Identify More, Observe Less: Mediation Analysis Synthetic Control

by

Giovanni Mellace and Alessandra Pasquini

Discussion Papers on Business and Economics
No. 12/2019

FURTHER INFORMATION
Department of Business and Economics
Faculty of Business and Social Sciences
University of Southern Denmark
Campusvej 55, DK-5230 Odense M
Denmark
E-mail: lho@sam.sdu.dk / http://www.sdu.dk/ivoe

ISSN 2596-8157

Identify More, Observe Less: Mediation Analysis Synthetic Control

Giovanni Mellace* and Alessandra Pasquini**

*University of Southern Denmark; **Tor Vergata University

October 17, 2019

Abstract: The synthetic control method (SCM) allows estimation of the causal effect of an intervention in settings where panel data on just a few treated units and control units are available. We show that the existing SCM as well as its extensions can be easily modified to estimate how much of the "total" effect goes through observed causal channels. The additional assumptions needed are arguably very mild in many settings. Furthermore, in an illustrative empirical application we estimate the effects of adopting the euro on labor productivity in several countries and show that a reduction in the Economic Complexity Index helped to mitigate the negative short run effects of adopting the new currency in some countries and boosted the positive effects in others.

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 Simone De Angelis and participants at several seminars, workshops, and conferences. Addresses for correspondence: Giovanni Mellace (giome@sam.sdu.dk) and Alessandra Pasquini (alessandra.pasquini@uniroma2.it).

1 Introduction

The Synthetic Control Method (SCM) introduced by Abadie and Gardeazabal (2003), and further developed in Abadie et al. (2010, 2015) is becoming very popular in program evaluation. SCM is attractive, as it allows estimating the causal effect of an intervention even when data on only one treated and a few control units are available. This is possible by using information on the pre-intervention period to construct a "synthetic control", which mimics what would have happened to the treated unit in the post-intervention period in the absence of the intervention. Gobillon and Magnac (2016) compare SCM to other interactive fixed-effects models and find that it performs very well as soon as in 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 proposes a new method that combines both approaches. Doudchenko and Imbens (2016) propose a modification of SCM where the weights are not constrained to be positive and do not necessarily add up to 1. Ben-Michael et al. (2018) extend it to relax weights constraints and to correct for possible covariate imbalance, demonstrating that the synthetic control method can be seen as an inverse propensity score weighting estimator. Finally, Athey et al. (2017) propose a new method that includes synthetic control and other panel data methods as a special case.

Although all those methods are very well suited for estimating the "total" effect of an intervention, they are mostly uninformative about the causal mechanisms that generated this effect. Often an intervention may first have an impact on an intermediate outcome (hereafter referred to also as the "mediator"), which induces an impact on the final outcome. In the presence of a mediator, the total effect can generally be decomposed into a direct effect of the intervention and an indirect effect generated through it. Policy conclusions that ignore the presence of such intermediate outcomes might be misleading. Indeed, if the direct and indirect effects are both large and similar in magnitude but with opposite signs, one might wrongly conclude that the policy had no impact, by just looking at the total effect. Moreover, it is often important to quantify the direct and indirect effects to better target the intervention. Consider the huge decrease in tobacco consumption after the introduction of California's anti-tobacco law, Proposition 99, estimated in Abadie et al. (2010). Proposition 99 not only increased the 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 prices and how much of it is due to investments in informational campaigns.

The analysis of direct and indirect effects of an intervention may give additional information on macroeconomic mechanisms as well. Consider the impact of euro adoption on labor productivity. It may be highly relevant to determine whether an increase/decrease in labor productivity is due to a variation in the level of specialization of exporting firms and in the type of goods exported or to other factors. As we will show in our empirical application, estimating only the total effect could be highly misleading in this framework.

Mediation analysis is a standard approach to deal with these kind of issues. The main challenge in mediation analysis is that the identification of the direct and indirect effects requires knowledge about the potential outcome an individual would get if the potential mediator was set to the value it would have taken under the opposite treatment status than the one observed. This is never observed for any individual. A large part of the literature focuses on identification and estimation of direct and indirect effects 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, 2017). The idea behind this approach is that once we control for observed characteristics and the conditionally independent treatment the potential outcomes are independent of the potential mediators. To the best of our knowledge, none of the existing methods are specifically designed for panel data and cannot directly be applied in settings with one or a few treated and a few control units.[1]

This motivates the introduction of our mediation analysis synthetic control (MASC)

[1] One exception is Deuchert et al. (2018), however, who consider a framework with a randomized intervention with non-perfect compliance.

method, a generalization of the SCM that allows decomposing the total effect of an intervention into its direct and indirect components. As we will discuss below, in contrast to the standard framework, where the additional assumptions required to identify the direct and indirect effects are usually much stronger, MASC requires virtually the same assumptions on unobserved confounders as a standard SCM. The only additional assumption is that we are able to find control units with mediator values similar to those of the treated unit. Moreover, the plausibility of the assumptions required for MASC can be judged in a similar manner by checking the overlap in pre-treament outcomes in a standard SCM.

MASC can be easily implemented by using existing SCM algorithms and any of the new extensions. Indeed, as we will discuss in more detail below, to identify the direct and indirect effects MASC re-weights control unit post-intervention outcomes by choosing weights that minimize the distance between treated and synthetic unit in pre-intervention observable characteristics (including pre-intervention values of the outcome and the mediator) as well as in post-intervention values of the mediator. Intuitively, this allows us to mimic what would have happened to the treated unit in absence of the intervention if the mediator value were set to the potential mediator under treatment. As we mentioned above this is the main challenge of mediation analysis. 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 proposes possible inference procedures; Section 4 includes an empirical application to the introduction of the euro; and Section 5 concludes. All the technical proofs are relegated to the online appendix.

2 The Mediation Analysis Synthetic Control Method

Assume that we are interested in the effect of an intervention, D, implemented at time T, on an outcome, Y. Suppose that part of the effect of D on Y goes through an observed intermediate outcome (mediator), M. The total effect of the intervention on the final outcome can be decomposed into an indirect effect, which goes through M, and a residual effect, commonly known as the "direct effect", which could also go through other, possibly unobserved, causal pathways. Although often crucial for policy conclusions, identification of the direct and the indirect effects may be challenging. To see this, let D_{it} be a binary indicator that is equal to one if unit i is exposed to the intervention at time t. We will refer to units that are exposed to the intervention as "treated" and to those that are not exposed as "control". Using the potential outcome framework (see, e.g., Rubin 1974), for each unit, i, we can define the potential mediator at time t as follows:

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

$M_{it}(d)$ is the value that the mediator of unit i would take, at time t, if D_{it} is 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 period, $t < T$, and that we can observe only one of the two potential mediators for each unit in the post-intervention period, $t \geq T$.

Similarly, for each unit i at time t, we define the potential outcomes as

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

$Y_{it}^{d,d'}$ 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 outcome is a function of both the treatment and the potential mediator. Under SUTVA, and assuming no anticipation effects in the pre-intervention period, 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, we have for each unit four instead of two potential outcomes. As usual, only one between $Y_{it}^{0,0}$ and $Y_{it}^{1,1}$ can be observed for each unit in each period, while $Y_{it}^{0,1}$ and $Y_{it}^{1,0}$ are never observed for any unit in any period. Assuming no anticipation effects, $Y_{it} = Y_{it}^{0,0}$ in the pre-intervention period for all units.

Following the synthetic control literature, we will define our parameters of interest with respect to a single treated unit. This is in contrast with the standard mediation analysis 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 units(Vansteelandt and VanderWeele 2012; Huber et al. 2017). Nonetheless, if more than one unit is exposed to the intervention (see (Gobillon and Magnac, 2016; Adhikari, 2015)) our method can be easily used to decompose the average treatment effect on the treated.

We assume that we observe J units ordered such that units 1 through n are treated, while units $n + 1$ through J are controls. Without loss of generality, we will present our results for the first treated unit, unit 1, only. Since we have four potential outcomes instead of two, we can now define more parameters than in the standard synthetic control framework each measuring the effect implied by a different thought experiment. Indeed, each potential outcome represents a different state of the world, and one can in principle define effects by calculating the difference between a pair of potential outcomes. Intuitively, $Y_{1t}^{0,0}$ and $Y_{1t}^{1,1}$ measure the value that the outcome of the first treated unit would take with and without intervention. On the other hand, $Y_{1t}^{0,1}$ and $Y_{1t}^{1,0}$ measure the values that the outcome of the treated unit would take if the value of the mediator was pushed to the value it would take under the opposite treatment status. Intuitively, given that policy makers typically cannot choose which value the mediator takes, $Y_{1t}^{0,1}$ is arguably a more interesting counterfactual than and $Y_{1t}^{1,0}$. Indeed, to reproduce $Y_{1t}^{0,1}$ policy makers would need to push M to $M(0)$ in the absence of the intervention by, for example, implementing alternative policies that target the mediator directly. In contrast to reproduce $Y_{1t}^{1,0}$ they would need to implement the intervention and at the same time push M to $M(0)$, neutralizing the effect on the mediator. This would require implementing the intervention simultaneously with additional policies that have the opposite effect on the mediator.

The effects of interest with regard to unit 1 are the total effect, α_{1t}, which compares the outcomes the treated unit would get with and without the intervention; the direct effect, $\theta_{1t}(M_{1t}(1))$, which compares the treated potential outcome *with* the intervention and the outcome *without* intervention but where the mediator is set to the value it would have taken with the intervention; and the indirect effect, $\delta_{1t}(0)$, which measures the effect of pushing the mediator to its level under the intervention but without implementing the intervention. All parameters are assumed to be zero in the pre-intervention period and are in the post-intervention period defined as

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

It is easy to see that the total effect, α_{1t}, can be decomposed into[2]:

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

The decomposition above shows that if α_{1t} is identified, identifying $\theta_{1t}(M_{1t}(1))$ automatically implies identification of $\delta_{it}(0) = \alpha_{1t} - \theta_{1t}(M_{1t}(1))$.

The idea behind 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. In other words, SCM creates a synthetic value of $Y_{1t}^{0,0}$ in the post-intervention period. This is done by re-weighting the post-treatment outcomes of control units by using weights that are chosen so that they minimize the distance between the pre-intervention observable characteristics (including pre-intervention outcomes) of the treated and synthetic units. The main assumption is that $Y_{1t}^{0,0}$ lies in the convex hull of the non-treated post-intervention outcomes. Thus, it can be written as a linear combination of the latter.

MASC generalizes this idea to create "synthetic" values of $Y_{1t}^{0,1}$ in the post intervention period. 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 choosing the weights that minimize the distance between treated and synthetic with respect to post-treatment values of the mediator as well will mimic what would have

[2]In the mediation literature the following alternative decomposition is often also considered:

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

where $\delta_{1t}(1) = Y_{1t}^{1,1} - Y_{1t}^{1,0}$ and $\theta_{1t}(M_{1t}(0)) = Y_{1t}^{1,0} - Y_{1t}^{0,0}$. We decide not to focus on this decomposition for two reasons. First, as we argue above, a policy maker would need to be able to neutralize the effect of the treatment on the mediator to reproduce $Y_{1t}^{1,0}$. Second, identification of $Y_{1t}^{1,0}$ requires additional assumptions and the ability to observe more than one treated unit.

happened to the treated unit in the absence of the intervention but fixing the mediator value to the its value with the intervention, $M_{1t}(1)$.

