PERUVIAN ECONOMIC ASSOCIATION

The economic effects of international administrations: The cases of Kosovo and East Timor

César Urquizo

Diego Winkelried

Working Paper No. 70, June 2016

The economic effects of international administrations: The cases of Kosovo and East Timor*

César Urquizo
Universidad del Pacífico (Lima, Perú)
c.urquizoubillus@up.edu.pe

Diego Winkelried†
Universidad del Pacífico (Lima, Perú)
winkelried_dm@up.edu.pe

This version: June 20, 2016

Abstract

Does the involvement of foreign third parties in the post-conflict management of a country in the wake of a civil war have positive or negative economic effects? The approaches used to address this question in the social and political sciences literature are mostly qualitative, and are not sufficiently supported by quantitative evidence. This document attempts a quantitative analysis of the post-conflict economic performance of Kosovo and East Timor under the UN-sponsored international administrations established in both countries in the late 1990s. By using the synthetic control impact evaluation technique, we compute suitable counterfactual scenarios for each country to estimate the intervention effects of interest. A robust negative effect of the intervention is found for Kosovo, whereas the effect on East Timor is positive.

JEL Classification : C54, D74, P16.
Keywords : impact evaluation, synthetic control, Kosovo, UNMIK, East Timor, UNTAET.

*We would like to thank Alberto Abadie, Farid Kahhat, Ramiro Gil-Serrate, Alonso Gurmendi, Roberto Heimowitz and Pablo Lavado for insightful comments. We are also grateful to participants at the 2015 Annual Congress of the Peruvian Economic Association, and at Research Seminars of the Central Reserve Bank of Peru for their suggestions. Financial support of the Research Center at the Universidad del Pacifico (CIUP) is gratefully acknowledged. The opinions expressed in this paper are our own.

†Corresponding author. Department of Finance, Universidad del Pacífico, Av. Salaverry 2020, Lima 11, Peru. Phone: (+511) 219 0100 ext. 2238.

1 Motivation

Comparative research, prominent in social and political sciences, provides evidence supporting hypotheses of interest by means of a meticulous analysis of the similarities and differences among a small number of selected cases. A particular question that researchers have endeavored to answer from a comparative perspective is whether the resolution of conflicts through foreign interventions are economically beneficial for countries involved in violent internal conflicts (see, *inter alia*, Collier et al., 2003). In this study we focus on two of the most emblematic United Nations (UN) international administrations of the late 20th century (see Wilde, 2001; Wolfrum, 2005): the UN Interim Administration Mission in Kosovo (UNMIK) and the UN Transitional Administration in East Timor (UNTAET). Under an international administration, the UN, beyond peacekeeping, may act as an interim government in post-conflict territories. Both the UNMIK and the UNTAET launched in the second half of 1999, were fully operational by 2000 (which we use as "intervention date"), and have since gradually transferred responsibilities to autonomous governments. We inquire if these interventions had a positive or negative effect on post-conflict economic performance and attempt to make a quantitative assessment of such effects.

A rough answer to this question can be provided by a difference-in-difference analysis, a technique often used to identify causal effects in observational data (see Bertrand et al., 2004, for a review). Figure 1 shows, for both countries, the average growth of per capita gross national income in U.S. dollars, an available measure of economic performance, a decade before and a decade after the intervention. It also shows the growth rates of "reasonable" control groups, constructed as the average of neighbor countries (Albania and former Yugoslav republics for Kosovo, and Southeast Asian countries for East Timor).[1]

In the case of Kosovo, an average growth of 12.5 percent is recorded after the intervention; i.e., 19.5 percentage points higher than the average growth rate in the pre-intervention period (−7.0 percent). This difference captures the effect of the intervention, but also the influence of a number of economic drivers that may be unrelated to the intervention. To isolate the former effect, we note that the average growth rate in neighbor countries, which are "similar" to Kosovo but did not experience an intervention (i.e., a "proximity control") changed from −1.0 percent before the intervention to 11.0 percent afterwards. Thus, this 12.0 percentage point difference in the average growth rates can be attributed to factors other than the intervention, rendering an estimated intervention effect of $19.5 - 12.0 = 7.5$ percent.

