the mediator as the treated unit. Similar to standard SCM, we want to find a $(1\times(J-n))$ vector of positive and adding up to 1 weights $W_t^* = (w_{n+1,t}^*, ..., w_{Jt}^*)$ such that in the post intervention period

$$Y_{1t}^{0,1} = \sum_{i=n+1}^{J} w_{it}^* Y_{it}.$$

Notice that by assumption [1] $Y_{1t}^{0,1}$ depends on the value that M takes at time t thus the weights need to be calculated at each post-intervention period. As we show in the appendix W_t^* exists if $\forall\ t = 1, ..., T-1$, it also satisfies

$$\begin{aligned} \sum_{j=n+1}^{J} w_{tj}^* Y_{jt} &= Y_{1t}, \\ \sum_{j=n+1}^{J} w_{tj}^* Z_j &= Z_1, \end{aligned}$$

and $\forall\ t = 1, ..., T-1, \ldots, t$, it also satisfies

$$\sum_{j=n+1}^{J} w_{tj}^* M_{jt} = M_{1t}.$$

The vector of weight W_t^* is then estimated in a similar way as L^*. The only difference is that we now need to include the mediator in the distance. More formally, let $\Omega_1^{\theta_t(1)} = (X_1, Y_{11}, \ldots, Y_{1T}, M_{1,t})$, $\omega_{0i}^{\theta_t(1)} = (X_i, Y_{i1}, \ldots, Y_{iT}, M_{i,t})$, and $\Omega_0^{\theta_t(1)} = (\omega_{n+1}^{\theta_t(1)}, \ldots, \omega_J^{\theta_t(1)})'$,

[1] It is easy to let $Y_{1t}^{0,1}$ depends on all the values that the mediator takes between T and t. This is done by using $\Omega_1^{\theta_t(1)} = (X_1, Y_{11}, \ldots, Y_{1T}, M_{1,T+1}, \ldots, M_{1,t})$ and $\omega_{0i}^{\theta_t(1)} = (X_i, Y_{i1}, \ldots, Y_{iT}, M_{i,T+1}, \ldots, M_{i,t})$

then

$$\begin{aligned} W_t^* &= \min_{w_{n+1,t},\dots,w_{Jt}} ||\Omega_1^{\theta_t(1)} - W_t\Omega_0^{\theta_t(1)}|| \\ & s.t. \quad w_{n+1,t} \leq 0, ..., w_{Jt} \leq 0, \sum_{i=n+1}^{J} w_{it} = 1, \end{aligned}$$

Let $E(\hat{Y}_{1t}^{0,1}) = \sum_{i=n+1}^{J} w_{it}^* Y_{it}$, as we show in the appendix if W_t^* exists

$$E(\hat{Y}_{1t}^{0,1}) = Y_{1t}^{0,1} + o(T)$$

Then, we can estimate the direct effect as $\theta_{1t}(M_{1t}(1))$ and the indirect effect $\delta_{it}(0)$

$$\hat{\theta}_{1t}(M_{1t}(1)) = Y_{1t} - \hat{Y}_{1t}^{0,1}, \qquad \hat{\alpha}_{it} - \hat{\theta}_{1t}(M_{1t}(1)),$$

respectively. Intuitively W^* only exists if a sort of sequential conditional independence assumption holds. In other words, as for SCM, the potential outcomes need to be independent from the intervention once we condition on observable characteristics including pre-intervention outcomes which allow for some selection on unobservables. Additionally, the potential outcomes need to be independent on the potential mediators conditional on the intervention and the observable characteristics.

If there is more than one treated we can also find a "syntetic" $Y_{it}^{1,0}$ this is done in two steps.

In a first step we estimate $M_{1t}(0)$ by $\hat{M}_{1t}(0) = \sum_{i=n+1}^{J} k_{it}^* M_{it}$ with $K_t^* = (k_{n+1,t}^*, \dots, k_{Jt}^*)$ estimated with a standard SCM. Note that also those weight need to be calculated for each post-intervention period t. In a second step we need to find a vector of positive and adding up to 1 weights $Q_t^* = (q_{2t}^*, ..., q_{nt}^*)$ such that $Y_{it}^{1,0} = \sum_{i=2}^{n} q_{it}^* Y_{it}$. Q^* is estimated with a SCM but using only the treated. More specifically, let $\Omega_1^{\delta_t(1)} = (X_1, Y_{11}, \dots, Y_{1T}, \hat{M}_{1t}(0))$,

$\omega_{0i}^{\theta_t(1)} = (X_i, Y_{i1}, \ldots, Y_{iT}, M_{i,t})$, and $\Omega_0^{\theta_t(1)} = (\omega_2^{\theta_t(1)}, \ldots, \omega_n^{\theta_t(1)})'$, then

$$\begin{aligned} Q_t^* &= \min_{q_{n+1,t},\ldots,q_{Jt}} ||\Omega_1^{\theta_t(1)} - Q_t\Omega_0^{\theta_t(1)}|| \\ & \quad s.t. \quad q_{n+1,t} \leq 0, ..., q_{Jt} \leq 0, \sum_{i=n+1}^{J} q_{it} = 1, \end{aligned}$$

Let $E(\hat{Y}_{1t}^{1,0}) = \sum_{i=2}^{n} q_{it}^* Y_{it}$, as we show in the appendix if Q_t^* exists

$$E(\hat{Y}_{1t}^{1,0}) = Y_{1t}^{1,0} + o(T)$$

Then, we can estimate the indirect effect $\delta_{it}(1)$ and the direct effect as $\theta_{1t}(M_{1t}(0))$ as

$$\hat{\delta}_{1t}(1) = Y_{1t} - \hat{Y}_{1t}^{1,0}, \qquad \hat{\alpha}_{it} - \hat{\delta}_{1t}(1),$$

respectively. Intuitively Q^* only exists under the same sequential conditional independence assumption discussed above. However if the number of treated is too small it is very hard to estimate Q^*.

3 Inference

Inference can be carried over in the same as in the standard synthetic control method. Abadie et al. (2015) suggest to run a series of placebo tests consist in estimating the effect (in our case also the direct and indirect ones) of the intervention either pre-intervention periods or for unit not exposed to the intervention. Abadie et al. (2015) criticize the former type of (in-time) placebo tests arguing the other shocks in the past might have affected treated and control unit differently. The latter type of (in-space) placebo consist estimating the effect of the intervention treating untreated units as treated. The synthetic controls obtained are then used to build the root mean squared prediction error (RMSPE) for both pre- and post- intervention periods. The ratio between the post- and pre- RMSPE of the untreated units can then be used to build the distribution of the treatment effects (total, direct and indirect) in absence of treatment. This distribution can be compared with the correspondent ratio of the treated units (for more details see Abadie et al. (2015)).