In both MASC and SCM, the main identification assumption is that the unobserved confounders are either time invariant or, if time varying, they change in the same way for all units. To further illustrate our approach, in the spirit of Abadie et al. (2010), 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 a $(v \times 1)$ 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 an $(r \times 1)$ vector of observed covariates that includes all the variables included in Z_i but might also include other observable variables, which affects the treatment and the outcome but not the mediator; 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

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

As mentioned above, for the total effect we can just use the standard SCM. In particular, we assume that there exists a $(1 \times (J-n))$ vector of weights $L^* = (l^*_{n+1}, ..., l^*_J)$ that are positive, adding up to 1, and such that in the post-intervention period

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

As in Abadie et al. (2015) we assume that $\forall\ t = 1, ..., T-1$, L^* also satisfies

$$\begin{aligned}
\sum_{j=n+1}^{J} l^*_j Y_{jt} &= Y_{1t}, \\
\sum_{j=n+1}^{J} l^*_j M_{jt} &= M_{1t}, \\
\sum_{j=n+1}^{J} l^*_j X_j &= X_1.
\end{aligned}$$

This justifies the choice of weights that minimize the distance between the observable characteristics of the treated unit and the control units in the pre-treatment period. More formally, let $\Omega^{\alpha}_1 = (X_1, Y_{11}, \ldots, Y_{1,T-1}, M_{11}, \ldots, M_{1,T-1})$ be a $((2(T-1)+r) \times 1)$ vector, $\omega^{\alpha}_{0i} = (X_i, Y_{i1}, \ldots, Y_{i,T-1}, M_{i1}, \ldots, M_{i,T-1})$ be a $(1 \times (2(T-1)+r))$ vector, and

$\Omega_0^\alpha = (\omega_{0,n+1}^\alpha, \ldots, \omega_{0J}^\alpha)'$. Then

$$L^* = \min_{l_{n+1},\ldots,l_J} ||\Omega_1^\alpha - L\Omega_0^\alpha||$$
$$s.t. \quad l_{n+1} \leq 0, ..., l_J \leq 0, \sum_{i=n+1}^{J} l_i = 1,$$

where $||\Omega_1^\alpha - L\Omega_0^\alpha|| = \sqrt{(\Omega_1^\alpha - L\Omega_0^\alpha)'(\Omega_1^\alpha - L\Omega_0^\alpha)}$. It is also possible to give more weight to specific observable characteristics by using the alternative distance $||\Omega_1^\alpha - L\Omega_0^\alpha||_V = \sqrt{(\Omega_1^\alpha - L\Omega_0^\alpha)' V (\Omega_1^\alpha - L\Omega_0^\alpha)}$ (see Abadie et al. 2010, for a data driven procedure to choose V).

Let $\hat{Y}_{1t}^{0,0} = \sum_{i=n+1}^{J} l_i^* Y_{it}$, Abadie et al. (2010) show that, if L^* exists, for $t \geq T$

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

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

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

The estimation of $Y_{1t}^{0,1}$ in MASC requires additional constraints but no extra assumptions on the unobservable in the potential outcomes equations. Our goal is to construct a "synthetic" unit, which is identical to the treated unit, 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 weights $W_t^* = (w_{n+1,t}^*, ..., w_{Jt}^*)$ that are positive, adding up to 1, and such that in the post-intervention period

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

Notice that, in our simple factor model, $Y_{1t}^{0,1}$ depends on the value that M takes at time t only. [3] Also notice that the weights need to be calculated in each post-intervention period

[3]It is easy to let $Y_{1t}^{0,1}$ depend on all the values that the mediator takes between T and t. This is done by replacing $\Omega_1^{\theta_{t'}(1)}$ and $\omega_{0i}^{\theta_{t'}(1)}$ defined below

in this model. Let $t' \geq T$ be the time at which we want to estimate the direct effect. Similar to Abadie et al. (2010) we assume that $W^*_{t'}$ exists and satisfies $\forall\ t = 1, ..., T-1$:

$$\begin{aligned}
\sum_{j=n+1}^{J} w^*_{jt'} Y_{jt} &= Y_{1t}, \\
\sum_{j=n+1}^{J} w^*_{jt'} X_{j} &= X_{1},
\end{aligned}$$

and $\forall\ t = 1, ..., T-1, t'$,

$$\sum_{j=n+1}^{J} w^*_{jt'} M_{jt} = M_{1t}.$$

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

$$\begin{aligned}
W^*_{t'} &= \min_{w_{n+1,t'}, \ldots, w_{Jt'}} ||\Omega_1^{\theta_{t'}(1)} - W_{t'}\Omega_0^{\theta_{t'}(1)}||_V \\
&\quad s.t.\ \ w_{n+1,t'} \leq 0, ..., w_{Jt'} \leq 0, \sum_{i=n+1}^{J} w_{it'} = 1,
\end{aligned}$$

where $||\Omega_1^{\theta_{t'}(1)} - W_{t'}\Omega_0^{\theta_{t'}(1)}||_V = \sqrt{\left(\Omega_1^{\theta_{t'}(1)} - W_{t'}\Omega_0^{\theta_{t'}(1)}\right)' V \left(\Omega_1^{\theta_{t'}(1)} - W_{t'}\Omega_0^{\theta_{t'}(1)}\right)}$. Notice that we only have one mediator in the post-intervention period and several pre-intervention variables. Thus, we suggest to choose V such that equal weights are given to pre- and post- intervention information.

Let $\hat{Y}^{0,1}_{1t'} = \sum_{i=n+1}^{J} w^*_{it'} Y_{it'}$, as we show in the appendix, if $W^*_{t'}$ exists, under standard regularity conditions:

$$E(\hat{Y}^{0,1}_{1t'}) = Y^{0,1}_{1t'} + o(T).$$

This allows us to estimate the direct effect as $\theta_{1t'}(M_{1t}(1))$ and the indirect effect as $\delta_{it'}(0)$

with $\Omega_1^{\theta_{t'}(1)} = (X_1, Y_{11}, \ldots, Y_{1,T-1}, M_{11}, \ldots, M_{1,T-1}, M_{1,T}, \ldots, M_{1,t'})$ and $\omega_{0i}^{\theta_{t'}(1)} = (X_i, Y_{i1}, \ldots, Y_{i1,T-1}, M_{i1}, \ldots, M_{i,T-1}, M_{i,T}, \ldots, M_{i,t'})$, respectively.

since

$$\hat{\theta}_{1t'}(M_{1t'}(1)) = Y_{1t'} - \hat{Y}^{0,1}_{1t'}, \qquad \hat{\delta}_{1t'}(0) = \hat{\alpha}_{1t'} - \hat{\theta}_{1t'}(M_{1t'}(1)),$$

respectively, as it implies

$$\lim_{T\to\infty} E(\hat{\theta}_{1t'}(M_{1t}(1)_{1t}) = \theta_{1t'}(M_{1t}(1)_{1t}) \quad \forall \quad t \geq T,$$
$$\lim_{T\to\infty} E(\hat{\delta}_{it'}(0)) = \delta_{it'}(0) \quad \forall \quad t \geq T$$

Intuitively, $W^*_{t'}$ only exists if, in addition to the assumptions needed for a standard SCM, there is also overlap in the post-intervention values of the mediation. Similarly to the standard SCM, the plausibility of the existence of $W^*_{t'}$ can be graphically assessed by looking at the overlap in the pre-intervention period between the observed outcome, Y_{1t}, and the synthetic outcome $\hat{Y}^{0,1}_{1t}$.

3 Inference

Inference can be carried over in a similar manner as for the standard synthetic control method. For example, one can run similar placebo tests as the one suggested in Abadie et al. (2015), estimating the effects (in our case also the direct and indirect effects) of the intervention either before its implementation or for units not exposed to it. Abadie et al. (2015) criticize the former type of placebo tests (often called in-time placebos), arguing that there may be other shocks in the past affecting treated and control units differently. They suggest to base inference on the ratio between post- and pre- intervention root mean square prediction error (RMSPE). For unit i and synthetic outcome $\hat{Y}^{d,d'}$ the pre-intervention RMSPE can be defined as

$$RMSPE_i^{\hat{Y}^{d,d'},pre} = \frac{\sum_{t=1}^{T-1}(Y_{i,t} - \hat{Y}^{d,d'}_{i,t})^2}{T-1}.$$

The post RMSPE is defined similarly

$$RMSPE_i^{\hat{Y}^{d,d'},post} = \frac{\sum_{t=T}^{t'}(Y_{i,t'} - \hat{Y}_{i,t'}^{d,d'})^2}{t' - T - 1}.$$

The test statistic can then be defined as

$$Test_i^{Y^{d,d'}} = \frac{RMSPE_i^{\hat{Y}^{d,d'},post}}{RMSPE_i^{\hat{Y}^{d,d'},pre}}.$$

Under the null hypothesis of a zero effect of the intervention the distribution of $Test_i Y^{d,d'}$ can be calculated using the non-treated units, as, for example, in Firpo and Possebom (2018) and Chernozhukov et al. (2018).

Another possibility is to follow the approach outlined in Chernozhukov et al. (2018) which can be easily adapted to our framework. Firpo and Possebom (2018) generalize this method and show that it performs better than other classical inference methods. Ferman and Pinto (2017) reconsider the ratio of post- and pre-RMSPE and show that it performs better and is less sensitive to violations of the assumptions than the post-intervention RMSPE on its own. The authors propose a different test statistic which is robust serial correlation in the temporary shocks. A similar procedure has been proposed by Chernozhukov et al. (2018).

Yet another inference procedure is described in Gobillon and Magnac (2016). This procedure is based on different steps. First of all, the outcome of the treated unit is reduced by the treatment effect. In our framework, the mediator of treated unit should be reduced as well. this is easily done using standard SCM to estimate of the potential mediator in the absence of the intervention. Then, 10,000 samples without replacement of treated units are drawn from 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 MASC. Finally, the estimated values are used to estimate the distribution of the treatment effects. 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 the total effects 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.
   - 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 this inference procedure, just as in Abadie et al. (2015), is based on the strong assumption that the disturbances across units are exchangeable. Indeed, the basic ideas behind these methods is that the error terms of the placebos can be used to approximate the error term of the treated unit.

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 each of the units in the donor pool one by one (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 not sensitive to the chosen donor pool and can be considered to be robust.

4 Decomposing the Impact of Adopting the Euro on Productivity: A Backstage Story

In this section we use MASC to estimate the causal effect of adoption of the euro on labor productivity in several European countries and investigate the role of the Economic Complexity Index as a possible causal mechanism. Several studies on the impact of euro adoption focus on trade outcomes. Rose (2000) finds a huge effect of being in a monetary union on trade in ex-ante analysis. This study has been widely criticized because its results, obtained from small and low-income countries, were extrapolated to big and high-income countries. Subsequent studies, based on ex-post analysis, found lower, although still positive and significant, impacts (Micco et al. 2003; Flam and Nordström 2006; Mancini-Griffoli and Pauwels 2006; De Nardis and Vicarelli 2003; Bun and Klaassen 2007; Berger and Nitsch 2008; Chintrakarn 2008; Saia 2017) with the exception of Silva and Tenreyro (2010), who found a positive although non-significant effect (see also Baldwin and Taglioni 2007 and Baldwin et al. 2008 for two literature reviews on gravity models).