Likewise, the analysis for East Timor produces an estimated intervention effect of 6.6 percent, computed as the difference between the 12.1 percentage point increase in the country's average growth rates (19.5 percent after the intervention and 7.4 prior to it) and the 5.5 percentage point increase in neighbor countries' average growth (11.5 percent after the intervention and 6.0 percent prior to it). Thus, this preliminary analysis indicates that the UN interventions had a positive impact, even of comparable magnitudes, on the economic performance of both intervened countries. Such a finding supports the idea, put forward by studies like Collier et al. (2008) or Kane and Jones (2012), that the deployment of foreign troops may have positive effects on the economic growth of a conflict country.

However, the above estimates are valid only under special circumstances, for instance when the intervention is randomly assigned; i.e., whether a country is intervened or not is not related to the country's characteristics or history, which seems to be a rather unrealistic assumption. A more subtle source of bias, even under the randomization of the intervention, arises if the control group is not fully comparable to the intervened unit. A result of this potential bias is that the estimated intervention effects would be different if, in the analysis of Figure 1, the control group is taken as the average of all developing countries in our sample (a striking 12.9 percent for Kosovo and about 5.5 percent for East Timor). Given that the neighbor countries form, at least qualitatively, a better basis for comparison, the "rest of the world" average can be easily regarded as not comparable to the intervened countries. Nonetheless, a more appropiate quantitative approach is called for to answer whether the neighbor countries themselves form such a desirable comparison group, despite what intuition or conventional wisdom might suggest.

[1] Details on the data definitions and sources are given in section 4.1.

There is a large, well-developed literature exploring statistical methods to overcome these difficulties in observational studies like ours; see Lee (2005) for a textbook account, and Bertrand et al. (2004) and Imbens and Wooldridge (2009) for comprehensive surveys. Nonetheless, these methods rely heavily on having a large number of control and intervened units, which does not necessarily happen in comparative case studies dealing with aggregate data. However, Abadie and Gardeazabal (2003) and Abadie et al. (2010, 2014) propose the so-called synthetic control method as a formal approach for inferring on intervention effects even in cases where a single intervened unit is available.

The purpose of this study is to assess the economic effect of the UNMIK and the UNTAET through the lens of the synthetic control methodology. This approach consists of generating a counterfactual scenario for the intervened unit (i.e., the synthetic control) as a weighted average of units in the control group, where the weights are chosen so that the counterfactual resembles as much as possible the intervened unit prior to the intervention. The method shares several similarities with traditional evaluation techniques like matching (see Lee, 2005), especially in their attempt to provide two fully comparable units that only differ in their intervention status. Nonetheless, since the identification of causal effects in the synthetic control method relies on its ability to mimic the pre-intervention history of the intervened unit, the data requirements are different from those in traditional methods. In particular, the estimation operates along relatively long time-series and small cross-sectional dimensions, making it well-suited for comparative case studies.

It is important to stress that, even though we give a similar treatment to the UNMIK and the UNTAET in our empirical work, and besides the fact that the intervention year is coincidentally the same in both cases, we acknowledge that —for a number of historical, political and geographical reasons —the experiences of these UN international administrations may not be comparable (observers, in fact, tend to be strongly critical of the UNMIK and to regard the UNTAET as successful). To put differently, we actually evaluate the effect of two "different" interventions, and our main concern is assessing the internal validity of the results for each country under study.[2] Our analysis lacks external validity, and we do not attempt to estimate an unconditional "UN international administration" effect.

The rest of the paper is organized as follows. Section 2 provides a brief review of the literature on the effects of foreign interventions on the economic performance of the intervened territory and, on the other hand, on the synthetic control method in comparative studies. This section also presents a basic background history of the conflicts and resolutions in Kosovo and East Timor. Section 3 discusses the synthetic control method and how inferences can be obtained under this approach. Section 4 presents our empirical exploration and the main findings of the study. In a nutshell, we find that, a decade after the intervention date, the UNMIK had a negative effect on the economic performance of Kosovo, whereas the UNTAET was followed by a significant positive effect on the economic performance in East Timor. Section 5 provides conclusions and avenues for further research.

2 Historical background and literature review

2.1 Historical background

Despite several similarities between the UNMIK and the UNTAET, especially in their mandate, their judicial functions, and the gradual transfer of responsibilities to self-determined local authorities, it would not be correct to assume that Kosovo and East Timor went through similar transitional processes once the international administrations began. This is so because the very internal conflicts leading to the interventions were substantially different, and so was the international support for each intervention. Thus, different intervention effects are likely to emerge.