Another inference procedure is described in Gobillon and Magnac (2016). This procedure is based on different steps. In our framework, the mediator of treated units should be reduced as well, exploiting the estimation of the potential mediator in absence of treatment. Later on, 10'000 samples without replacement of a number of units equal to the number of treated units have to be selected from the group containing all units. For each of the 10'000 samples the selected units should be used as treated units and the rest of the group as control, to apply the mediation synthetic control method. Finally, the estimated values should be used as distributions to make inference on the estimated effects. To be more clear, in our framework, this method would consist in the following steps:

1. Substitute Y_{it} with $Y'_{it} = Y_{it} - \hat{\alpha}_t$ for $i = 1, ..., n$ and $t \geq T$. Where α_t is given by the average among all the total effect estimated.
2. Substitute M_{it} with $M'_{it} = M_{it} - E(M_{it} - \hat{M}_{it}(0))$ for $i = 1, ..., n$ and $t \geq T$.
3. Iterate 10'000 times:
    - Select n units.
    - Apply MASC on selected unit.
    - Calculate the average total, direct and indirect effects
4. Use the calculated effects to determine the distribution of the real effects and do inference.

We refer to Gobillon and Magnac (2016) for more details. Note that both Abadie et al. (2015) and Gobillon and Magnac (2016) inference procedures are based on the strong assumption that the disturbances across units are exchangeable. Indeed, the basic ideas behind these methods is that the noise of the placebos can be used to approximate the noises of the treated units.

In this framework there is a second source of uncertainty. Unfortunately, the choice of the control units (donor pool) can dramatically affect the results. To solve this issue

Abadie et al. (2015) suggest to make a sensitivity test excluding one by one each of the units in the donor pool (if the donor pool is particularly big one can select a sample with replacement from the donor pool). If the estimated effects do not change much the results are sensitive to the chosen donor and can be considered robust.

4 Simulation study

4.1 Data-Generating Process

In the simulation, we considered only one post-intervention period. We follow the simulation study in Gobillon and Magnac (2016) and use the following linear factor model for the mediator:

$$M_{it} = \gamma_t + \vartheta_t \varrho_i + \psi_{it} D_{it} + \nu_{it}, \tag{4.1}$$

where γ_t, ϑ_t and ϱ_i are defined as before and drown from a continuous uniform distribution, with the exception of the first element of ϑ_t which is set to one, to allow for individual fixed effects. The ψs and νs are drown from a normal distribution.

For the outcome is generated generating by the following linear factor model:

$$Y_{it} = \zeta_t + \lambda_t \mu_i + \kappa M_{it} + \phi M_{it} D_{it} + \tau D_{it} + \epsilon_{it} \tag{4.2}$$

Where ζ_t, λ_t and μ_i are defined as before and drown, as in Gobillon and Magnac (2016) once and for all, from a continuous uniform distribution, the first element of λ_t is 1 allowing for individual fixed effects. In our DGP the total, direct and indirect effects are given by

$$\begin{aligned} \alpha_t &= \kappa\psi + \phi\psi + \tau, \\ \theta_t(M_t(1)) &= \phi\psi + \tau, \\ \theta_t(M_t(0)) &= \tau, \\ \delta_t(1) &= \phi\psi + \kappa\psi, \\ \delta_t(0) &= \kappa\psi. \end{aligned} \tag{4.3}$$

4.2 Results

coming soon

5 Empirical application

coming soon

6 Conclusions

We introduced MASC a generalization of the Synthetic Control Method to decompose the total effect of an intervention in its direct and indirect effects in setting where possibly only few treated and control units are available. MASC is very easy to implement as it only requires the already available public available SCM algorithm.

A Derivation of "Synthetic" Y_{1t}^{01}

To easy the notation the subscript t is dropped from the weights. Following Abadie et al. (2010), consider a generic vector of weights $W = (w_{n+1}, ..., w_J)'$ such that $w_j \geq 0$ for all $j = n+1, ..., J$ and $w_{n+1} + ... + w_J = 1$. With these weights the synthetic value of Y_{1t}^{01} is given by

$$\sum_{j=n+1}^{J} w_j Y_{jt} = \zeta_t + \eta_t \sum_{j=n+1}^{J} w_j X_j + \lambda_t \sum_{j=n+1}^{J} w_j \mu_j + \sum_{j=n+1}^{J} w_j \varphi_t(0) M_{jt}(0) + \sum_{j=n+1}^{J} w_j \epsilon_{jt}.$$

The difference between the real potential outcome and the synthetic one is then

$$\begin{aligned} Y_{1t}^{0,1} - \sum_{j=n+1}^{J} w_j Y_{jt} &= \eta_t (X_1 - \sum_{j=n+1}^{J} w_j X_j) + \lambda_t (\mu_1 - \sum_{j=n+1}^{J} w_j \mu_j) \\ &+ \varphi_t(0) \left(M_{1t}(I\{t \geq T\}) - \sum_{j=n+1}^{J} w_j M_{jt}(0) \right) + \sum_{j=n+1}^{J} w_j (\epsilon_{1t} - \epsilon_{jt}). \end{aligned}$$

Let Y_i^P be the $((T-1)\times 1)$ vector with tth element equal to Y_{it}, ϵ_i^P the $((T-1)\times 1)$ vector with tth element equal to ϵ_{it}, η^P the $((T-1)\times r)$ matrix with tth row equal to η_t and λ^P the $((T-1)\times F)$ matrix with tth row equal to λ_t. Moreover, let $\varphi^P(0)$ be the $((T-1)\times 1)$ vector with tth element equal to $\varphi_t(0)$ and $M_i^P(0)$ the $((T-1)\times 1)$ vector with tth element equal to $M_{it}(0)$. We can now write

$$\begin{aligned} Y_1^P - \sum_{j=n+1}^{J} w_j Y_j^P &= \eta^P(X_1 - \sum_{j=n+1}^{J} w_j X_j) + \lambda^P(\mu_1 - \sum_{j=n+1}^{J} w_j \mu_j) \\ &+ \varphi^P(0)\left(M_{1t}^P(0) - \sum_{j=n+1}^{J} w_j M_{jt}^P(0)\right) + \left(\epsilon_1^P - \sum_{j=n+1}^{J} w_j \epsilon_j^P\right) \end{aligned} \tag{A.1}$$

Note that we have $M_{1t}^P(0)$ as $t < T$. It is easy to see that:

$$\begin{aligned} \lambda^P(\mu_1 - \sum_{j=n+1}^{J} w_j \mu_j) &= Y_1^P - \sum_{j=n+1}^{J} w_j Y_j^P - \eta^P(X_1 - \sum_{j=n+1}^{J} w_j X_j) \\ &- \varphi^P(0)(M_{1t}^P(0) - \sum_{j=n+1}^{J} w_j M_{jt}^P(0)) - (\epsilon_1^P - \sum_{j=n+1}^{J} w_j \epsilon_j^P) \\ &\Downarrow \\ \mu_1 - \sum_{j=n+1}^{J} w_j \mu_j &= (\lambda^P)^{-1}\left\{ Y_1^P - \sum_{j=n+1}^{J} w_j Y_j^P - \eta^P(X_1 - \sum_{j=n+1}^{J} w_j X_j) \right. \\ &- \varphi^P(0)(M_{1t}^P(0) - \sum_{j=n+1}^{J} w_j M_{jt}^P(0)) \\ &- \left. (\epsilon_1^P - \sum_{j=n+1}^{J} w_j \epsilon_j^P) \right\}. \end{aligned} \tag{A.2}$$