Other studies focus on the impact of euro adoption on GDP. Pesaran et al. (2007) and Žúdel and Melioris (2016) find positive effects for, respectively, the UK and Sweden (in the hypothetical scenario of euro adoption by these two countries) and Slovakia. Puzzello and Gomis-Porqueras (2018), who apply the SCM to multiple European countries, find heterogeneous effects instead. According to their results, euro adoption had a negative effect for Belgium, France, Germany, and Italy, while Ireland benefited from the common currency. No significant impact is found for the Netherlands. Differently from previous studies, Gabrielczak and Serwach (2017) use the SCM to estimate the impact of adopting the euro on the economic complexity of Slovenia exports showing that it increased after euro adoption.

Our first contribution is to estimate the causal effect of euro adoption on labor productivity. To the best of our knowledge, ours is the first study focusing on labor productivity. As the introduction of a common currency can to a certain extent be considered as a liberalization policy, since it facilitates trade by reducing costs and uncertainties due to volatility in the exchange rates, our analysis contributes to the broader literature on the impact of liberalization on labor productivity. Second, by using MASC we are able to investigate an important causal channel, namely the economic complexity index. This gives us a better understanding of how countries reacted to the introduction of the euro. In particular, we show that to deal with the more competitive environment induced by the common currency the economy in most of the countries experienced an increase in the degree of specialization that translated into an increase in labor productivity.

4.1 The introduction of the euro

The first commitment among some European countries to create an economic and monetary union was made in 1971. The countries involved were the European Community members at that time, namely: France, West Germany, Italy, Belgium, Luxembourg, and the Netherlands. Other countries joined in the following years, namely Denmark, UK, and Ireland in 1973; Greece in 1981; and Spain and Portugal in 1986. In 1998 the euro was introduced officially and the exchange rates of the participants were fixed. The introduction involved almost all European Community countries, with the exceptions of Denmark and the UK, which decided not to participate (Denmark pegged its national currency to the euro), and Greece and Sweden, which did not meet the standards required for joining the common currency.[4] The actual introduction of the euro was not completed until 2002 (see Puzzello and Gomis-Porqueras 2018 for more details on the process).

4.2 Data and MASC Implementation

We use country level panel data over the period 1986-2007. The treatment period starts in 1998 with the official introduction of the euro. Our outcome is labor productivity measured as output per hour worked. Data on labor productivity were taken from Penn World Table version 9.0. To measure economic complexity (our mediator), we follow Gabrielczak and Serwach (2017) and use the Economic Complexity Index (Hidalgo and Hausmann 2009; Hausmann et al. 2014). The Economic Complexity Index is built and provided by the Atlas of Economic Complexity of the Center for International Development at Harvard University. To build the index, the complexity of each product is defined according to the number of countries exporting it and their level of export complexity (http://atlas.cid.harvard.edu/rankings/).

We implement MASC separately on five different countries adopting the euro in 1998: Belgium, France, Ireland, Italy, and the Netherlands. We exclude Germany becauseof

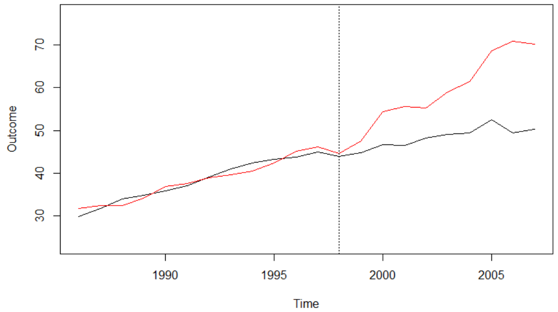
[4]Greece joined in 2001, while Sweden decided not to participate after a referendum in 2003.

the German reunification implies that the assumptions of our method as well as synthetic control are likely to be violated. We furthermore exclude Luxembourg for data availability reasons and Portugal, Spain, Greece, Austria, and Finland because the pre-intervention overlap was poor in those countries. Details on the covariates as well as the donor pool we used for each treated country are reported in appendix E.

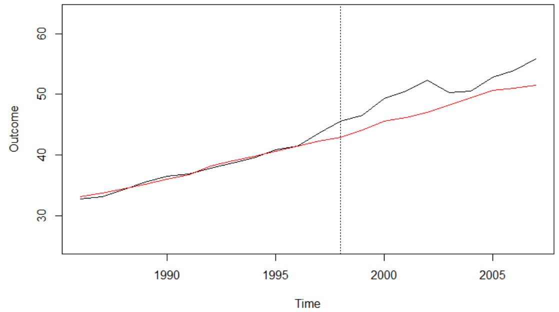
To guarantee a good fit of the mediator when estimating the total effect we select the weights to minimize both the root mean squared prediction error of the pre-treatment outcomes and the one of the pre-treatment mediators. For the direct effect, instead, we use a similar procedure, but we assign half of the weights to post- intervention constraints (notice that the post- intervention constraints to identifying the direct effect at time t' were imposed on the mediator over the period T to t'). For the calculation of the direct effect we use a single year lag between the mediator and the outcome, but our results are robust to the choice of different time lags. For inference we follow Firpo and Possebom (2018) and derive the p-values from the placebo tests proposed in Abadie et al. (2015).

4.2.1 Results and Discussion

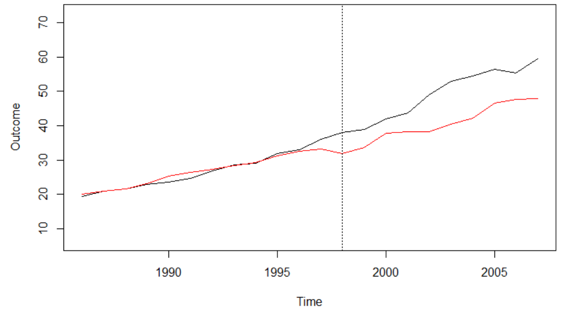
Figures 1 and 2 compare trends in the outcomes of each treated unit and their synthetic controls constructed for calculating the total effect in the upper panel and the direct effect in the lower one. The two figures show good pre-treatment fits for all countries. The pre- and post-treatment fits for the mediator are fairly good for most countries (graphs are available upon request). By graphical inspection we can already see that the total effect is negative for Belgium and Italy, positive for France and Ireland, and close to zero for the Netherlands. The indirect effects have a positive sign.



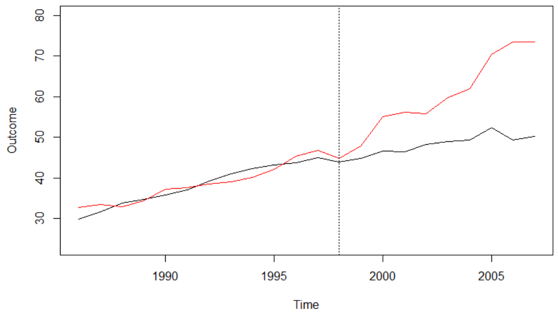
(a) Belgium



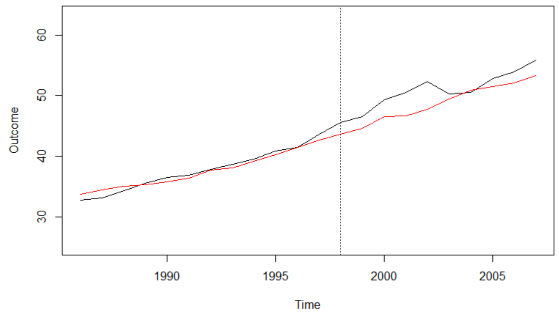
(b) France



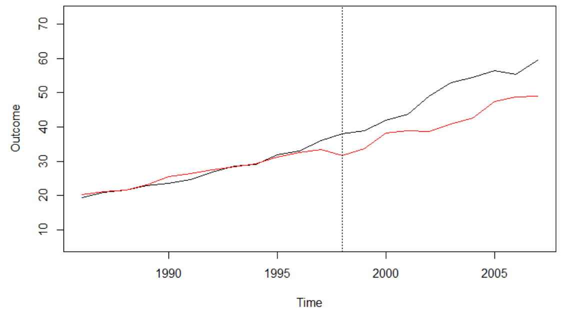
(c) Ireland



(d) Belgium

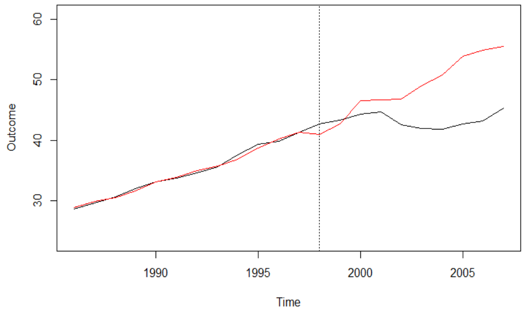


(e) France

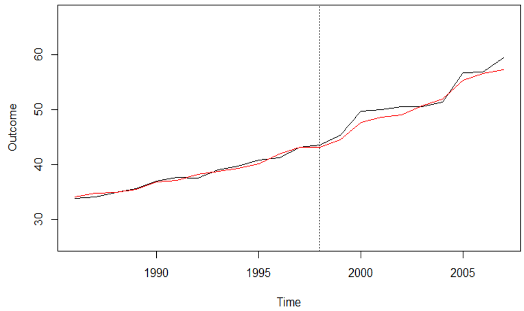


(f) Ireland

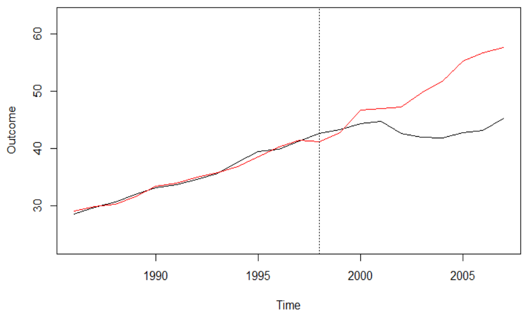
Figure 1: Differences in the outcome between treated and synthetic units $Y^{0,0}$. The black line represents the treated outcome. The red line represents the synthetic unit outcome.



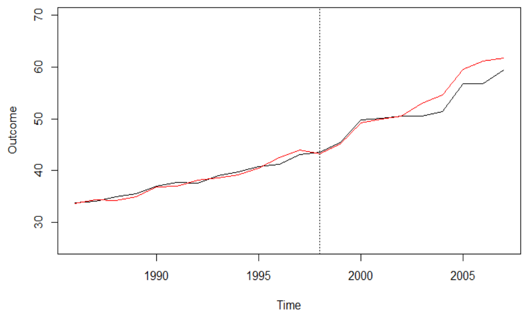
(a) Italy



(b) Netherlands



(c) Italy



(d) Netherlands

Figure 2: Differences in the outcome between treated and synthetic units $Y^{0,1}$. The black line represents the treated outcome. The red line represents the synthetic unit outcome.

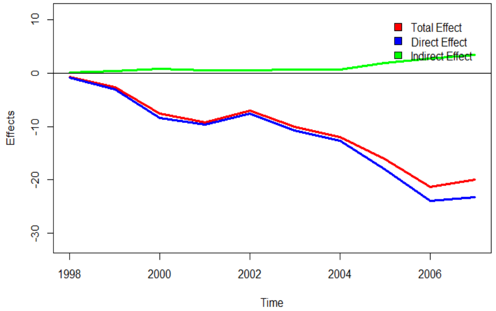
This is confirmed by the total, direct, and indirect effects estimates displayed in table 1 and figure 3. [5]

[5] All results are fairly robust to leaving one country out from the donor pool with few exceptions. For example, the results for Belgium and Ireland are sensitive to the exclusion of Norway. This is due to a lack of overlap in the pre-intervention period. See Appendix F.

