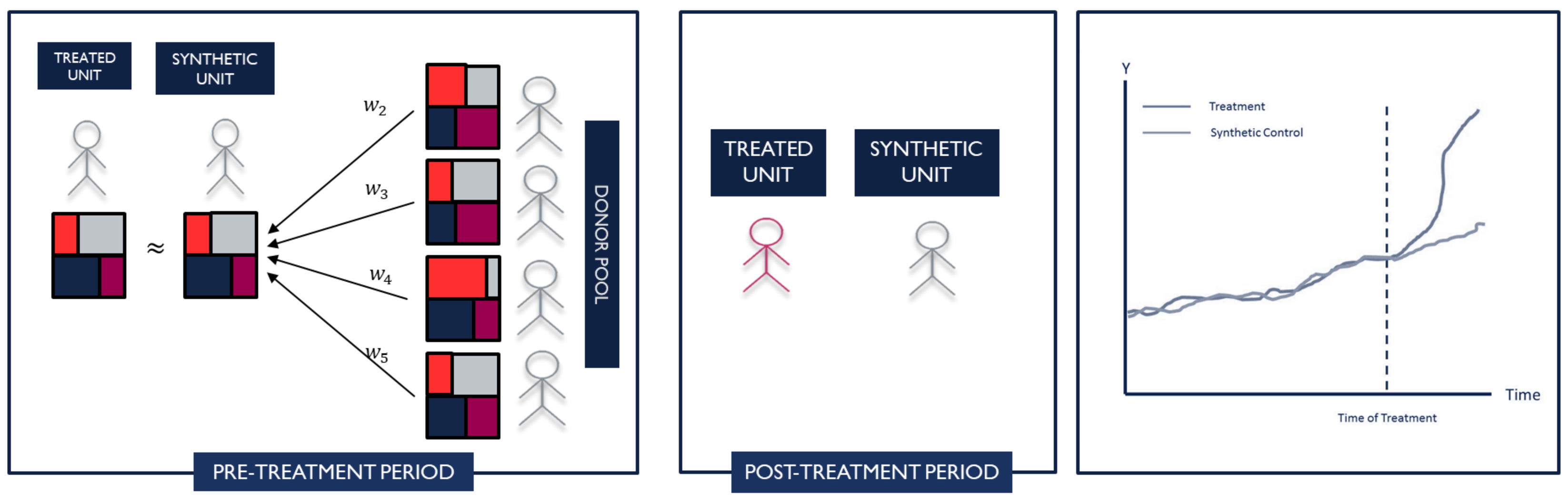
The inclusive Synthetic Control Method

Roberta Di Stefano* and Giovanni Mellace**

*Sapienza University of Rome **University of Southern Denmark

The SCM

- The synthetic control method (SCM) uses a weighted average of control units' outcomes to reproduce the counterfactual outcome of the treated unit.
- The weights are chosen to minimize the distance in the pre-intervention observed characteristics of the treated and the synthetic control.
- The causal effect ($\boldsymbol{\theta_{1t}}$) is estimated as the difference between the observed outcome of treated and one of the synthetic control in the post-intervention period.



iSCM

Two *scenarios* in which using the SCM might be problematic:

1. Some of the treated units need to be included in the donor pool to have a better pre-treatment fit.
2. Some of the control units in the donor pool are affected by the intervention indirectly.

Our *inclusive synthetic control method* (iSCM) allows using units affected by the intervention in the donor pool.

ADVANTAGES:

- You can use your favorite SC-type estimator.
- Units potentially affected either directly or indirectly by the treatment can be included in the donor pool.
- It does not impose any assumptions on treatment effect heterogeneity.

LIMITATIONS:

- The number of "potentially affected" units must be known.
- We cannot have too many of those units.
- It requires the existence of at least one "pure control" unit.

How it works

We assume that SUTVA holds such that the observed outcome $\boldsymbol{Y_{jt}}$ is equal to

- $\boldsymbol{Y^N_{jt}}$ in the pre-intervention period for everyone and in the post intervention period for pure control units.
- $\boldsymbol{Y^S_{jt}}$ in the post intervention period for potentially affected units.
- $\boldsymbol{Y^I_{jt}}$ in the post intervention period for the main treated unit.

Under this assumption any SCM-type estimator for the main treated will converge to

$$\begin{aligned}\hat{Y}^N_{1t} &= \sum_{j=2}^{J} \hat{w}_j Y_{jt} \\ &= \sum_{j=m+1}^{J} \hat{w}_j Y^N_{jt} + \sum_{j=2}^{m} \hat{w}_j \underbrace{(Y^N_{jt} + \gamma_{jt})}_{Y^S_{jt}} \\ &= \sum_{j=2}^{J} \hat{w}_j Y^N_{jt} + \sum_{j=2}^{m} \hat{w}_j \gamma_{jt} \\ &\to Y^N_{1t} + \sum_{j=2}^{m} \hat{w}_j \gamma_{jt}\end{aligned}$$

Similar results hold for each potentially affected unit.