Assume that

Assumption 1. $\sum_{t=1}^{T-1} \lambda_t' \lambda_t$ *is non-singular.*

Assumption 1 is equivalent to assume no perfect-collinearity among unobserved common factors. Under Assumption 1 $(\lambda^{P'}\lambda^P)^{-1}$. Going back to A.2, we can then multiply and divide the right hand side by $(\lambda^{P'}\lambda^P)$ to obtain

$$\mu_1 - \sum_{j=n+1}^{J} w_j\mu_j = (\lambda^{P'}\lambda^P)^{-1}\lambda^{P'}\left\{Y_1^P - \sum_{j=n+1}^{J} w_jY_j^P - \eta^P(X_1 - \sum_{j=n+1}^{J} w_jX_j)\right.$$
$$\left. - -\varphi^P(0)(M_{1t}^P(0) - \sum_{j=n+1}^{J} w_jM_{jt}^P(0)) - (\epsilon_1^P - \sum_{j=n+1}^{J} w_j\epsilon_j^P)\right\}.$$

Substituting in A.1 we have:

$$\begin{gathered} Y_{1t}^{0,1} - \textstyle\sum_{j=n+1}^{J} w_jY_{jt} = \lambda_t(\lambda^{P'}\lambda^P)^{-1}\lambda^{P'}(Y_1^P - \sum_{j=n+1}^{J} w_jY_j^P)+ \\ (\eta_t - \lambda_t(\lambda^{P'}\lambda^P)^{-1}\lambda^{P'}\eta^P)(X_1 - \textstyle\sum_{j=n+1}^{J} w_jX_j) - \lambda_t(\lambda^{P'}\lambda^P)^{-1}\lambda^{P'}[\varphi^P(0)(M_1^P(0)- \\ \textstyle\sum_{j=n+1}^{J} w_jM_{jt}^P(0))] + \varphi_t(0)(M_{1t}(d) - \sum_{j=n+1}^{J} w_jM_{jt}(0))- \\ \lambda_t(\lambda^{P'}\lambda^P)^{-1}\lambda^{P'}(\epsilon_1^P - \textstyle\sum_{j=n+1}^{J} w_j\epsilon_j^P) + \sum_{j=n+1}^{J} w_j(\epsilon_{1t} - \epsilon_{jt}). \end{gathered} \tag{A.3}$$

If we replace the generic weights with the optimal $w_{n+1}^*, ..., w_J^*$, we have

$$\begin{gathered} Y_{1t}^{0,1} - \textstyle\sum_{j=n+1}^{J} w_j^*Y_{jt} = \varphi_t(0)(M_{1t}(d) - \sum_{j=n+1}^{J} w_j^*M_{jt}(0)) - \lambda_t(\lambda^{P'}\lambda^P)^{-1}\lambda^{P'}[\epsilon_1^P - \sum_{j=n+1}^{J} w_j^*\epsilon_j^P)]+ \\ \textstyle\sum_{j=n+1}^{J} w_j^*(\epsilon_{1t} - \epsilon_{jt}). \end{gathered} \tag{A.4}$$

Considering now the post-intervention period we have

$$\begin{gathered} Y_{1t}^{0,1} - \textstyle\sum_{j=n+1}^{J} w_j^*Y_{jt} = \varphi_t(0)(M_{1t}(1) - \sum_{j=n+1}^{J} w_j^*M_{jt}(0))- \\ \lambda_t(\lambda^{P'}\lambda^P)^{-1}\lambda^{P'}[\epsilon_1^P - \textstyle\sum_{j=n+1}^{J} w_j^*\epsilon_j^P)]+ \\ \textstyle\sum_{j=n+1}^{J} w_j^*(\epsilon_{1t} - \epsilon_{jt}) = \lambda_t(\lambda^{P'}\lambda^P)^{-1}\lambda^{P'}[\epsilon_1^P - \sum_{j=n+1}^{J} w_j^*\epsilon_j^P]+ \\ \textstyle\sum_{j=n+1}^{J} w_j^*(\epsilon_{1t} - \epsilon_{jt}) \quad \forall \quad t \geq T \end{gathered} \tag{A.5}$$

Unless specified otherwise from now on we always refer to the post-intervention period $t \geq T$. From here, the proof follow as in Abadie et al. (2010). We can write:

$$Y_{1t}^{0,1} - \sum_{j=n+1}^{J} w_j^*Y_{jt} = R_{1t} + R_{2t} + R_{3t}$$

where

$$R_{1t} = \lambda_t(\lambda^{P'}\lambda^P)^{-1}\lambda^{P'} \sum_{j=n+1}^{J} w_j^* \epsilon_j^P \tag{A.6}$$

$$R_{2t} = -\lambda_t(\lambda^{P'}\lambda^P)^{-1}\lambda^{P'}\epsilon_1^P \tag{A.7}$$

$$R_{3t} = \sum_{j=n+1}^{J} w_j^*(\epsilon_{jt} - \epsilon_{1t}) \tag{A.8}$$

Following Abadie et al. (2010), we impose the following assumptions

Assumption 2. $\epsilon_{it} \perp \epsilon_{jt}$ $\forall i \neq j$ *with* $i, j = 1, ..., J$.

Assumption 3. $\epsilon_{it} \perp \epsilon_{it'}$ $\forall t \neq t'$ *with* $t, t' = 1, ..., T$.

Assumption 4. $E(\epsilon_{it}|\{X_i, \mu_i, \{M_{it}(0)\}_{t=1}^{T}\}_{i=1}^{J}, \{M_{1t}(1)\}_{t \geq T}) = E(\epsilon_{it}) = 0$ *for* $i \in \{1, n+1, ..., J\}$