[2] It is also important to stress that the estimated effects may be capturing the effect of multiple interventions whose marginal effects are difficult, if not impossible, to disentangle. For instance, as we review later, the UNMIK received peacekeeping support from a coalition of countries and has changed its functions since its inception, whereas the UNTAET gave rise to additional UN administrations, with different scopes and functions. Thus, by "intervention effect" we mean the cumulative effect over the transitional processes that started with the establishment of the very first UN administration.

Next, we provide a brief description of the conflicts suffered by both countries and of their resolution processes. This section does not intend to be a comprehensive account of both experiences, since the goal is limited to summarizing qualitative arguments that may help us explain such differences. The interested reader is referred to sources like Vickers (1998) and King and Mason (2006) for Kosovo, and Martin and Mayer-Rieckh (2005) and Ofstad (2012) for East Timor. Strohmeyer (2001), Wilde (2001) and Wolfrum (2005) also provide a detailed account of the UN international administrations under study.

Kosovo

Kosovo, located in the Balkans region of Europe, is a partially recognized state that used to be a province of Serbia, itself part of the former Yugoslavia. Unlike the rest of Serbia, since the late 19th century a vast majority of its population has been of Albanian ethnicity. Hostilities between Albanians and Serbs have been commonplace eversince; and in the 1980s the ethnic tensions —fueled by political instability, an economic crisis and nationalist unrest —escalated into serious violence. The 1980s closed with the Serbian government basically revoking Kosovan autonomy (granted by the 1974 Yugoslav Constitution) by establishing Serbian control over law enforcement and the judiciary.

In the early 1990s, an unofficial Albanian parliament declared Kosovo an independent country, backed by a referendum among ethnic Albanians. Such actions were dubbed as illegal and unconstitutional by the non-Albanian population of Yugoslavia, and were not recognized internationally. In practice, Albanians organized a parallel state, including the Kosovo Liberation Army (KLA), a paramilitary organization that sought an "ethnically clean" Kosovo and the eventual creation of a "Greater Albania". The KLA initially carried out sporadic attacks against Serbian law enforcement agencies, but soon they became increasingly frequent and violent. As a response, Serbian paramilitaries and the Yugoslav armed forces pursued a campaign of retribution, starting the so-called Kosovo War in 1998.

During 1998, major atrocities were committed, causing a large number of refugees to flee the region. After these events, a coalition of international powers led by the U.S. asked for a cease-fire, the withdrawal of the armed forces, and unrestricted access to international monitors. All these demands were accepted by the Yugoslav government, but never implemented. The KLA resumed hostilities, causing a very strong military response from the Yugoslav government. The United Nations Security Council (UNSC) condemned these actions and authorized the North Atlantic Treaty Organization (NATO) to bombard Serbia in March 1999. After 11 weeks of air strikes, the Yugoslav government signed a peace treaty. In June 1999, the UNMIK was established.

The UNMIK was set up to provide the region with political stability after its destructive conflict, backed by the NATO-led peacekeeping Kosovo Force (KFOR). Even though the UNMIK did manage to eventually (but not smoothly) calm a volatile situation, paving the way for Kosovo's proclamation of independence in 2008, experts agree that it failed to establish a stronger state and facilitate a definitive settlement with Serbia. It is argued that one important reason for this was that the UNMIK marginalized the local population from the decision-making process, thereby eroding its credibility and the population's support for international involvement. In 2008, the UNMIK transferred most of its functions to the newly-formed government and to the European Union Rule of Law Mission in Kosovo (EULEX), which marked, for all practical purposes, the end of the UN international administration.

East Timor

East Timor is an insular country located in Southeast Asia. It was a Portuguese colony from the 16th century to 1975, when the Revolutionary Front for an Independent East Timor (FRETILIN) declared a short-lived independence. In the same year, the territory was occupied by Indonesia, marking the beginning of a long war between the Indonesian army and the Armed Forces for the National Liberation of East Timor (FALINTIL), the military wing of the FRETILIN. The report of the Commission for Reception, Truth and Reconciliation in East Timor (CVAR) states that both sides were guilty of numerous atrocities.