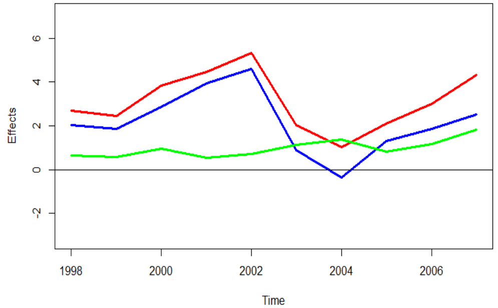
Table 1: Effects of Euro Adoption by Country

Year	Belgium			France			Ireland			Italy			Netherlands		
	Total	Direct	Indirect	Total	Direct	Indirect	Total	Direct	Indirect	Total	Direct	Indirect	Total	Direct	Indirect
1998	-0.71	-0.81	0.10	2.69***	2.02	0.66	6.26	6.35***	-0.09***	1.73***	1.48***	0.25***	0.46	0.33***	0.13***
P-val	(0.90)	(0.40)	(0.30)	(0.00)	(0.11)	(0.22)	(0.20)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.70)	(0.00)	(0.00)
1999	-2.68	-3.04	0.36	2.47	1.88	0.59	5.33	5.27	0.06	0.61	0.61	-0.01	0.83	0.27	0.56
P-val	(0.80)	(0.90)	(1.00)	(0.22)	(0.22)	(0.56)	(0.80)	(0.80)	(0.90)	(0.70)	(0.90)	(1.00)	(0.70)	(0.90)	(0.90)
2000	-7.60	-8.43*	0.82	3.83	2.88	0.95	4.12	3.67*	0.45***	-2.23***	-2.30***	0.07	2.1*	0.61	1.50
P-val	(0.20)	(0.10)	(0.40)	(0.11)	(0.22)	(0.56)	(0.30)	(0.10)	(0.00)	(0.00)	(0.00)	(0.60)	(0.10)	(0.50)	(0.40)
2001	-9.17	-9.70	0.52	4.47***	3.94	0.53	5.45	4.92	0.53	-1.88	-2.20	0.32	1.34	0.07	1.28
P-val	(0.70)	(0.60)	(0.90)	(0.00)	(0.22)	(0.44)	(0.80)	(0.60)	(0.20)	(0.20)	(0.30)	(0.70)	(0.80)	(1.00)	(0.80)
2002	-7.08	-7.60	0.51	5.32	4.60	0.72	10.76	10.35*	0.41	-4.21***	-4.63***	0.41	1.45*	-0.07	1.52
P-val	(0.20)	(0.20)	(0.80)	(0.11)	(0.11)	(0.56)	(0.20)	(0.10)	(0.20)	(0.00)	(0.00)	(0.50)	(0.10)	(1.00)	(0.80)
2003	-10.04***	-10.78***	0.74***	2.03	0.91	1.12	12.43***	11.90***	0.53***	-7.21***	-7.79***	0.58***	-0.11***	-2.46***	2.35***
P-val	(0.00)	(0.00)	(0.00)	(0.56)	(0.67)	(0.67)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
2004	-11.98***	-12.65***	0.67***	1.03	-0.36	1.39	12.35	11.91	0.44	-8.88	-9.78*	0.91	-0.51	-3.11	2.61
P-val	(0.00)	(0.00)	(0.00)	(0.56)	(0.56)	(0.67)	(0.90)	(0.80)	(0.40)	(0.20)	(0.10)	(0.70)	(1.00)	(0.70)	(0.90)
2005	-16.16	-18.09*	1.93*	2.12	1.30	0.82	9.69	8.89	0.8***	-11.19***	-12.53***	1.34***	1.33	-2.84	4.17***
P-val	(0.50)	(0.10)	(0.10)	(0.56)	(0.56)	(0.44)	(0.70)	(0.20)	(0.00)	(0.00)	(0.00)	(0.00)	(0.50)	(0.20)	(0.00)
2006	-21.31	-24.01	2.70	3.01	1.86	1.15	7.54	6.40	1.13	-11.76	-13.52*	1.76	0.22	-4.33	4.54
P-val	(0.80)	(0.80)	(0.90)	(0.33)	(0.56)	(0.67)	(0.90)	(0.80)	(0.40)	(0.20)	(0.10)	(0.70)	(1.00)	(0.70)	(0.90)
2007	-19.91***	-23.30***	3.40***	4.34	2.51	1.83	11.65***	10.47***	1.18***	-10.18***	-12.27***	2.09***	2.18***	-2.26***	4.44***
P-val	(0.00)	(0.00)	(0.00)	(0.11)	(0.44)	(0.56)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)

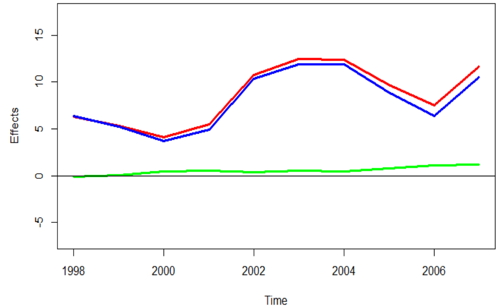
Table 2: P-values are displayed in brackets. * Significant at 10%. ** Significant at 5%. *** Significant at 1%.



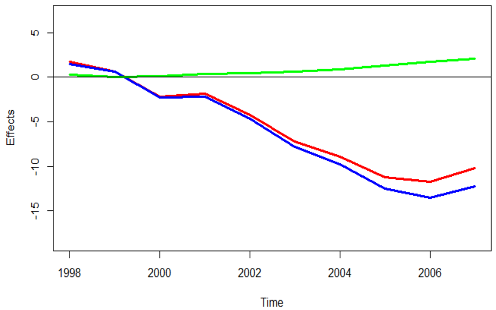
(a) Belgium



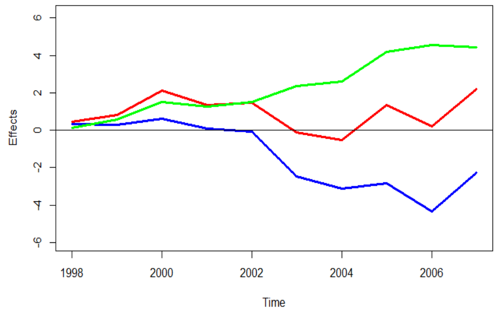
(b) France



(c) Ireland



(d) Italy



(e) Netherlands

Figure 3: Total (red), direct (blue), and indirect effects (green) evolution over time.

Our results show that Belgium and Italy saw a decrease in labor productivity, at least in the short run, while France and Ireland recorded a small increase. In the Netherlands the total effect was mostly small and non-statistically significant. In all countries a decrease in economic complexity helped facing the potentially detrimental short run effects of a tougher competition induced by the common currency. This can be explained using arguments from competition theory (see, e.g., Bayar 2002). Indeed, a common currency removes the trade risks deriving from changes in the exchange rates and decreases trade costs. Therefore, firms have to face a higher level of competition due to the reduced trade costs. Moreover, countries are not able to increase their competitiveness through currency depreciation. Hence, firms have to specialize in producing products at the highest productivity level. In addition, firms with a low productivity level might be forced out of the market. At a macro level, a higher level of specialization by some firms and the exit of unproductive ones would result in a lower complexity index and higher productivity levels. Another possible explanation can be found in the theories of economies of scale. Euro adoption made the realization of economies of scale easier, so that firms were able to specialize on their most competitive products, while importing the needed intermediate products from abroad (Barro and Tenreyro 2007). This, in turns, allowed them to increase their productivity level and lowered the economy complexity index. We find, in all countries but Ireland (and the Netherlands during the first years), a decrease in the economy complexity index and a positive indirect effect. Thus, the negative impact of euro adoption on the complexity index mitigated the potential short term negative impact of the introduction of a common currency (Belgium, Italy, and the Netherlands) or amplified its benefit (France and Ireland). [6]

The results of the direct effects estimations are more heterogeneous. Indeed, the direct effect is positive for France and Ireland and negative for Belgium, Italy, and the Netherlands. One possible explanation for these differences is that the economies in those countries are characterized by different returns of scale. The data on increasing returns displayed in Midelfart-Knarvik et al. (2000) corroborate this hypothesis. Indeed, France

[6] Results are available from the author upon request.

has the highest returns, followed by the Netherlands, Belgium, and Italy. Ireland's returns to scale was not particularly high, however, this country experienced a strong focus on productivity after the introduction of the euro (Petrakos et al. 2006).

In general, the fact that, for some countries, the direct and the indirect effects go in opposite directions, shows the advantages of using a method like MASC that allows decomposing the total effect. Indeed, looking exclusively at the total effects of adopting the euro on labor productivity would leave a policy maker with only partial evidence. In particular, looking at the total effect for the Netherlands for example, one might conclude that euro adoption did not have any impact on labor productivity. Our results suggest that the short run effect of joining the common currency would have been negative had the Dutch economy not reacted by increasing the level of specialization.

5 Conclusions

We introduced a new methodology called mediation analysis synthetic control (MASC). This method combines the synthetic control method (Abadie and Gardeazabal 2003; Abadie et al. 2010, 2015) with the mediation analysis approach and allows us to identify direct and indirect effects in frameworks with selection on unobservables and a low number of treated units and control units. This method is very intuitive and easy to implement (i.e., publicly available SCM algorithms can be employed). Even though introduced for the procedure presented in Abadie et al. (2010, 2015), it can be easily extended to new approaches, such as Athey et al. (2017); Xu (2017); Kreif et al. (2016); Ben-Michael et al. (2018); and Doudchenko and Imbens (2016). Finally, after estimating the "total" effect of the introduction of the euro on labor productivity in several European countries, we showed that an increase in the degree of specialization in those economies either helped to mitigate the potentially negative short run effects of adopting a common currency or amplified its positive effects.