Taking the expected value on both sides of A.7 we get

$$E(R_{2t}) = E(-\lambda_t(\lambda^{P'}\lambda^P)^{-1}\lambda^{P'}\epsilon_1^P) = -\lambda_t(\lambda^{P'}\lambda^P)^{-1}\lambda^{P'}E(\epsilon_1^P) = 0$$

where the second equality follows from the fact that $-\lambda_t(\lambda^{P'}\lambda^P)^{-1}\lambda^{P'}$ is constant and the third equality follows from assumption 4. Taking the expectation on both sides of A.8

$$\begin{aligned} E(R_{3t}) &= E(\sum_{j=n+1}^{J} w_j^*(\epsilon_{jt} - \epsilon_{1t}))' = \sum_{j=n+1}^{J}[E(w_j^*\epsilon_{jt}) - E(w_j^*\epsilon_{1t})] \\ &= = \sum_{j=n+1}^{J}[E(w_j^*)E(\epsilon_{jt}) - E(w_j^*)E(\epsilon_{1t})] = 0 \end{aligned}$$

where the third equality follows from the fact that weights $W^* = w_{n+1}^*, ..., w_J^*$ are determined using constraints on covariates, pre-treatment period outcomes and the mediator at time $t \geq T$. Hence, under assumptions 2, 3 and 4 they are independent from the error terms at time $t \geq T$. The fourth equality follows from assumption 4. The reaming A.6

can be rewritten as:

$$R_{1t} = \sum_{j=n+1}^{J} w_j^* \sum_{s=1}^{T-1} \lambda_t (\sum_{h=1}^{T-1} \lambda_h' \lambda_h)^{-1} \lambda_s' \epsilon_{js} \tag{A.9}$$

As in Abadie et al. (2010) we further assume that

Assumption 5. *Let $\varsigma(M)$ be the smallest eigenvalue of*

$$\frac{1}{M} \sum_{t=T-M+1}^{T-1} \lambda_t' \lambda_t, \tag{A.10}$$

$\varsigma(M) \geq \underline{\varsigma} > 0$ for each positive integer M.

Assumption 6.

$$\exists \ \ \underline{\lambda} \ s.t. \ |\lambda_{tf}| \leq \underline{\lambda} \quad \forall \ t{=}1,...,T \ and \ f{=}1,...,F \ . \tag{A.11}$$

Assumption 5 guarantees that the matrix $\sum_{t=1}^{T} \lambda_t' \lambda_t$ and, consequently, its inverse are symmetric and positive definite. As $A = \sum_{t=1}^{T} \lambda_t' \lambda_t$ is symmetric and positive definite for the Cauchy-Schwarz inequality we have that

$$\begin{aligned} \left(\lambda_t \left(\sum_{h=1}^{T-1} \lambda_h' \lambda_h \right)^{-1} \lambda_s' \right)^2 &= |\langle \lambda_t, A \lambda_s' \rangle|^2 \leq ||A\lambda_t||^2 ||A\lambda_s||^2 \\ &= \left(\lambda_t \left(\sum_{h=1}^{T-1} \lambda_h' \lambda_h \right)^{-1} \lambda_t' \right) \left(\lambda_s \left(\sum_{h=1}^{T-1} \lambda_h' \lambda_h \right)^{-1} \lambda_s' \right) \end{aligned} \tag{A.12}$$

Since A is a symmetric matrix $B = \frac{1}{T-1} A$ is symmetric as well. Thus, it can be decomposed as $B = GOG^{-1}$. Where G is orthogonal and $G^{-1} = G'$ and O is a diagonal matrix with the eigenvalues of B as elements. Thus,

$$\lambda_t (\sum_{h=1}^{T-1} \lambda_h' \lambda_h)^{-1} \lambda_t' = \frac{1}{T-1} (\lambda_t B \lambda_t') = \frac{1}{T-1} (\lambda_t GOG' \lambda_t') \tag{A.13}$$

Defining $b_t = \lambda_t G$ we have

$$\lambda_t (\sum_{h=1}^{T-1} \lambda_h' \lambda_h)^{-1} \lambda_t' = \tfrac{1}{T-1} (b_t O b_t') = \tfrac{1}{T-1} (b_{t1}^2 \tfrac{1}{\varsigma_1} + ... + b_{tF}^2 \tfrac{1}{\varsigma_F}) \tag{A.14}$$

where ς_i are the eigenvalues of matrix B. From assumption 5, imposing $M = T-1$, we'll have that $\frac{1}{\varsigma_i} \leq \frac{1}{\underline{\varsigma}}$ for $i = 1, ..., F$. Indeed the eigenvalues of the inverse of a matrix are given by the inverse of the matrix eigenvalues, and B is the inverse of the matrix in assumption 5. Consequently:

$$\lambda_t(\sum_{h=1}^{T-1} \lambda_h' \lambda_h)^{-1} \lambda_t' = \frac{1}{T-1} \sum_{f=1}^{F} \frac{b_{tf}^2}{\varsigma_f} \leq \frac{1}{(T-1)\underline{\varsigma}} \sum_{f=1}^{F} b_{tf}^2 = \frac{1}{(T-1)\underline{\varsigma}} ||b_t||^2 = \frac{1}{(T-1)\underline{\varsigma}} ||\lambda_t G||^2 \tag{A.15}$$

But, as said before, G is an orthogonal matrix, consequently it is isometric, hence $||\lambda_t G|| = ||\lambda_t||$. Consequently:

$$\lambda_t(\sum_{h=1}^{T-1} \lambda_h' \lambda_h)^{-1} \lambda_t' \leq \frac{1}{(T-1)\underline{\varsigma}} ||\lambda_t||^2 = \frac{\sum_{f=1}^{F} \lambda_{tf}^2}{(T-1)\underline{\varsigma}} \leq \frac{\sum_{f=1}^{F} \underline{\lambda}^2}{(T-1)\underline{\varsigma}} = \frac{F\underline{\lambda}^2}{(T-1)\underline{\varsigma}} \tag{A.16}$$

where the last inequality follows from assumption 6. We can apply the same reasoning to the content of the second parenthesis in the final part of A.12. We'll have then:

$$(\lambda_t(\sum_{h=1}^{T-1} \lambda_h' \lambda_h)^{-1} \lambda_s')^2 \leq (\lambda_t(\sum_{h=1}^{T-1} \lambda_h' \lambda_h)^{-1} \lambda_t')(\lambda_s(\sum_{h=1}^{T-1} \lambda_h' \lambda_h)^{-1} \lambda_s') \leq \frac{F\underline{\lambda}^2}{(T-1)\underline{\varsigma}} \tag{A.17}$$

Following Abadie et al. (2010) let's define:

$$\overline{\epsilon_j^L} = \sum_{s=1}^{T-1} \lambda_T (\sum_{h=1}^{T-1} \lambda_h' \lambda_h)^{-1} \lambda_s' \epsilon_{js} \tag{A.18}$$

for $j = n+1, ..., J$. Let's assume:

Assumption 7. *$\exists$ a p^{th} moment of $|\epsilon_{jt}|$ for some even p and for $j = 2, ..., J$ and $t = 1, ..., T-1$*