Ignoring estimation bias we have that

$$\vartheta_t = \hat{\Omega}^{-1} \hat{\beta}_t.$$

Where:

- ϑ_t is the $(m \times 1)$ vector of effects of interest
- $\hat{\Omega}$ is the $(m \times m)$ matrix of estimated weights
- $\hat{\beta}_t$ is the $(m \times 1)$ vector of biased estimated effects

Thus, using Cramer's rule:

$$\vartheta_{jt} = \frac{\det(\hat{\Omega}_{j,t})}{det(\hat{\Omega})}, \qquad j = 1, \dots, m.$$

Why using iSCM?

1. Implement SCM, including units potentially affected by the intervention (i.e., other treated units or units affected by spillovers).

- If all potentially affected units receive low or zero weights, they induce a negligible bias and can be used as pure controls in a SCM.
- If some of the potentially affected units receive high weights, proceed with step 2.

2. Implement the "restricted" SCM, i.e., excluding units potentially affected by the intervention and:

   (a) Compare the bias in terms of predictors $(X_1 - X_0\widehat{W})$ between "restricted" SCM and the "unrestricted" SCM;

   (b) Compare Root Mean Squared Prediction Errors (RMSPEs) in the pre-intervention period of the "restricted" SCM and "unrestricted" SCM.

$$RMSPE = \left(\frac{1}{T_0} \sum_{t=1}^{T_0} (Y_{1t} - \sum_{j \neq 1} \widehat{w}_j Y_{jt})^2 \right)^{1/2}.$$

- If $(X_1 - X_0^{res}\widehat{W}^{res}) < (X_1 - X_0\widehat{W}^{unres})$ and $RMSPE^{res} < RMSPE^{unres}$, use the "restricted" SCM.
- If $(X_1 - X_0^{res}\widehat{W}^{res}) > (X_1 - X_0\widehat{W}^{unres})$ and/or $RMSPE^{res} > RMSPE^{unres}$, we advise using our iSCM.

Empirical Application

We estimate the causal effect of German reunification on West Germany's per capita GDP allowing for spillover effects from West Germany to Austria.

Applying iSCM with only 1 potentially affected unit requires the following steps:

- Construct Synthetic West Germany.
- Estimate the biased treatment effec $\boldsymbol{\hat{\theta}_t}$ and the weight assigned to Austria $\boldsymbol{\hat{w}_A}$.
- Construct synthetic Austria, including West Germany in the donor pool.
- Estimate the bias spillover effect $\boldsymbol{\hat{\gamma}_t}$ and the weight assigned to West Germany $\boldsymbol{\hat{l}_{WG}}$.

Country	Synthetic West Germany Weights	Synthetic Austria Weights	Restricted Synthetic West Germany Weights
West Germany	-	0.33	-
Austria	0.42	-	-
Australia	0	0	0
Belgium	0	0.12	0.12
Denmark	0	0	0
France	0	0	0
Greece	0	0	0
Italy	0	0	0
Japan	0.16	0.21	0.22
Netherlands	0.09	0.30	0.30
New Zealand	0	0	0
Norway	0	0.03	0
Portugal	0	0	0
Spain	0	0	0
Switzerland	0.11	0	0.09
UK	0	0	0
USA	0.22	0	0.40

The effects are then estimated as:

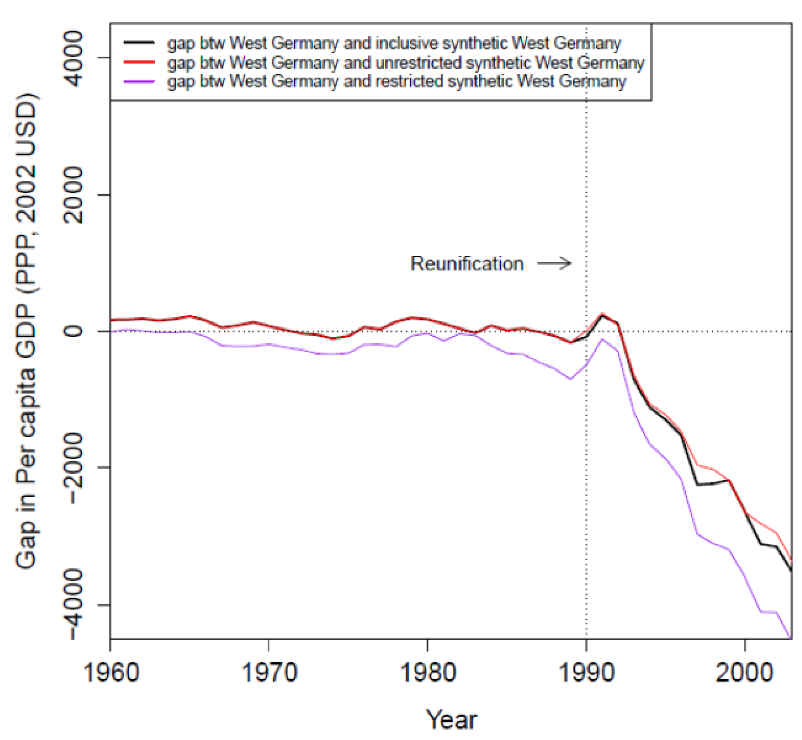
$$\frac{\det(\Omega_{WG,t})}{\det(\Omega)} = \frac{\hat{\theta}_t + 0{,}42\hat{\gamma}_t}{1 - 0{,}42(0{,}33)} \qquad \frac{\det(\Omega_{A,t})}{\det(\Omega)} = \frac{\hat{\gamma}_t + 0{,}33\hat{\theta}_t}{1 - 0{,}42(0{,}33)}$$