Appendix

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 (and considering the factor model introduced in the text) the synthetic value of Y_{1t}^{01} is given by

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

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

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

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

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

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

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

Similar to Abadie et al. (2010) 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 and implies that $(\lambda^{P'}\lambda^P)^{-1}$ exists. We can then multiply both sides of A.2 by $(\lambda^{P'}\lambda^P)^{-1}\lambda^{P'}$ to get

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

Substituting in A.1 and considering a generic post-intervention period $t' \geq T$, we have

$$\begin{aligned}
Y_{1t'}^{0,1} - \sum_{j=n+1}^{J} w_jY_{jt'} &= \lambda_{t'}(\lambda^{P'}\lambda^P)^{-1}\lambda^{P'}\left(Y_1^P - \sum_{j=n+1}^{J} w_jY_j^P\right) \\
&+ \left(\eta_{t'} - \lambda_{t'}(\lambda^{P'}\lambda^P)^{-1}\lambda^{P'}\eta^P\right)\left(X_1 - \sum_{j=n+1}^{J} w_jX_j\right) \\
&- \lambda_{t'}(\lambda^{P'}\lambda^P)^{-1}\lambda^{P'}\left[\varphi^P(0)(M_1^P(0) - \sum_{j=n+1}^{J} w_jM_j^P(0))\right] \\
&+ \varphi_{t'}(0)\left(M_{1t'}(1) - \sum_{j=n+1}^{J} w_jM_{jt'}(0)\right) \\
&- \lambda_{t'}(\lambda^{P'}\lambda^P)^{-1}\lambda^{P'}\left(\epsilon_1^P - \sum_{j=n+1}^{J} w_j\epsilon_j^P\right) + \sum_{j=n+1}^{J} w_j(\epsilon_{1t'} - \epsilon_{jt'}).
\end{aligned}$$

If we now assume, as we did in the main text, that there exists a set of positive and summing up to 1 weights W^* that satisfies, $\forall\ t = 1, ..., T-1$

$$\begin{aligned}\sum_{j=n+1}^{J} w_j^* Y_{jt} &= Y_{1t}, \\ \sum_{j=n+1}^{J} w_j^* X_j &= X_1,\end{aligned}$$

and $\forall\ t = 1, ..., T-1, t'$, also satisfies

$$\sum_{j=n+1}^{J} w_j^* M_{jt} = M_{1t},$$

replacing in the post-intervention period, the generic weights with W^*, we get

$$Y_{1t'}^{0,1} - \sum_{j=n+1}^{J} w_j^* Y_{jt'} = -\lambda_{t'}(\lambda^{P'}\lambda^{P})^{-1}\lambda^{P'}\left(\epsilon_1^P - \sum_{j=n+1}^{J} w_j^* \epsilon_j^P\right) + \sum_{j=n+1}^{J} w_j^*(\epsilon_{1t'} - \epsilon_{jt'}).$$

From here, the proof is identical to the one 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

$$R_{1t'} = \lambda_{t'}(\lambda^{P'}\lambda^{P})^{-1}\lambda^{P'} \sum_{j=n+1}^{J} w_j^* \epsilon_j^P \tag{A.3}$$

$$R_{2t'} = -\lambda_{t'}(\lambda^{P'}\lambda^{P})^{-1}\lambda^{P'}\epsilon_1^P \tag{A.4}$$

$$R_{3t'} = \sum_{j=n+1}^{J} w_j^*(\epsilon_{jt'} - \epsilon_{1t'}) \tag{A.5}$$

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. *$E(\epsilon_{it}|X_i, \mu_i, M_{it}(I\{t \geq T\})) = E(\epsilon_{it}) = 0$ for $i \in \{1, n+1, ..., J\}$ and for $t = 1, ..., t'$*

Taking the expected value on both sides of A.4 we get

$$\begin{aligned} E(R_{2t'}) &= E(-\lambda_{t'}(\lambda^{P'}\lambda^P)^{-1}\lambda^{P'}\epsilon_1^P) \\ &= -\lambda_{t'}(\lambda^{P'}\lambda^P)^{-1}\lambda^{P'}E(\epsilon_1^P) \\ &= 0 \end{aligned}$$

where the second equality follows from the fact that $-\lambda_{t'}(\lambda^{P'}\lambda^P)^{-1}\lambda^{P'}$ is non-stochastic and the third equality follows from assumption 4. Taking the expectation on both sides of A.5

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

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 which under assumptions 2, 3 and 4 are independent from the error terms at time $t' \geq T$. The fourth equality follows from assumption 4. The remaining A.3 can be rewritten as:

$$R_{1t'} = \sum_{j=n+1}^{J} w_j^* \sum_{s=1}^{T-1} \lambda_{t'}(\sum_{h=1}^{T-1} \lambda_h' \lambda_h)^{-1} \lambda_s' \epsilon_{js} \tag{A.6}$$

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

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

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

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

Assumption 6.

$$\exists \quad \underline{\lambda} \text{ s.t. } |\lambda_{tf}| \leq \underline{\lambda} \quad \forall \ t{=}1,...,t' \text{ and } f{=}1,...,F \ .$$

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

$$\begin{aligned} \left(\lambda_t \left(\sum_{h=1}^{T-1} \lambda_h' \lambda_h \right)^{-1} \lambda_s' \right)^2 &= |\langle \lambda_t, A\lambda_s' \rangle|^2 \leq ||A\lambda_t||^2 ||A\lambda_s||^2 \\ &= \left(\lambda_t \left(\sum_{h=1}^{T-1} \lambda_h' \lambda_h \right)^{-1} \lambda_t' \right) \left(\lambda_s \left(\sum_{h=1}^{T-1} \lambda_h' \lambda_h \right)^{-1} \lambda_s' \right) \end{aligned} \tag{A.7}$$

Where $A = \left(\sum_{h=1}^{T-1} \lambda_h' \lambda_h \right)^{-1}$. Since A is a symmetric matrix $B = (T-1)A$ 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,

$$\lambda_t \left(\sum_{h=1}^{T-1} \lambda_h' \lambda_h \right)^{-1} \lambda_t' = \frac{1}{T-1}(\lambda_t B \lambda_t') = \frac{1}{T-1}(\lambda_t GOG' \lambda_t')$$

Defining $b_t = \lambda_t G$ we have

$$\lambda_t \left(\sum_{h=1}^{T-1} \lambda_h' \lambda_h \right)^{-1} \lambda_t' = \frac{1}{T-1}(b_t O b_t') = \frac{1}{T-1} \left(b_{t1}^2 \frac{1}{\varsigma_1} + \ldots + b_{tF}^2 \frac{1}{\varsigma_F} \right)$$

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:

$$\begin{aligned}\lambda_t\left(\sum_{h=1}^{T-1}\lambda_h'\lambda_h\right)^{-1}\lambda_t' &= \frac{1}{T-1}\sum_{f=1}^{F}\frac{b_{tf}^2}{\varsigma_f} \leq \frac{1}{(T-1)\underline{\varsigma}}\sum_{f=1}^{F}b_{tf}^2 \\ &= \frac{1}{(T-1)\underline{\varsigma}}||b_t||^2 = \frac{1}{(T-1)\underline{\varsigma}}||\lambda_t G||^2\end{aligned}$$

As we noticed before, G is an orthogonal and thus isometric matrix, hence $||\lambda_t G|| = ||\lambda_t||$. Consequently,

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

where the last inequality follows from assumption 6. Applying the same idea to the second part of A.7 we get

$$\begin{aligned}\left(\lambda_t\left(\sum_{h=1}^{T-1}\lambda_h'\lambda_h\right)^{-1}\lambda_s'\right)^2 &\leq \left(\lambda_t\left(\sum_{h=1}^{T-1}\lambda_h'\lambda_h\right)^{-1}\lambda_t'\right)\left(\lambda_s\left(\sum_{h=1}^{T-1}\lambda_h'\lambda_h\right)^{-1}\lambda_s'\right) \\ &\leq \left(\frac{F\underline{\lambda}^2}{(T-1)\underline{\varsigma}}\right)^2\end{aligned} \tag{A.8}$$

Following Abadie et al. (2010) we define

$$\overline{\epsilon_j^L} = \sum_{s=1}^{T-1}\lambda_T(\sum_{h=1}^{T-1}\lambda_h'\lambda_h)^{-1}\lambda_s'\epsilon_{js} \tag{A.9}$$

for $j = n+1, ..., J$. Assume that

Assumption 7. *The p^{th} moment of $|\epsilon_{jt}|$ for some even p exists 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:

$$
\begin{aligned}
\sum_{j=n+1}^{J} w_j^* |\overline{\epsilon_j^L}| &= \sum_{j=n+1}^{J} w_j^* |\overline{\epsilon_j^L} * 1| \leq \left(\sum_{j=n+1}^{J} w_j^* |\overline{\epsilon_j^L}|^p \right)^{1/p} \left(\sum_{j=n+1}^{J} w_j^* |1|^q \right)^{1/q} \\
&= \left(\sum_{j=n+1}^{J} w_j^* |\overline{\epsilon_j^L}|^p \right)^{1/p} \left(\sum_{j=n+1}^{J} w_j^* \right)^{1/q} = \left(\sum_{j=n+1}^{J} w_j^* |\overline{\epsilon_j^L}|^p \right)^{1/p} \leq \left(\sum_{j=n+1}^{J} |\overline{\epsilon_j^L}|^p \right)^{(1/p)}
\end{aligned}
$$

where the last equality follow from $w_{n+1}^* + ... + w_J^* = 1$ and the last inequality follows from the condition that $w_{n+1}^* \leq 1, ..., w_J^* \leq 1$. Applying Hölder's Inequality again we get

$$
E\left[\sum_{j=n+1}^{J} w_j^* |\overline{\epsilon_j^L}| \right] \leq \left(E\left[\sum_{j=n+1}^{J} |\overline{\epsilon_j^L}|^p \right] \right)^{1/p} \tag{A.10}
$$

Applying Rosenthal's Inequality we have

$$
\begin{aligned}
E\left[|\overline{\epsilon_j^L}|^p \right] &= E\left[\left| \sum_{s=1}^{T-1} \lambda_t \left(\sum_{h=1}^{T-1} \lambda_h' \lambda_h \right)^{-1} \lambda_s' \epsilon_{js} \right| \right] \\
&\leq C(p) \max \left(\sum_{s=1}^{T-1} E\left[\left| \lambda_t \left(\sum_{h=1}^{T-1} \lambda_h' \lambda_h \right)^{-1} \lambda_s' \epsilon_{js} \right|^p \right] \right. \\
&\quad , \left. \left(\sum_{s=1}^{T-1} E\left[\left| \lambda_t \left(\sum_{h=1}^{T-1} \lambda_h' \lambda_h \right)^{-1} \lambda_s' \epsilon_{js} \right|^2 \right] \right)^{p/2} \right)
\end{aligned}
$$

where $C(p)$ is the pth moment of -1 plus a Poisson random variable with mean 1 (see Abadie et al. (2010)). Consider the two elements of $max(.)$. For the first element, we have

$$
\begin{aligned}
\sum_{s=1}^{T-1} E\left[\left| \lambda_t \left(\sum_{h=1}^{T-1} \lambda_h' \lambda_h \right)^{-1} \lambda_s' \epsilon_{js} \right|^p \right] &= \sum_{s=1}^{T-1} E\left[\left(\lambda_t \left(\sum_{h=1}^{T-1} \lambda_h' \lambda_h \right)^{-1} \lambda_s' \right)^{2*(p/2)} |\epsilon_{js}|^p \right] \\
&\leq \sum_{s=1}^{T-1} E\left[\left(\frac{F\underline{\lambda}^2}{(T-1)\,\underline{\varsigma}} \right)^{2*(p/2)} |\epsilon_{js}|^p \right] \\
&= \left(\frac{F\underline{\lambda}^2}{\underline{\varsigma}} \right)^p \frac{1}{(T-1)^p} \sum_{s=1}^{T-1} E\left(|\epsilon_{js}|^p \right)
\end{aligned}
$$

where the first equality follows from the distributivity of the power and the inequality follows from A.8. For the second element in $max\,(.)$, we have

$$\left(\sum_{s=1}^{T-1} E\left[\left|\lambda_t \left(\sum_{h=1}^{T-1}\lambda_h'\lambda_h\right)^{-1}\lambda_s'\epsilon_{js}\right|^2\right]\right)^{p/2} \leq \left[\sum_{s=1}^{T-1} E\left(\left(\frac{F\overline{\lambda}^2}{(T-1)\,\underline{\varsigma}}\right)^2 \epsilon_{js}^2\right)\right]^{p/2}$$

$$= \left(\frac{F\overline{\lambda}^2}{\underline{\varsigma}}\right)^p \left[\sum_{s=1}^{T-1}\frac{1}{(T-1)^2}E\left(\epsilon_{js}^2\right)\right]^{p/2}$$

where the first inequality follows from A.8. Putting all these results together have