One such transgression, the so-called "Santa Cruz massacre" of the early 1990s, was broadcasted and drew international attention on the conflict in East Timor. As a result of international pressure, along with financial difficulties faced by the Indonesian government to support the occupation, in 1999 the Indonesian government agreed with the UN to hold a referendum, whereby the Timorese would be given the choice of greater autonomy within Indonesia or independence. The United Nations Mission in East Timor (UNAMET) was established in June 1999 to organize and conduct the consultation.

In September 1999, violent attacks carried out by anti-independence militants quickly spread throughout the country, as the referendum results were announced (78.5 percent voted for independence). The UNSC supported an Australian-led peacekeeping force to enter East Timor to manage the crisis, and in October 1999 it established the UNTAET in lieu of the UNAMET. The purpose of the UNTAET was to help East Timor in its transition to democracy, self-governance and economic reconstruction. Among other things, it facilitated the preparation of the CVAR report to support internal reconciliation.

The functions of the UNTAET were completely transferred to the government in May 2002, after the declaration of independence. As the UNATET was coming to an end, the UN Mission of Support to East Timor (UNMISET) was established to manage the peacekeeping efforts. As the country achieved a stable and fully functional political system, experts regard the UN intervention in East Timor as successful.

2.2 Literature review

Foreign interventions

An increasing body of studies on transitional justice focuses on whether the transition processes should be local or global. Should countries in conflict face these situations by themselves or with the help of a third party? Even though authors such as Weinstein (2005) point out that autonomous processes within societies should correct security problems, thus making international intervention unnecessary, the bulk of the literature agrees that foreign intervention, in general, is necessary.

For instance, Collier et al. (2008) argue that, since impoverished societies tend to be structurally insecure, definite solutions to conflicts would need external assistance. Also, Roberts (2006) indicates that the support of third parties is beneficial when the country facing the transition has a non-legitimate government. Furthermore, Triponel and Pearson (2010) suggest that the intervention of third parties can help achieve accountability and national reconciliation, especially when large scale atrocities and human rights violations have occurred.

It has also been acknowledged that the success of foreign interventions depends on a number of factors. Benvenisti (1993) argues that they are more likely to be successful if the occupying power intervenes as little as possible in the affairs of the local population. Gurmendi (2015) states that the principles of national ownership and citizen participation are key to guaranteeing a successful transition from conflict to peace, since relevant stakeholders are more prone to commit to the transition process when they have a more active participation in it. Hence, a multilateral approach when intervening in a country with security problems should be taken as the preferred strategy. Authors like Fox (2012) elaborate further, arguing that the more unilateral the transformation mission, the greater the incentives for occupying countries to put aside local interests in favor of their own.

Quantitative analysis include Collier et al. (2008), who estimate the benefits of stopping a conflict by providing packages including post-conflict aid, military spending limits, peacekeeping and guarantees; and find cost-to-benefit ratios ranging from 1:4 to 1:18. These authors also find that countries affected by civil violence lose, on average, over two percentage points of growth a year, and take about 14 years to get back to normal. Kane and Jones (2012) also evaluate the effects of the deployment of U.S. troops in conflict countries and find a positive long-term economic effect. A similar conclusion is reached by Kane (2012).

To our best knowledge, this study is the first to consider a foreign intervention as a form of natural experiment, and consequently to use tools for impact evaluation to assess its effect on economic outcomes.

Synthetic controls

The synthetic control method provides a data-driven, objective approach for constructing a suitable counterfactual with non-experimental, aggregate data. The synthetic control is a weighted average of several control units that is constructed to be approximately equal to the intervened unit regarding a number of pre-intervention characteristics. A critical assumption is that a synthetic control that is approximately equal to the intervened unit in the pre-intervention period would have also been approximately equal to the intervened unit in the post-intervention period, had it not been for the intervention. Thus, the effect of interest is estimated as the post-intervention differences between outcomes of the intervened and the synthetic control.