Using Hölder's Inequality and taking into account that $0 \leq w_j^* \leq 1$ for $j = n+1, ..., J$ we have that:

$$\sum_{j=n+1}^{J} w_j^* |\overline{\epsilon_j^L}| = \sum_{j=n+1}^{J} w_j^* |\overline{\epsilon_j^L} * 1| \leq (\sum_{j=n+1}^{J} w_j^* |\overline{\epsilon_j^L}|^p)^{1/p} (\sum_{j=n+1}^{J} w_j^* |1|^q)^{1/q} = (\sum_{j=n+1}^{J} w_j^* |\overline{\epsilon_j^L}|^p)^{1/p} (\sum_{j=n+1}^{J} w_j^*)^{1/q} = (\sum_{j=n+1}^{J} w_j^* |\overline{\epsilon_j^L}|^p)^{1/p} \tag{A.19}$$

where the last equality follow from $w^*_{n+1} + ... + w^*_J = 1$. Applying Hölder's Inequality again:

$$E[\sum_{j=n+1}^{J} w_j^* |\overline{\epsilon_j^L}|] \leq (E[\sum_{j=n+1}^{J} w_j^* |\overline{\epsilon_j^L}|^p])^{1/p} \tag{A.20}$$

Now, applying Rosenthal's Inequality:

$$E[|\overline{\epsilon_j^L}|^p] = E[|\sum_{s=1}^{T-1} \lambda_t (\sum_{h=1}^{T-1} \lambda'_h \lambda_h)^{-1} \lambda'_s \epsilon_{js}|] \leq$$
$$C(p) max(\sum_{s=1}^{T-1} E[|\lambda_t (\sum_{h=1}^{T-1} \lambda'_h \lambda_h)^{-1} \lambda'_s \epsilon_{js}|^p], (\sum_{s=1}^{T-1} E[|\lambda_t (\sum_{h=1}^{T-1} \lambda'_h \lambda_h)^{-1} \lambda'_s \epsilon_{js}|^2])^{p/2}) \tag{A.21}$$

where $C(p)$ is the pth moment of minus 1 plus a Poisson random variable with parameter equal to one (see Abadie et al. (2010)). Let's consider the two elements contained in $max(.)$ separately. For the first element:

$$\begin{gathered}\sum_{s=1}^{T-1} E[|\lambda_t (\sum_{h=1}^{T-1} \lambda'_h \lambda_h)^{-1} \lambda'_s \epsilon_{js}|^p] = \\ \sum_{s=1}^{T-1} E[(\lambda_t (\sum_{h=1}^{T-1} \lambda'_h \lambda_h)^{-1} \lambda'_s)^{2*(p/2)} |\epsilon_{js}|^p] \leq \sum_{s=1}^{T-1} E[(\frac{F\underline{\lambda}^2}{(T-1)\underline{\varsigma}})^{2*(p/2)} |\epsilon_{js}|^p] = \\ (\frac{F\underline{\lambda}^2}{\underline{\varsigma}})^p \frac{1}{(T-1)^p} \sum_{s=1}^{T-1} E(|\epsilon_{js}|^p)\end{gathered} \tag{A.22}$$

where the first equality follows from the distributivity of the power and the inequality follows from A.17. For the second element in $max(.)$:

$$(\sum_{s=1}^{T-1} E[|\lambda_t (\sum_{h=1}^{T-1} \lambda'_h \lambda_h)^{-1} \lambda'_s \epsilon_{js}|^2])^{p/2} \leq [\sum_{s=1}^{T-1} E((\frac{F\overline{\lambda}^2}{(T-1)\underline{\varsigma}})^2 \epsilon_{js}^2)]^{p/2} = (\frac{F\overline{\lambda}^2}{\underline{\varsigma}})^p [\sum_{s=1}^{T-1} \frac{1}{(T-1)^2} E(\epsilon_{js}^2)]^{p/2} \tag{A.23}$$

where the first inequality follows from A.17. From these results we can write:

$$E[|\overline{\epsilon_j^L}|^p] \leq C(p) (\frac{F\overline{\lambda}^2}{\underline{\varsigma}})^p \max(\frac{1}{(T-1)^p} \sum_{s=1}^{T-1} E(|\epsilon_{js}|^p), [\sum_{s=1}^{T-1} \frac{1}{(T-1)^2} E(\epsilon_{js}^2)]^{p/2}) \tag{A.24}$$

As Abadie et al. (2010) let's define $\sigma_{js}^2 = E|\epsilon_{js}|^2$, $\sigma_j^2 = (1/(T-1) \sum_{s=1}^{T-1} \sigma_{js}^2)$, $\overline{\sigma^2} = max_{j=n+1,...,J} \sigma_j^2$ and $\overline{\sigma} = \sqrt{\overline{\sigma^2}}$. Similarly, let's define $\tau_{p,jt} = E|\epsilon_{jt}|^p$, $\tau_{p,j} = \frac{1}{(T-1)} \sum_{t=1}^{T-1} \tau_{p,jt}$, and $\overline{\tau_p} = max_{j=n+1,...,J} \tau_{p,j}$. We can write the first value inside $max(.)$ as:

$$\frac{1}{(T-1)^p} \sum_{s=1}^{T-1} E(|\epsilon_{js}|^p) = \frac{1}{(T-1)^{p-1}} \frac{1}{(T-1)} \sum_{t=1} T - 1 \tau_{pjt} = \frac{1}{(T-1)^{p-1}} \tau_{pj} \tag{A.25}$$

Equally, we can write the second value inside $max(.)$ as:

$$[\sum_{s=1}^{T-1} \tfrac{1}{(T-1)^2} E(\epsilon_{js}^2)]^{p/2} = (\tfrac{1}{T-1} \tfrac{1}{T-1} \sum_{s=1}^{T-1} \sigma_{js}^2)^{p/2} = (\tfrac{1}{T-1} \sigma_j^2)^{p/2} \tag{A.26}$$

Consequently, calling $\varpi = C(p)(\frac{F\overline{\lambda}^2}{\underline{\varsigma}})^p$:

$$\begin{aligned}
E\left[|\overline{\epsilon_j^L}|^p\right] &\leq \varpi \max\left(\frac{1}{(T-1)^{p-1}}\tau_{pj}, \left(\frac{1}{T-1}\sigma_j^2\right)^{p/2}\right) \\
\sum_{j=n+1}^{J} E\left[|\overline{\epsilon_j^L}|^p\right] &= E\left[\sum_{j=n+1}^{J} |\overline{\epsilon_j^L}|^p\right] \\
&\leq \varpi \max\left(\frac{1}{(T-1)^{p-1}}\sum_{j=n+1}^{J}\tau_{pj}, \sum_{j=n+1}^{J}\left(\frac{1}{T-1}\sigma_j^2\right)^{p/2}\right) \\
&= \varpi \max\left(\frac{J-n-1}{(T-1)^{p-1}}\frac{1}{J-n-1}\sum_{j=n+1}^{J}\tau_{pj}, \frac{1}{(T-1)^{p/2}}\sum_{j=n+1}^{J}\sigma_j^{2*p/2}\right) \\
\left(E\left[\sum_{j=n+1}^{J} |\overline{\epsilon_j^L}|^p\right]\right)^{1/p} &\leq \varpi^{1/p} \max\left(\frac{\left(\frac{J-n-1}{(T-1)^{p-1}}\right)^{1/p}}{(J-n-1)^{1/p}}\left(\sum_{j=n+1}^{J}\tau_{pj}\right)^{1/p}, \frac{\left(\sum_{j=n+1}^{J}\sigma_j^{2*p/2}\right)^{1/p}}{(T-1)^{(p/2)*(1/p)}}\right) \\
&= \varpi^{1/p} \max\left(\left(\frac{J-n-1}{(T-1)^{p-1}}\right)^{1/p}\overline{\tau}_p^{1/p}, \frac{1}{(T-1)^{1/2}}\left(\sum_{j=n+1}^{J}\overline{\sigma}^{2*(p/2)}\right)^{1/p}\right)
\end{aligned}$$