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

As Abadie et al. (2010), we define $\sigma_{js}^2 = E|\epsilon_{js}|^2$, $\sigma_j^2 = (1/(T-1)\sum_{s=1}^{T-1}\sigma_{js}^2)$, $\overline{\sigma^2} = max_{j=n+1,\ldots,J}\sigma_j^2$ and $\overline{\sigma} = \sqrt{\overline{\sigma^2}}$. Similarly, we define $\tau_{p,jt} = E|\epsilon_{jt}|^p$, $\tau_{p,j} = \frac{1}{(T-1)}\sum_{t=1}^{T-1}\tau_{p,jt}$, and $\overline{\tau_p} = max_{j=n+1,\ldots,J}\tau_{p,j}$. We can write the first element of $max(.)$ as

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

Similarly, the second element can be written as

$$\left[\sum_{s=1}^{T-1}\frac{1}{(T-1)^2}E(\epsilon_{js}^2)\right]^{p/2} = \left(\frac{1}{T-1}\frac{1}{T-1}\sum_{s=1}^{T-1}\sigma_{js}^2\right)^{p/2} = \left(\frac{1}{T-1}\sigma_j^2\right)^{p/2}$$

Thus , defining $\varpi = C(p)(\frac{F\overline{\lambda}^2}{\underline{\varsigma}})^p$, we have

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

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

$$
\begin{aligned}
\left(E\left[\sum_{j=n+1}^{J}|\overline{\epsilon_j^L}|^p\right]\right)^{1/p} &\leq \varpi^{1/p} max\left(\frac{(J-n-1)^{1/p}\,\overline{\tau}_p^{1/p}}{(T-1)^{1-1/p}}, \frac{(J-n-1)\,\overline{\sigma}^{2*(p/2)}}{(T-1)^{1/2}}\right)^{1/p} \\
&= \varpi^{1/p}\,(J-n-1)^{1/p}\, max\left(\frac{\overline{\tau}_{1/p}}{(T-1)^{1-\frac{1}{p}}}, \frac{\sqrt{\overline{\sigma}^2}}{(T-1)^{1/2}}\right) \quad \text{(A.11)}
\end{aligned}
$$

this implies

$$
\begin{aligned}
E\left[|R_{1t'}|\right] &= E\left[\left|\sum_{j=n+1}^{J} w_j^*\epsilon_j^L\right|\right] \\
&\leq E\left[\sum_{j=n+1}^{J} w_j^*|\epsilon_j^L|\right] \\
&\leq \left(E\left[\sum_{j=n+1}^{J}|\epsilon_j^L|^p\right]\right)^{1/p} \\
&\leq \varpi^{1/p}(J-n-1)^{1/p} max\left(\frac{\overline{\tau}_p^{1/p}}{(T-1)^{1-\frac{1}{p}}}, \frac{\overline{\sigma}}{(T-1)^{1/2}}\right)
\end{aligned}
$$

where, in the second equation, the first equality follows from A.4 and A.9, the first inequality follows from the triangular inequality, the second follows from A.10 and the third from A.11. It follows that

$$E|R_{1t'}| \leq C(p)^{1/p} \frac{\overline{\lambda^2} F}{\underline{\varsigma}} (J-n-1)^{1/p} \max \left\{ \frac{\overline{\tau_p^{1/p}}}{(T-1)^{1-1/p}}, \frac{\overline{\sigma}}{(T-1)^{1/2}} \right\}.$$

Thus, the difference between the expected value of $Y_{1t}^{0,1}$ and its synthetic counterpart can be bounded by something that goes to zero when the number of pre-intervention periods goes to infinity, namely

$$E\left(Y_{1t'}^{0,1} - \sum_{j=n+1}^{J} w_j^* Y_{jt'}\right) = E(R_{1t'}) = o(T).$$

B Identification of $\delta_{it'}(1)$

Finding a "synthetic" value of $Y_{1t}^{1,0}$ is more challenging and requires more than 1 treated unit. First, we need to 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 remaining 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. If the number of treated is big enough, we can also create a "synthetic" $Y_{it'}^{1,0}$. 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 $K_{t'}^* = (k_{n+1,t'}^*, \ldots, k_{Jt'}^*)$ chosen with a standard SCM. Note that also those weights need to be calculated for each 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 $Y_{it'}^{1,0} = \sum_{i=2}^{n} q_{it'}^* Y_{it'}$. $Q_{t'}^*$ is estimated with a SCM but using only the other treated units. More specifically, let $\Omega_1^{\delta_{t'}(1)} = (X_1, Y_{11}, \ldots, Y_{1,T-1}, M_{11}, \ldots, M_{1,T-1}, \hat{M}_{1t'}(0))$, $\omega_{0i}^{\theta_{t'}(1)} =$

$(X_i, Y_{11}, \ldots, Y_{i,T-1}, M_{11}, \ldots, M_{i,T-1}, M_{i,t'})$, and $\Omega_0^{\theta_{t'}(1)} = (\omega_2^{\theta_{t'}(1)}, \ldots, \omega_n^{\theta_{t'}(1)})'$, then

$$\begin{aligned} Q_{t'}^* &= \min_{q_{n+1,t'},\ldots,q_{Jt'}} ||\Omega_1^{\theta_{t'}(1)} - Q_{t'}\Omega_0^{\theta_{t'}(1)}||_V \\ & s.t. \quad q_{n+1,t'} \leq 0, ..., q_{Jt'} \leq 0, \sum_{i=n+1}^{J} q_{it'} = 1, \end{aligned}$$

where the distance and V are defined as above for $Y_{it'}^{0,1}$.

Let $\hat{Y}_{1t'}^{1,0} = \sum_{i=2}^{n} q_{it'}^* Y_{it'}$, similar as before, we assume that $Q_{t'}^*$ exists and satisfies $\forall\ t = 1, ..., T-1$

$$\begin{aligned} \sum_{j=2}^{n} q_{jt'}^* Y_{jt} &= Y_{1t}, \\ \sum_{j=2}^{n} q_{jt'}^* X_j &= X_1, \end{aligned}$$

$$\sum_{j=2}^{n} q_{jt'}^* M_{jt} = M_{1t},$$

$\forall\ t = 1, ..., T-1$ and

$$\sum_{j=2}^{n} q_{jt'}^* M_{jt'} = \hat{M}_{1t'}(0).$$

Under extra standard conditions and assuming that $\rho_{t'}(\cdot)$ is a linear function, as we show in the appendix

$$E(\hat{Y}_{1t'}^{1,0}) = Y_{1t'}^{1,0} + o(T).$$

The latter assumption can admittedly be restrictive in many applications. However, it is substantially weaker than assuming a constant $\rho_{t'}$. Then, we can estimate the indirect effect $\delta_{it'}(1)$ and the direct effect as $\theta_{1t'}(M_{1t}(0))$ as

$$\hat{\delta}_{1t'}(1) = Y_{1t'} - \hat{Y}_{1t'}^{1,0}, \qquad \hat{\theta}_{1t'}(M_{1t'}(0)) = \hat{\alpha}_{it'} - \hat{\delta}_{1t'}(1),$$

respectively. Intuitively, $Q_{t'}^*$ exists under the similar assumptions as the one discussed in the main text. However, if the number of treated is too small $\hat{Y}_{1t'}^{1,0}$ will be a very poor approximation of $Y_{1t'}^{1,0}$. In this settings it is only possible to estimate $\delta_{it'}(0)$ and $\theta_{it'}(1)$.

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

To create a synthetic Y_{1t}^{10} we need to impose the standard SCM assumptions on the mediator which are:

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 $i \in \{1, n+1, ..., J\}$ and for $t = 1, ..., t'$*

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

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

Assumption 13.

$$\exists \quad \underline{\vartheta} \text{ s.t. } |\vartheta_{tv}| \leq \underline{\vartheta} \quad \forall \text{ t=1,...,t' and v=1,...,} V\,. \tag{C.2}$$

Assumption 14. *$\exists$ a p^{th} moment of $|\nu_{jt}|$ for some even p and for $j = n+1, ..., J$ and $t = 1, ..., t'$*

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

As for Y_{1t}^{01} we drop the subscript t from the weight and we write

$$\begin{aligned}\sum_{j=2}^{n} q_j Y_{jt} = \zeta_t \quad &+ \quad \eta_t \sum_{j=2}^{n} q_j X_j + \lambda_t \sum_{j=2}^{n} q_j \mu_j + \varphi_t \left(I\{t \geq T\}\right) \sum_{j=2}^{n} q_j M_{jt} \left(I\{t \geq T\}\right) \\ &+ \quad \sum_{j=2}^{n} q_j \rho_t \left(M_{jt} \left(I\{t \geq T\}\right)\right) I\{t \geq T\} + \sum_{j=2}^{n} q_j \epsilon_{jt}.\end{aligned}$$

Thus,

$$
\begin{aligned}
Y_{1t}^{1,0} - \sum_{j=2}^{n} q_j Y_{jt} &= \eta_t \left(X_1 - \sum_{j=2}^{n} q_j X_j \right) + \lambda_t \left(\mu_1 - \sum_{j=2}^{n} q_j \mu_j \right) \\
&+ \varphi_t \left(I\{t \geq T\} \right) \left(M_{1t}(0) - \sum_{j=2}^{n} q_j M_{jt}(I\{t \geq T\}) \right) \\
&+ \left(\rho_t \left(M_{1t}\left(0\right)\right) - \sum_{j=2}^{n} q_j \rho_t \left(M_{jt}\left(I\{t \geq T\}\right)\right) \right) I\{t \geq T\} \\
&+ \sum_{j=2}^{n} q_j \left(\epsilon_{1t} - \epsilon_{jt} \right)
\end{aligned}
$$

Using the same notation as before in the pre-intervention period we have

$$
\begin{aligned}
Y_1^P - \sum_{j=2}^{n} q_j Y_j^P &= \eta^P \left(X_1 - \sum_{j=2}^{n} q_j X_j \right) + \lambda^P \left(\mu_1 - \sum_{j=2}^{n} q_j \mu_j \right) \\
&+ \varphi^P \left(0\right) \left(M_1^P\left(0\right) - \sum_{j=2}^{n} q_j M_j^P\left(0\right) \right) + \left(\epsilon_1^P - \sum_{j=2}^{n} q_j \epsilon_j^P \right)
\end{aligned}
$$

Thus

$$
\begin{aligned}
\lambda^P \left(\mu_1 - \sum_{j=2}^{n} q_j \mu_j \right) &= Y_1^P - \sum_{j=2}^{n} q_j Y_j^P - \eta^P \left(X_1 - \sum_{j=2}^{n} q_j X_j \right) \\
&- \varphi^P \left(0\right) \left(M_1^P\left(0\right) - \sum_{j=2}^{n} q_j M_j^P\left(0\right) \right) - \left(\epsilon_1^P - \sum_{j=2}^{n} q_j \epsilon_j^P \right)
\end{aligned}
$$

Multiplying both sides by $(\lambda^{P\prime}\lambda^P)^{-1}\lambda^{P\prime}$ we get