The popularity of this approach has increased since first introduced by Abadie and Gardeazabal (2003). It has been widely used to evaluate the effect of exogenous events in political sciences. Among others, Dorsett (2013) measures the impact of terrorism in Northern Ireland on its economic performance, in a similar way as Abadie and Gardeazabal (2003) do for the Basque Country; Kennedy (2013) assesses the effect of color revolutions on a series of political liberalization indicators in Serbia; Bove and Nisticò (2014) test whether coups d'état have a positive effect on military spending; and Lee (2014) evaluates the effect of super-majority vote requirements on tax burdens. Moreover, the approach has proven successful to study the effects of passing specific laws on cigarette sales (Abadie et al., 2010), college enrollment (Klasik, 2013), and immigration (Bohn et al., 2014).

Likewise, the method has also found a wide range of economic applications. For instance, Sanso-Navarro (2011) estimates the cost of the UK not adopting the euro, in terms of U.S. foreign direct investment; Billmeier and Nannicini (2013) study the economic effects of the waves of liberalization during the 1980s and 1990s in developing countries; Jinjarak et al. (2013) evaluate the influence of the implementation of control policies, and their subsequent relaxation, on capital inflows in Brazil; Kiesel and Villas-Boas (2013) estimate the effects of the costs of labeling on consumer choice; Baccini et al. (2014) measure the effect of corporate tax cuts on foreign direct investment in Russia; Bassock et al. (2014) analyze whether universal public pre-school crowds out private provision of childcare in the U.S.; Cavallo et al. (2013) and Barone and Mocetti (2014) study the effects of natural disasters (earthquakes) on economic and social indicators; and Ando (2015) addresses whether the vicinity to a nuclear plant affects housing prices.

Other empirical applications include the estimation of the impact of police deployment on crime (DeAngelo and Hansen, 2014; Saunders et al., 2015) and the construction of counterfactuals to measure energy intensity (Jimenez and Mercado, 2014). To our best knowledge, this study is the first contribution within this growing body of literature that uses synthetic control methods to assess the economic effects of UN international administrations.

3 Synthetic control methods

Next we discuss methodological issues regarding the synthetic control approach, following mainly Abadie and Gardeazabal (2003) and Abadie et al. (2010). First, we introduce the rationale of the method for comparative case studies, and then we describe some practical guidance for its implementation. Finally, we discuss how inference can be carried out within this framework.

3.1 Framework

Consider a panel with information for $n+1$ countries and T time periods. For concreteness, we assume that only "country 0" is exposed uninterruptedly to the intervention of interest, from period T_0 onwards, where with $1 \leq T_0 < T$ is the number of pre-intervention observations. The remaining n countries can be used as possible controls, and we will refer to the group of potential control countries as the *donor pool*.

Define Y_{it}^N as the outcome of interest that would be observed in country i and in period t without any intervention, for units $i = 0,1,\ldots,n$, and $t = 1,2,\ldots,T$. Similarly, let Y_{0t}^I be the outcome observed in the country exposed to the intervention ($i = 0$), and in its aftermath $t = T_0+1,\ldots,T$. For countries that are

not affected by the intervention, we have that $Y_{it}^I = Y_{it}^N$ for all t, whereas for the intervened country $Y_{0t}^I = Y_{0t}^N$ in the pre-intervention period and $Y_{0t}^I = Y_{0t}^N + \alpha_t$ afterwards.[3] Here, α_t is the effect of the intervention for unit in period t, and the purpose is to estimate it for $t = T_0 + 1, \ldots, T$.

Clearly, $\alpha_t = Y_{0t}^I - Y_{0t}^N$. The fundamental inference problem is that, for a given t, we observe Y_{it} which is either equal to Y_{it}^N (for $i = 1, \ldots, n$ and all t, and for $i = 0$ only for $t = 1, \ldots, T_0$) or equal to Y_{it}^I (for $i = 0$ after the intervention, $t = T_0 + 1, \ldots, T$). In particular, we are not able to observe Y_{0t}^N after the intervention, and the synthetic control method provides a elegant solution to overcome this difficulty.