where the last equality follows from $\frac{1}{J-n-1}\sum_{j=n+1}^{J}\tau_{pj} = E(\tau_{pj}) \leq max_j(\tau_{pj}) = \overline{\tau_p}$. We obtain that:

$$\begin{aligned}
\left(E\left[\sum_{j=n+1}^{J}|\overline{\epsilon_j^L}|^p\right]\right)^{1/p} &\leq \varpi^{1/p} max\left(\frac{(J-n-1)^{1/p}\,\overline{\tau}_p^{1/p}}{(T-1)^{1-1/p}},\frac{(J-n-1)\,\overline{\sigma}^{2*(p/2)}}{(T-1)^{1/2}}\right)^{1/p} \\
&= \varpi^{1/p}\,(J-n-1)^{1/p}\,max\left(\frac{\overline{\tau}_{1/p}}{(T-1)^{1-\frac{1}{p}}},\frac{\sqrt{\overline{\sigma}^2}}{(T-1)^{1/2}}\right) \\
E\left[|R_{1t}|\right] &= E\left[|\sum_{j=n+1}^{J}w_j^*\epsilon_j^L|\right] \\
&\leq E\left[\sum_{j=n+1}^{J}w_j^*|\epsilon_j^L|\right] \\
&\leq \left(E\left[\sum_{j=n+1}^{J}|\epsilon_j^L|^p\right]\right)^{1/p} \\
&\leq \varpi^{1/p}J^{1/p}max\left(\frac{\overline{\tau}_p^{1/p}}{(T-1)^{1-\frac{1}{p}}},\frac{\overline{\sigma}}{(T-1)^{1/2}}\right)
\end{aligned}$$

where, in the second equation, the first equality follows from A.7 and A.18, the first inequality follows from the triangular inequality, the second follows from A.20 and the third from the first equation. It follows:

$$E|R_{1t}| \leq C(p)^{1/p}\frac{\overline{\lambda^2}F}{\underline{\varsigma}}(J-n)^{1/p}max\{\frac{\overline{\tau_p^{1/p}}}{(T-1)^{1-1/p}},\frac{\overline{\sigma}}{(T-1)^{1/2}}\}. \tag{A.27}$$

This means that the bias of the estimator can be bounded by a function that goes to zero when the number of pre-treatment periods goes to infinity. Hence, we can write the bias of the estimator as:

$$E(Y_{1t}^{0,1}-\sum_{j=n+1}^{J}w_j^*Y_{jt}) = E(R_{1t}) = g_2(T) \tag{A.28}$$

where $g_2(T)$ is a function going to zero for T going to infinity.

B Extra assumptions on the mediator needed for Y_{1t}^{10}

For estimating Y_{1t}^{10} we need to impose the standard SCM assumption on the mediator. These assumptions are listed below.

Assumption 8. $\sum_{t=1}^{T-1} \vartheta_t' \vartheta_t$ *is non-singular.*

Assumption 9. $\nu_{it} \perp \nu_{jt}$ $\forall i \neq j$ *with* $i, j \in \{1, n+1, ..., J\}$.

Assumption 10. $\nu_{it} \perp \nu_{it'}$ $\forall t \neq t'$ *with* $t, t' = 1, ..., T$.

Assumption 11. $E(\nu_{it}|\{Z_i, \varrho_i\}_{i\in\{1,n+1,...,J\}}) = E(\nu_{it}) = 0$ *for* $i \in \{1, n+1, ..., J\}$

Assumption 12. $\kappa(M) \geq \underline{\kappa} > 0$ *for each positive integer M, where* $\kappa(M)$ *is the smallest eigenvalue of*

$$\frac{1}{M} \sum_{t=T-M+1}^{T-1} \vartheta_t' \vartheta_t. \tag{B.1}$$

Assumption 13.

$$\exists \quad \underline{\vartheta} \text{ s.t. } |\vartheta_{tv}| \leq \underline{\vartheta} \quad \forall \text{ t=1,...,} T \text{ and v=1,...,} V\,. \tag{B.2}$$

Assumption 14. $\exists$ *a* p^{th} *moment of* $|\nu_{jt}|$ *for some even* p *and for* $j = n+1, ..., J$ *and* $t = 1, ..., T$

C Derivation of "Synthetic" Y_{1t}^{10}

A generic synthetic value of the outcome variable for the treated group is given by:

$$\sum_{j=2}^{n} q_j Y_{jt} = \zeta_t + \eta_t \sum_{j=2}^{n} q_j X_j + \lambda_t \sum_{j=2}^{n} q_j \mu_j + \sum_{j=2}^{n} q_j \varphi_t(D_{jt}) M_{jt}(D_{jt}) + \sum_{j=2}^{n} q_j \rho_t(M_{jt}(D_{jt})) D_{jt} + \sum_{j=2}^{n} q_j \epsilon_{jt} \tag{C.1}$$

where we have that D_{jt} takes value 0 for $t < T$ and 1 otherwise, because we're dealing with all treated units.

$$\begin{gathered} Y_{1t}^{1,0} - \textstyle\sum_{j=2}^{n} q_j Y_{jt} = \eta_t(X_1 - \sum_{j=2}^{n} q_j X_j) + \lambda_t(\mu_1 - \sum_{j=2}^{n} q_j \mu_j) + \\ \varphi_t(D_{1t}) M_{1t}(0) - \textstyle\sum_{j=2}^{n} q_j \varphi_t(D_{jt}) M_{jt}(D_{jt}) + \rho_t(M_{1t}(0)) D_{1t} - \sum_{j=2}^{n} q_j \rho_t(M_{jt}(D_{jt})) D_{jt} + \\ \textstyle\sum_{j=2}^{n} q_j(\epsilon_{1t} - \epsilon_{jt}) \end{gathered} \tag{C.2}$$