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

Substituting in D.1 and considering a generic post-intervention period t', we have

$$\begin{aligned}
Y_{1t'}^{1,0} - \sum_{j=2}^{n} q_j Y_{jt'} &= \left(\lambda^{P'}\lambda^{P}\right)^{-1}\lambda^{P'}\left(Y_1^P - \sum_{j=2}^{n} q_j Y_j^P\right) \\
&+ \left(\eta_{t'} - \left(\lambda^{P'}\lambda^{P}\right)^{-1}\lambda^{P'}\eta^{P}\right)\left(X_1 - \sum_{j=2}^{n} q_j X_j\right) \\
&- \left(\lambda^{P'}\lambda^{P}\right)^{-1}\lambda^{P'}\varphi^{P}(0)\left(M_1^P(0) - \sum_{j=2}^{n} q_j M_j^P(0)\right) \\
&+ \varphi_{t'}(1)\left(M_{1t'}(0) - \sum_{j=2}^{n} q_j M_{jt'}(1)\right) \\
&+ \left(\rho_{t'}\left(M_{1t'}(0)\right) - \sum_{j=2}^{n} q_j \rho_{t'}\left(M_{jt}(1)\right)\right) \\
&- \left(\lambda^{P'}\lambda^{P}\right)^{-1}\lambda^{P'}\left(\epsilon_1^P - \sum_{j=2}^{n} q_j \epsilon_j^P\right) + \sum_{j=2}^{n} q_j\left(\epsilon_{1t'} - \epsilon_{jt'}\right)
\end{aligned}$$

Assume, as we did in the main text, that there exists weights $q_2^*, ..., q_n^*$ that satisfy $\forall t = 1, ..., T-1$

$$\begin{aligned}
\sum_{j=2}^{n} q_j^* Y_{jt} &= Y_{1t}, \\
\sum_{j=2}^{n} q_j^* X_j &= X_1, \\
\sum_{j=2}^{n} q_j^* M_{jt} &= M_{1t},
\end{aligned}$$

and it also satisfies

$$\sum_{j=2}^{n} q_j^* M_{jt'} = \hat{M}_{1t'}(0).$$

Substituting the generic weights with $q_2^*, ..., q_n^*$ in the post-intervention period t', we get

$$\begin{aligned}
Y_{1t'}^{1,0} - \sum_{j=2}^{n} q_j^* Y_{jt'} &= \left(\rho_{t'}\left(M_{1t'}(0)\right) - \sum_{j=2}^{n} q_j^* \rho_{t'}\left(M_{jt'}(1)\right)\right) \\
&- \left(\lambda^{P'}\lambda^{P}\right)^{-1}\lambda^{P'}\left(\epsilon_1^P - \sum_{j=2}^{n} q_j^* \epsilon_j^P\right) + \sum_{j=2}^{n} q_j^*\left(\epsilon_{1t'} - \epsilon_{jt'}\right)
\end{aligned}$$

Note that, as by assumption $\sum_{j=2}^{n} q_j^* M_{t'} = \hat{M}_{1t'}(0)$ and $\hat{M}_{1t'}(0)$ is estimated using a standard SCM

$$E\left(\varphi_{t'}(1)\left(M_{1t'}(0) - \sum_{j=2}^{n} q_j^* M_{jt'}(1)\right)\right) = o(T).$$

As we mention in the main text, for identification we have to impose an extra assumption, namely

Assumption 15. *$\rho_{t'}(.)$ is a linear function*

Under assumption 15 we have

$$\begin{aligned} E\left[\left(\rho_{t'}\left(M_{1t'}(0)\right) - \sum_{j=2}^{n} q_j^* \rho_{t'}\left(M_{jt'}(1)\right)\right)\right] &= E\left[\left(\rho_{t'}\left(M_{1t'}(0)\right) - \rho_{t'}\left(\sum_{j=2}^{n} q_j^* M_{jt'}(1)\right)\right)\right], \\ &= E\left[\left(\rho_{t'}\left(M_{1t'}(0)\right) - \rho_{t'}\left(\hat{M}_{1t'}(0)\right)\right)\right], \\ &= \rho_{t'}\left(M_{1t'}(0)\right) - \rho_{t'}\left(E(\hat{M}_{1t'}(0))\right) = o(T). \end{aligned}$$

Thus,

$$Y_{1t'}^{1,0} - \sum_{j=2}^{n} q_j^* Y_{jt'} = -\left(\lambda^{P'}\lambda^{P}\right)^{-1} \lambda^{P'}\left(\epsilon_1^P - \sum_{j=2}^{n} q_j^* \epsilon_j^P\right) + \sum_{j=2}^{n} q_j^*\left(\epsilon_{1t'} - \epsilon_{jt'}\right).$$

This, with an analogous as the one above therefore omitted proof, can be shown to imply

$$E(Y_{1t'}^{1,0} - \sum_{j=2}^{n} q_j^* Y_{jt'}) = o(T).$$

E Constraints and Donor Pool

For each treated country, we use different covariates and donor pools, trying to balance between three different goals: obtaining satisfying pre-(post-)intervention fits for the outcome and (or) the mediator, obtaining robust results and obtaining a synthetic unit with covariates similar enough to those of the treated unit. In addition, data availability is an issue in some countries.

The variable we used came from different sources. We use data on capital stock in constant international dollars and on investments as a percentage of GDP, coming from the International Monetary Fund database. Data on the percentage of population older than 25 with secondary education and those with tertiary education, on the percentage of internet users, on the amount of trade (the sum of exports and imports of good and services) as a share of GDP, on natural resources rents (calculated as the difference between commodity price and its average cost of production) as a share of GDP, on the number of patents applications per millions of residents (as a measure of technological development) come from the World Bank. Data on total factor productivity at constant national prices come from Penn World Table version 9.0. Data on employment share by country come from the ILO database.

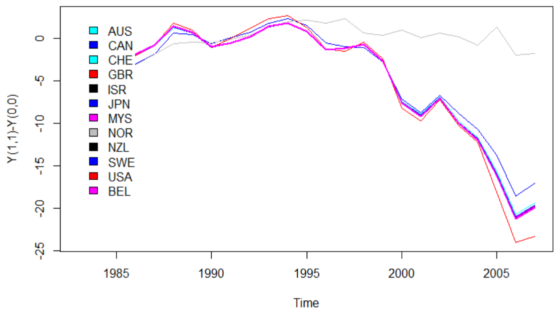
The donor pools are selected from a restricted group of countries. Indeed, following Puzzello and Gomis-Porqueras (2018), we exclude all countries which during pre- and/or post-treatment periods were affected by one or more conflicts with more than 1000 deaths for at least two years, and/or those which experienced defaults or rescheduling of domestic or foreign debt for at least three consecutive years and/or were autocracies. To determine which countries were affected by conflicts we use data from the Uppsala Conflict Data Program (https://ucdp.uu.se/). Information on defaults and rescheduling of debts are taken from Reinhart and Rogoff (2009). Autocracies are defined as all countries with a value smaller than zero of the polity index developped in the Polity IV Project of the Center for Systemic Peace (http://www.systemicpeace.org/polityproject.html). We further exclude Trinidad and Tobago, and Uruguay for data availability and Denmark because they peg their currency to the euro.

The variables we used as controls as well as the donor pool used for each treated country are presented in table 3.

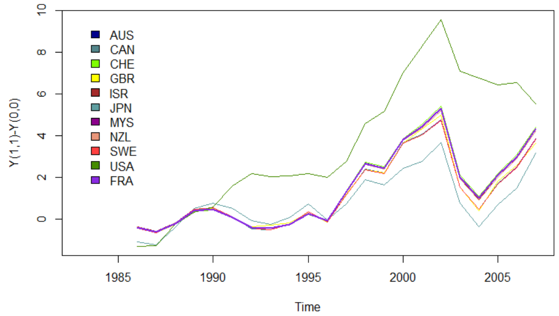
Table 3: Donor Pool and Control Variables

country	Constrained Variables	Donor Pool
France	Average outcome and average mediator over time frames: 1986-1989, 1990-1993, 1994-1997	Australia, Sweden, Canada, Switzerland, United Kingdom, Israel, Malasya, New Zealand, Japan, USA
Belgium	Average outcome and average mediator over time frames: 1986-1989, 1990-1993, 1994-1997	Australia, Sweden, Canada, Switzerland, United Kingdom, Norway, Israel, Malasya, New Zealand, Japan, USA
Ireland	All values of outcome and all values of mediator over time frames: 1986-1997	Australia, Sweden, Canada, United Kingdom, Norway, USA Japan, Switzerland, Israel, Malasya, New Zealand
Italy	Average outcome and average mediator over time frames: 1986-1989, 1990-1993, 1994-1997 Average capital stock over time frames: 1986-1990, 1991-1997 Average schooling over years: 1990, 1995 Average Internet users over time frame: 1992-1997	Australia, Sweden, Canada, Japan, Israel, Malasya, New Zealand, USA, Norway, Switzerland, United Kingdom
Netherlands	Average outcome and average mediator over time frames: 1986-1989, 1990-1993, 1994-1997 Average capital stock over time frames: 1986-1990, 1991-1997 Average schooling over years: 1990, 1995 Average rents of natural resources over time frames: 1986-1990, 1991-1997 Average trade over time frames: 1986-1990, 1991-1997	Australia, Sweden, Canada, Japan, Israel, Malasya, New Zealand, USA, Switzerland, United Kingdom

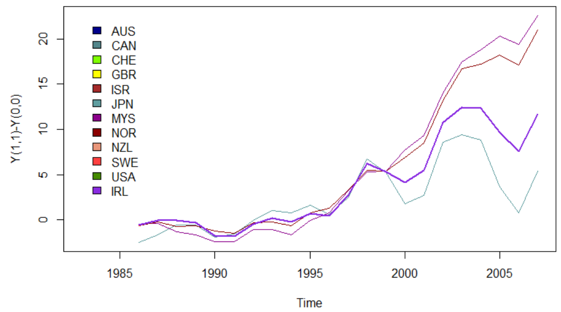
F Leave-one-out Robustness Checks



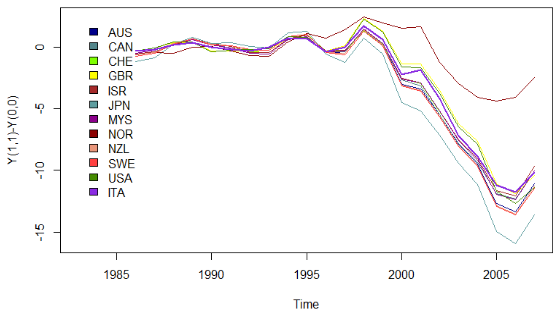
(a) Belgium



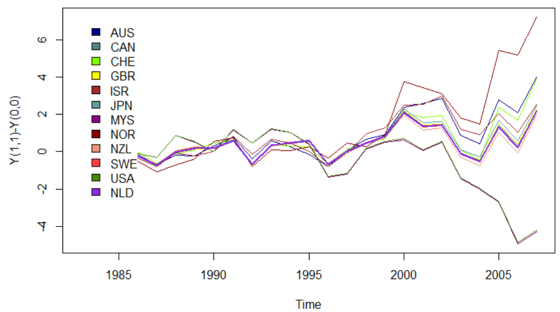
(b) France



(c) Ireland

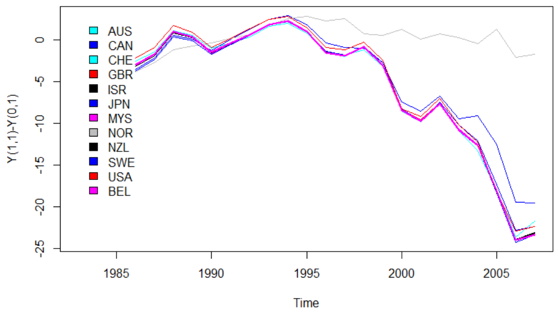


(d) Italy

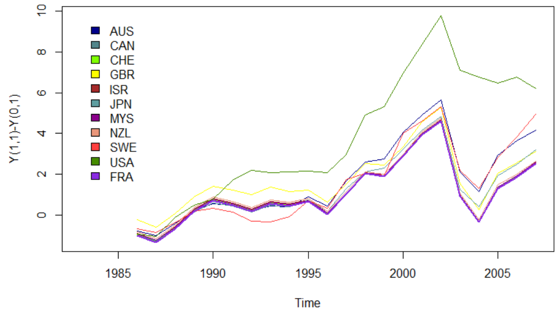


(e) Netherlands

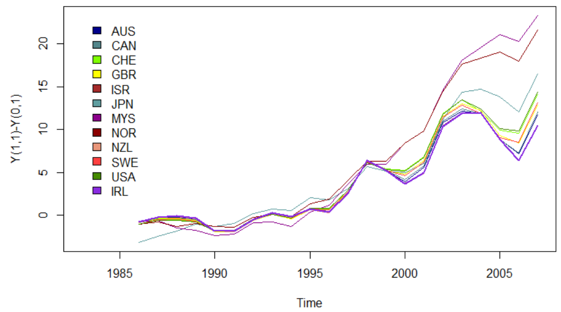
Figure 4: Leave-one-out robustness checks for the total effect.