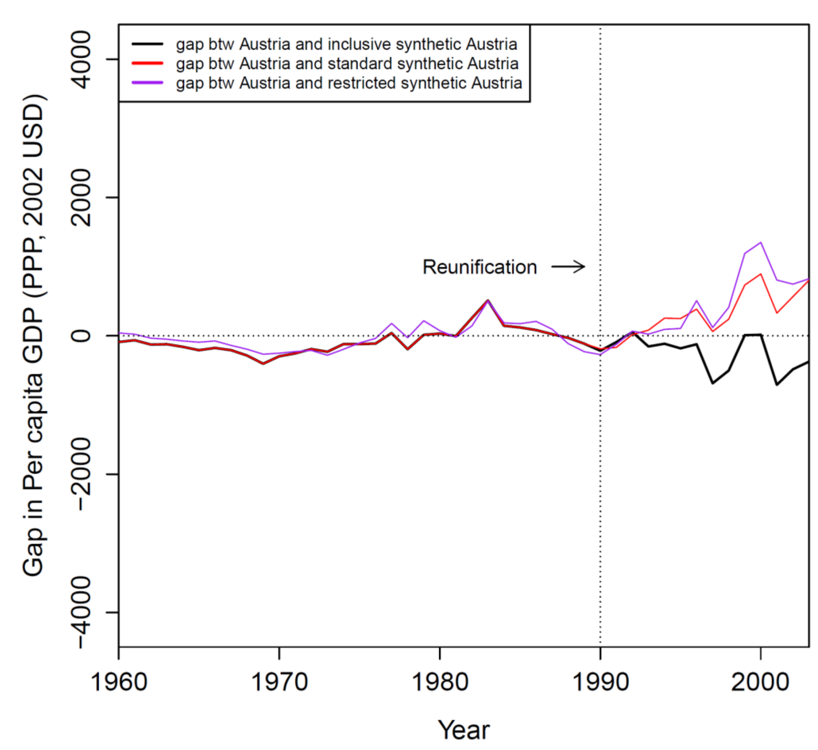
Table 4: Treatment Effects on West Germany

	$\widehat{\theta}_t^{unresSCM}$	$\widehat{\theta}_t^{iSCM}$	$\widehat{\theta}_t^{iSCM} - \widehat{\theta}_t^{unresSCM}$	$\widehat{\theta}_t^{resSCM}$
1990	7.58	-83.21	-90.79	-490.71
1991	268.31	229.89	-38.42	-113.50
1992	87.90	111.03	23.13	-291.56
1993	-642.23	-707.14	-64.91	-1187.41
1994	-1064.13	-1112.46	-48.33	-1656.09
1995	-1216.99	-1293.31	-76.32	-1860.16
1996	-1473.30	-1524.64	-51.34	-2169.88
1997	-1960.38	-2249.24	-288.86	-2970.69
1998	-2020.74	-2232.20	-211.47	-3104.82
1999	-2181.48	-2177.89	3.59	-3194.84
2000	-2645.30	-2638.79	6.51	-3595.01
2001	-2815.12	-3113.22	-298.10	-4109.12
2002	-2951.69	-3155.55	-203.86	-4116.73
2003	-3372.36	-3529.42	-157.06	-4559.62

Table 5: Treatment Effects on Austria

	$\widehat{\gamma}_t^{unresSCM}$	$\widehat{\gamma}_t^{iSCM}$	$\widehat{\gamma}_t^{iSCM} - \widehat{\gamma}_t^{unresSCM}$	$\widehat{\gamma}_t^{resSCM}$
1990	-188.12	-215.83	-27.71	-271.68
1991	-167.88	-91.33	76.55	-123.23
1992	18.01	54.98	36.97	67.13
1993	81.18	-154.29	-235.47	24.37
1994	255.56	-114.88	-370.43	93.01
1995	249.23	-181.43	-430.66	106.43
1996	385.64	-122.04	-507.69	508.32
1997	62.32	-686.65	-748.97	119.63
1998	240.62	-502.68	-743.29	410.14
1999	733.73	8.52	-725.21	1190.19
2000	894.16	15.48	-878.68	1350.67
2001	328.05	-708.61	-1036.66	805.76
2002	566.15	-484.60	-1050.76	747.11
2003	801.89	-373.36	-1175.25	823.97





Contacts

roberta.distefano@uniroma1.it giome@sam.sdu.dk