Let Y_i^P be the $((T-1)\times 1)$ vector with tth element equal to Y_{it}, $\varphi^P(D_i^P)$ be the $((T-1)\times 1)$ vector with tth element equal to $\varphi_t(D_{it})$, $M_i^P(D_i^P)$ be the $((T-1)\times 1)$ vector with tth

element equal to $M_{it}(D_{it})$ and $\rho^P(M_i^P(D_i^P))$ be the vector with tth element equal to $\rho_t(M_{it}(D_{it}))$, D_i^P be the $((T-1)\times 1)$ vector with tth element equal to D_{it}. Using the other notation as before, we can write:

$$\begin{gathered} Y_1^P - \textstyle\sum_{j=2}^n q_j Y_j^P = \eta^P(X_1 - \sum_{j=2}^n q_j X_j) + \lambda^P(\mu_1 - \sum_{j=2}^n q_j \mu_j) + \\ \varphi^P(D_1^P)M_1^P(0) - \textstyle\sum_{j=2}^n q_j \varphi^P(D_j^P)M_j^P(D_j^P) + \rho^P(M_1^P(0))D_1^P - \sum_{j=2}^n q_j \rho^P(M_j^P(D_j^P))D_j^P + \\ (\epsilon_1^P - \textstyle\sum_{j=2}^n q_j \epsilon_j^P) = \eta^P(X_1 - \sum_{j=2}^n q_j X_j) + \lambda^P(\mu_1 - \sum_{j=2}^n q_j \mu_j) + \\ \varphi^P(0)(M_1^P(0) - \textstyle\sum_{j=2}^n q_j M_j^P(0)) + (\epsilon_1^P - \sum_{j=2}^n q_j \epsilon_j^P) \end{gathered} \tag{C.3}$$

Where the second equality follow from the fact that in period $t < T$ we have $D_{jt} = 0$ for j=1,...,n. It is easy to see that:

$$\begin{gathered} \lambda^P(\mu_1 - \textstyle\sum_{j=2}^n q_j \mu_j) = Y_1^P - \sum_{j=2}^n q_j Y_j^P - \eta^P(X_1 - \sum_{j=2}^n q_j X_j) - \\ \varphi^P(0)(M_1^P(0) - \textstyle\sum_{j=2}^n q_j M_j^P(0)) - (\epsilon_1^P - \sum_{j=2}^n q_j \epsilon_j^P) \end{gathered} \tag{C.4}$$

$$\begin{gathered} \mu_1 - \textstyle\sum_{j=2}^n q_j \mu_j = \lambda^{P-1}\{Y_1^P - \sum_{j=2}^n q_j Y_j^P - \eta^P(X_1 - \sum_{j=2}^n q_j X_j) - \\ \varphi^P(0)(M_1^P(0) - \textstyle\sum_{j=2}^n q_j M_j^P(0)) - (\epsilon_1^P - \sum_{j=2}^n q_j \epsilon_j^P)\}. \end{gathered}$$

Multiplying and dividing on the right hand side for $(\lambda^{P'}\lambda^P)$ (we are able to do that thanks to assumption 1) we obtain:

$$\begin{gathered} \mu_1 - \textstyle\sum_{j=2}^n q_j \mu_j = (\lambda^{P'}\lambda^P)^{-1}\lambda^{P'}\{(Y_1^P - \sum_{j=2}^n q_j Y_j^P) - \eta^P(X_1 - \sum_{j=2}^n q_j X_j) - \\ \varphi^P(0)(M_1^P(0) - \textstyle\sum_{j=2}^n q_j M_j^P(0)) - (\epsilon_1^P - \sum_{j=2}^n q_j \epsilon_j^P)\}. \end{gathered} \tag{C.5}$$

Substituting in C.2 we have:

$$\begin{gathered} Y_{1t}^{1,0} - \textstyle\sum_{j=2}^n q_j Y_{jt} = (\lambda^{P'}\lambda^P)^{-1}\lambda^{P'}(Y_1^P - \sum_{j=2}^n q_j Y_j^P) + \\ (\eta_t - (\lambda^{P'}\lambda^P)^{-1}\lambda^{P'}\eta^P)(X_1 - \textstyle\sum_{j=2}^n q_j X_j) - (\lambda^{P'}\lambda^P)^{-1}\lambda^{P'}\varphi^P(0)(M_1^P(0) - \sum_{j=2}^n q_j M_j^P(0)) + \\ \varphi_t(D_{1t})M_{1t}(0) - \textstyle\sum_{j=2}^n q_j \varphi_t(D_{jt})M_{jt}(D_{jt}) + \rho_t(M_{1t}(0))D_{1t} - \sum_{j=2}^n q_j \rho_t(M_{jt}(D_{jt}))D_{jt} - \\ (\lambda^{P'}\lambda^P)^{-1}\lambda^{P'}(\epsilon_1^P - \textstyle\sum_{j=2}^n q_j \epsilon_j^P) + \sum_{j=2}^n q_j(\epsilon_{1t} - \epsilon_{jt}) \end{gathered} \tag{C.6}$$

If we substitute the generic weigths with $q_2^*, ..., q_n^*$, thanks to conditions **??**, we'll have:

$$Y_{1t}^{1,0} - \sum_{j=2}^{n} q_j^* Y_{jt} = \varphi_t(D_{1t})M_{1t}(0) - \sum_{j=2}^{n} q_j^* \varphi_t(D_{jt})M_{jt}(D_{jt}) + \rho_t(M_{1t}(0))D_{1t} - \sum_{j=2}^{n} q_j^* \rho_t(M_{jt}(d))D_{jt} - (\lambda^{P'}\lambda^{P})^{-1}\lambda^{P'}[\epsilon_1^P - \sum_{j=n+1}^{J} q_j^* \epsilon_j^P)] + \sum_{j=n+1}^{J} q_j^*(\epsilon_{1t} - \epsilon_{jt}). \tag{C.7}$$

Considering now the period of interest we'll have:

$$Y_{1t}^{1,0} - \sum_{j=2}^{n} q_j^* Y_{jt} = \varphi_t(1)(M_{1t}(0) - \sum_{j=2}^{n} q_j^* M_{jt}(1)) + \rho_t(M_{1t}(0)) * 1 - \sum_{j=2}^{n} q_j^* \rho_t(M_{jt}(1)) * 1 - (\lambda^{P'}\lambda^{P})^{-1}\lambda^{P'}[\epsilon_1^P - \sum_{j=n+1}^{J} q_j^* \epsilon_j^P)] + \sum_{j=n+1}^{J} q_j^*(\epsilon_{1t} - \epsilon_{jt}) \quad \forall \quad t \geq T \tag{C.8}$$

where we could substitute D_{jt} with 1 for all $j = 1, .., n$ because all units are treated. From now on, unless specified otherwise, we'll consider the period $t \geq T$. We now have to introduce another assumption:

Assumption 15. *$\rho_t(.)$ is a linear function*

Thanks to assumption 15 we can bring the sum inside $\rho_t(.)$. The condition in **??** for the period of interest allows us to substitute $\sum_{j=2}^{n} q_j^* M_{jt}(1)$ with $\hat{M}_{1t}(0)$. It follows:

$$Y_{1t}^{1,0} - \sum_{j=2}^{n} q_j^* Y_{jt} = \varphi_t(1)(M_{1t}(0) - \hat{M}_{1t}(0)) + \rho_t(M_{1t}(0)) - \rho_t(\sum_{j=2}^{n} q_j^* M_{jt}(1)) - (\lambda^{P'}\lambda^{P})^{-1}\lambda^{P'}[\epsilon_1^P - \sum_{j=n+1}^{J} q_j^* \epsilon_j^P)] + \sum_{j=n+1}^{J} q_j^*(\epsilon_{1t} - \epsilon_{jt}) = \varphi_t(1)(M_{1t}(0) - \hat{M}_{1t}(0)) + \rho_t(M_{1t}(0)) - \rho_t(\hat{M}_{1t}(0)) + R_{1t} + R_{2t} + R_{3t} \tag{C.9}$$

where to define R_{1t}, R_{2t}, R_{3t} we used the same notation as in A. Taking the expected values we'll have:

$$E(Y_{1t}^{1,0} - \sum_{j=2}^{n} q_j^* Y_{jt}) = E(\varphi_t(1)(M_{1t}(0) - \hat{M}_{1t}(0))) + E(\rho_t(M_{1t}(0)) - E(\rho(\hat{M}_{1t}(0))) + E(R_{1t}) + E(R_{2t}) + E(R_{3t}) = \varphi_t(1)M_{1t}(0) - \varphi_t(1)E(\hat{M}_{1t}(0)) + \rho_t(M_{1t}(0)) - \rho_t(E(\hat{M}_{1t}(0))) + E(R_{1t}) \tag{C.10}$$

where the second equality follow from the fact that $\varphi_t(1)$ and $M_{1t}(0)$ are constant and

$\rho_t(M_{1t}(0))$ is constant. We could bring the expected value inside $\rho_t(.)$ because it is a linear function and its only random component is inside $\hat{M}_{1t}(0)$. Finally, $E(R_{2t}) = 0$ and $E(R_{3t}) = 0$ can be demonstrated as before (see appendix A) adding the following assumption:

Assumption 16. $E(\epsilon_{it}|\{\{M_{jt}(1)\}_{j=2}^n\}, \hat{M}_{1t}(0)\}_{t\geq T}) = E(\epsilon_{it})$ *for* $i = 1, ..., n$

Substituting $E(\hat{M}_{1t}(0))$ with $M_{1t}(0) + g_0(T)$ in C.10 we'll have:

$$E(Y_{1t}^{1,0} - \sum_{j=2}^n q_j^* Y_{jt}) = \varphi_t(1)M_{1t}(0) - \varphi_t(1)M_{1t}(0) - \varphi_t(1)g_0(T) + \rho_t(M_{1t}(0)) - \rho_t(M_{1t}(0) + g_0(T)) + E(R_{1t}) = -\varphi_t(1)g_0(T) + \rho_t(M_{1t}(0)) - \rho_t(M_{1t}(0) + g_0(T)) + g_3(T) \tag{C.11}$$

where the second equality can be demonstrated as before (see appendix A) and $g_3(T)$ is a function going to 0 when T goes to infinity. We have to add the following assumption:

Assumption 17. $\varphi_T(1) < \infty$

This assumption is fairly reasonable, considering that if it is violated $Y^{0,1}$ and $Y^{1,1}$ wouldn't be finite. Taking the limit for T going to infinity of C.11 we'll have:

$$\lim_{T\to\infty} E(Y_{1t}^{1,0} - \sum_{j=2}^n q_j^* Y_{jt}) = \rho_t(M_{1t}(0)) - \rho_t(M_{1t}(0)) = 0 \tag{C.12}$$

where the fact that $\lim_{T\to\infty} \varphi_t(1)g_0(T) = 0$ is guaranteed by assumption 17. And the fact that $\lim_{T\to\infty} \rho(M_{1t}(0) + g_0(T)) = \rho_t(M_{1t}(0))$ is guaranteed by assumption 15. Indeed, if $\rho_t(.)$ is a linear function, it will be continuous on all its domain. Hence, it will be continuous in $M_{1t}(0)$ as well.

alala

References

Abadie, A., A. Diamond, and J. Hainmueller (2010): "Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program," *Journal of the American Statistical Association*, 105, 493–505.

——— (2015): "Comparative Politics and the Synthetic Control Method," *American Journal of Political Science*, 59, 495–510.

Abadie, A. and J. Gardeazabal (2003): "The Economic Costs of Conflict: A Case Study of the Basque Country," *American Economic Review*, 93, 113–132.

Adhikari, B. (2015): "When Does Introducing a Value-Added Tax Increase Economic Efficiency? Evidence from the Synthetic Control Method," *Tulane University Working Paper*.

Athey, S., M. Bayati, N. Doudchenko, G. Imbens, and K. Khosravi (2017): "Matrix Completion Methods for Causal Panel Data Modelss," *arXiv:1710.10251*.

Doudchenko, N. and G. W. Imbens (2016): "Balancing, Regression, Difference-In-Differences and Synthetic Control Methods: A Synthesis," *arXiv:1710.10251*.

Gobillon, L. and T. Magnac (2016): "Regional policy evaluation: interactive fixed effects and synthetic controls," *Review of Economics and Statistics*, 98, 535–551.

Huber, M. (2013): "Identifying causal mechanisms in experiments (primarily) based on inverse probability weighting," *Economics Working Paper Series, University of St. Gallen, School of Economics and Political Science*.

Huber, M., M. Lechner, and G. Mellace (2016): "The Finite Sample Performance of Estimators for Mediation Analysis Under Sequential Conditional Independence," *Journal of Business & Economic Statistics*, 34, 139–160.

——— (2017): "Why Do Tougher Caseworkers Increase Employment? The Role of Program Assignment as a Causal Mechanism," *Review of Economics and Statistics*, 99, 180–183.

Imai, K., L. Keele, and T. Yamamoto (2010): "Identification, Inference and Sensitivity Analysis for Causal Mediation Effects," *Statistical Science*, 25, 51–71.

Imai, K. and T. Yamamoto (2013): "Identification and Sensitivity Analysis for Multiple Causal Mechanisms: Revisiting Evidence from Framing Experiments," *Political Analysis*, 141–171.

Pearl, J. (2001): "Direct and Indirect Effects," *Proceedings of the Seventeenth Conference on Uncertainty in Artificial Intelligence, San Francisco, CA: Morgan Kaufmann*, 411–420.

Robins, J. M. (2003): "Semantics of causal DAG models and the identification of direct and indirect effects," in *Highly Structured Stochastic Systems*, ed. by P. Green, N. Hjort, and S. Richardson, Oxford: Oxford University Press, 70–81.

VANSTEELANDT, S. AND T. J. VANDERWEELE (2012): "Natural Direct and Indirect Effects on the Exposed: Effect Decomposition under Weaker Assumptions," *Biometrics*, 68, 1019–1027.

XU, Y. (2017): "Generalized Synthetic Control Method: Causal Inference with Interactive Fixed Effects Models," *Political Analysis*, 25, 57–76.