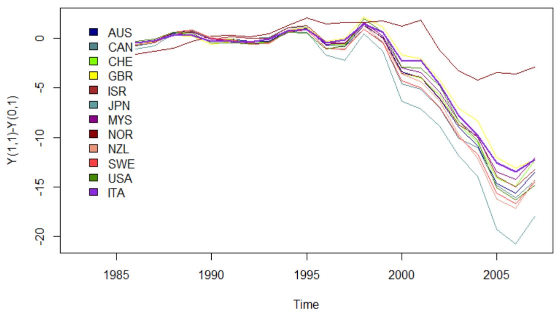
(a) Belgium



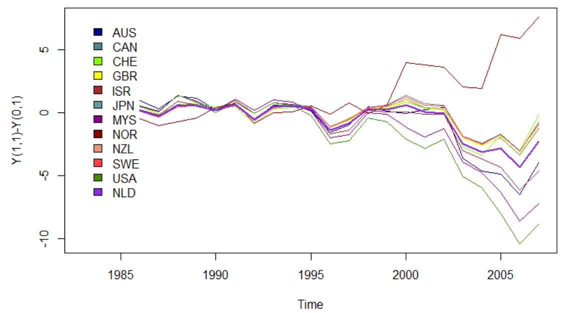
(b) France



(c) Ireland

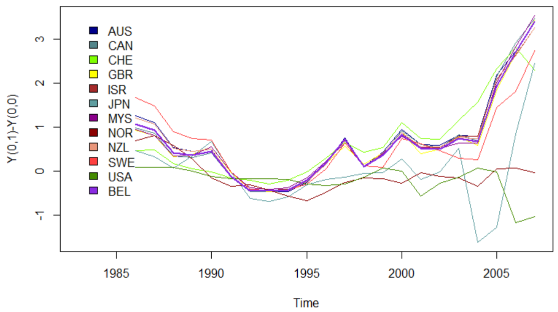


(d) Italy

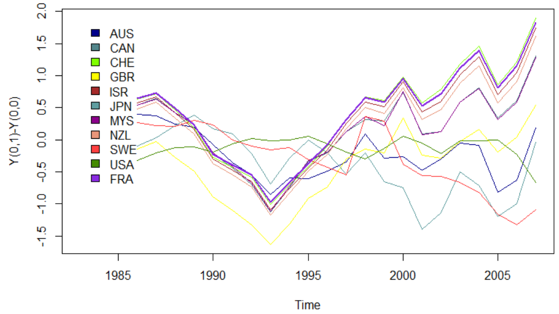


(e) Netherlands

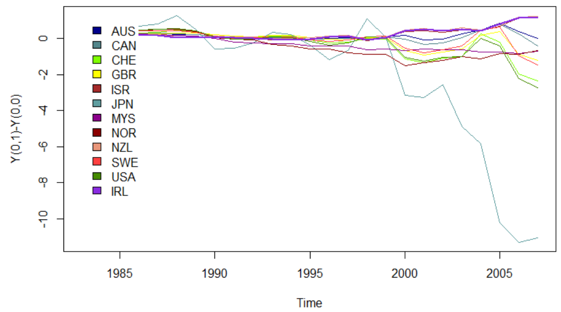
Figure 5: Leave-one-out robustness checks for the direct effect.



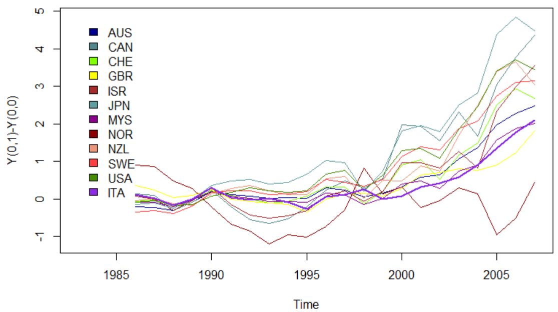
(a) Belgium



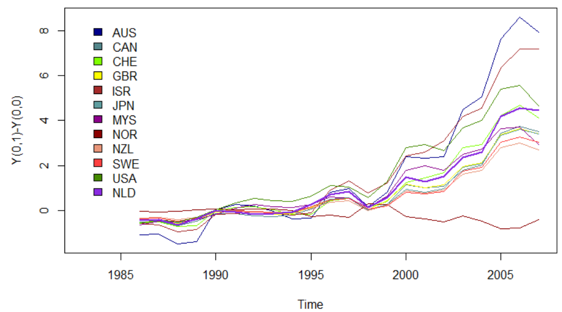
(b) France



(c) Ireland



(d) Italy



(e) Netherlands

Figure 6: Leave-one-out robustness checks for the indirect effect.

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*.

Baldwin, R., V. Di Nino, L. Fontagné, R. De Santis, and D. Taglioni (2008): "Study on the Impact of the Euro on Trade and Foreign Direct Investment," *Economic Papers*, 321.

Baldwin, R. and D. Taglioni (2007): "Trade Effects of the Euro: a Comparison of Estimators," *Journal of Economic Integration*, 22, 780–818.

Barro, R. J. and S. Tenreyro (2007): "Economic effects of currency unions," *Economic Inquiry*, 45, 1–23.

Bayar, G. (2002): "Effects of Foreign Trade Liberalization on the Productivity of Industrial Sectors in Turkey," *Emerging Markets Finance & Trade*, 38, 46–71.

Ben-Michael, E., A. Feller, and J. Rothstein (2018): "The Augmented Synthetic Control Method," *arXiv:1811.04170*.

Berger, H. and V. Nitsch (2008): "Zooming out: The trade effect of the euro in historical perspective," *Journal of International Money and Finance*, 27, 1244–1260.

Bun, M. J. G. and F. J. G. M. Klaassen (2007): "The Euro Effect on Trade is not as Large as Commonly Thought," *Oxford Bulletin of Economics and Statistics*, 69, 473–496.

Chernozhukov, V., K. Wüthrich, and Y. Zhu (2018): "An Exact and Robust Conformal Inference Method for Counterfactual and Synthetic Controls," *arXiv:1712.09089v6.*

Chintrakarn, P. (2008): "Estimating the Euro Effects on Trade with Propensity Score Matching," *Review of International Economics*, 16, 186–198.

De Nardis, S. and C. Vicarelli (2003): "Currency unions and trade: The special case of EMU," *Review of World Economics*, 139, 625–649.

Žúdel, B. and L. Melioris (2016): "Five Years in Balloon: Estimating the Effects of Euro Adoption in Slovakia Using the Synthetic Control Method," *OECD Economics Department Working Papers.*

Deuchert, E., M. Huber, and M. Schelker (2018): "Direct and indirect effects based on difference-in-differences with an application to political preferences following the Vietnam draft lottery," *Journal of Business & Economic Statistics*, forthcoming.

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

Ferman, B. and C. Pinto (2017): "Placebo Tests for Synthetic Controls," *MPRA Paper.*

Firpo, S. and V. Possebom (2018): "Synthetic control method: Inference, sensitivity analysis and confidence sets," *Journal of Causal Inference*, 6, 1–26.

Flam, H. and H. Nordström (2006): "Trade volume effects of the euro: Aggregate and sector estimates," *IIES Seminar Paper.*

Gabrielczak, P. and T. Serwach (2017): "The impact of the euro adoption on the complexity of goods in Slovenian exports," *Zb. Rad. Ekon. Fak. Rij.*, 35, 45–71.

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

Hausmann, R., C. A. Hidalgo, S. Bustos, M. Coscia, S. Chung, J. Jimenez, A. Simoes, and M. A. Yildirim (2014): "The atlas of economic com- plexity: mapping paths to prosperity," *Cambridge MIT Press.*

Hidalgo, C. A. and R. Hausmann (2009): "The building blocks of economic complexity," *Proceedings of the National Academy of Sciences*, 106, 10570–10575.

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.

Kreif, N., R. Grieve, D. Hangartner, A. J. Turner, S. Nikolova, and M. Sutton (2016): "Examination of the synthetic control method for evaluating health policies with multiple treated units," *Health Economics*, 25, 1514–1528.

Mancini-Griffoli, T. and L. L. Pauwels (2006): "Is there a euro effect on trade? An application of the end-to-sample structural break tests for panel data," *HEI Working Paper*.

Micco, A., E. Stein, and G. Ordoñez (2003): "The Currency Union Effect on Trade: Early Evidence from EMU," *Economic Policy*, 18, 315–356.

Midelfart-Knarvik, K. H., H. G. Overman, S. J. Redding, and A. J. Venables (2000): "The Location of European industry," *European Commission, Economic Paper*.

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

Pesaran, M. H., L. V. Smith, and R. P. Smith (2007): "What if the UK or Sweden had joined the euro in 1999? An empirical evaluation using a global VAR," *International Journal of Finance and Economics*, 12, 55–87.

Petrakos, G., D. Kallioras, and A. Ageliki (2006): "Determinants of Industrial Performance in the EU-15 Countries, 1980-2003," *46th Congress of the European Regional Science Association: "Enlargement, Southern Europe and the Mediterranean", August 30th - September 3rd, Volos, Greece.*

Puzzello, L. and P. Gomis-Porqueras (2018): "Winners and losers from the €uro," *European Economic Review*, 108, 129–152.

Reinhart, C. M. and K. S. Rogoff (2009): "This Time is different: Eight centuries of financial folly," *Princeton University Press.*

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.

Rose, A. K. (2000): "One money, one market: The effect of currency unions on trade," *Economic Policy*, 15, 8–45.

Rubin, D. B. (1974): "Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies," *Journal of Educational Psychology*, 66, 688–701.

Saia, A. (2017): "Choosing the open sea: The cost to the UK of staying out of the euro," *Journal of International Economics*, 108, 82–98.

Silva, J. M. C. S. and S. Tenreyro (2010): "Currency unions in prospect and retrospect," *Annual Review of Economics*, 2, 51–74.

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.