Consider that Y_{it}^N is determined by a factor model of the form:

$$Y_{it}^N = \delta_t + \boldsymbol{\theta}_t \boldsymbol{z}_i + \boldsymbol{\lambda}_t \boldsymbol{\mu}_i + \varepsilon_{it} \,, \tag{1}$$

where δ_t is a time-specific common factor; $\boldsymbol{z}_i$ is a $k_z \times 1$ vector with observable country-specific, time-invariant covariates; $\boldsymbol{\mu}_i$ is a $k_\mu \times 1$ vector with unobservable country-specific effects; $\boldsymbol{\theta}_t$ and $\boldsymbol{\lambda}_t$ are conformable vectors with time-varying loadings; and the zero-mean error term ε_{it} is identically and independently distributed across time and countries.[4]

Now consider a vector of dimension $n \times 1$ with weights $\boldsymbol{w} = (w_1, w_2, \ldots, w_n)'$ such that $0 \leq w_i \leq 1$ for $i = 1, 2, \ldots, n$ and $w_1 + w_2 + \cdots + w_n = 1$, and the weighted average (strictly speaking, a convex linear combination)

$$Y_{0t}^w = \sum_{i=1}^{n} w_i Y_{it} = \delta_t + \boldsymbol{\theta}_t \sum_{i=1}^{n} w_i \boldsymbol{z}_i + \boldsymbol{\lambda}_t \sum_{i=1}^{n} w_i \boldsymbol{\mu}_i + \sum_{i=1}^{n} w_i \varepsilon_{it} \,. \tag{2}$$

Note that if $\boldsymbol{w}$ is chosen such that

$$\sum_{i=1}^{n} w_i \boldsymbol{z}_i = \boldsymbol{z}_0 \,, \tag{3}$$

and

$$\sum_{i=1}^{n} w_i \boldsymbol{\mu}_i = \boldsymbol{\mu}_0 \,, \tag{4}$$

then the difference $Y_{0t}^N - Y_{0t}^w$ will be equal to a random variable with zero mean. Thus, under (3) and (4), Y_{0t}^w is an unbiased predictor of the unobservable Y_{0t}^N, and so

$$a_t = Y_{0t} - Y_{0t}^w \tag{5}$$

is an unbiased estimator of α_t, for $t = T_0 + 1, \ldots, T$. However, choosing $\boldsymbol{w}$ to satisfy (4) is not feasible since $\boldsymbol{\mu}_i$ is unobservable. Abadie et al. (2010, Appendix B) show that, under mild regularity conditions and a relatively large value of T_0, this requirement can be exchanged with

$$\sum_{i=1}^{n} w_i Y_{it} = Y_{0t} \quad \text{for} \quad t = 1, 2, \ldots, T_0 \,. \tag{6}$$

[3] Note that it is assumed that the intervention has no effects before the implementation period. This carries no loss of generality, since an intervention whose effects manifests before its implementation can be modeled by simply redefining T_0.

[4] It is interesting to note that (1) generalizes the traditional "fixed-effects" model, thoroughly applied in the impact evaluation literature. This model is obtained by imposing the restrictions $\boldsymbol{\lambda}_t = \boldsymbol{\lambda}$ for all every t, so it allows for the presence of unobserved confounding factors but restricts their effects to be constant through time; hence, the confounding factors $\boldsymbol{\mu}_i$ can be eliminated by differentiating over time. In contrast, differentiating would not necessarily eliminate $\boldsymbol{\mu}_i$ in a model like (1) that allows for the effects of unobserved characteristics to vary over time. Furthermore, Abadie et al. (2010) show that the synthetic control method is still valid in more general cases, for instance when ε_{it} is autocorrelated.

The linear combination Y_{0t}^{w} represents the synthetic control and serves as an observable counterfactual for the post-intervention outcome with no intervention Y_{0t}^{I}. Requirements (3) and (6) ensure that the synthetic control replicate the pre-intervention history of the outcome variable, and also that it is comparable to the intervened country along the dimensions collected in $\boldsymbol{z}_0$. Note that (1) suggests that the covariates in vector $\boldsymbol{z}_i$ are not affected by the intervention, so that this vector should also contain pre-intervention quantities. Thus, since the choice of $\boldsymbol{w}$ is made to set $Y_{0t} - Y_{0t}^{w}$ equal to zero in the pre-intervention period, and since (1) remains a valid representation of the counterfactual Y_{0t}^{N}, any difference $Y_{0t} - Y_{0t}^{w}$ after the intervention can be attributed, precisely, to the intervention.

3.2 Implementation

Equations (3) and (6) hold exactly if and only if $(Y_{01},\dots,Y_{0T_0},\boldsymbol{z}_0')$ belongs to the convex hull of $\{(Y_{11},\dots,Y_{1T_0},\boldsymbol{z}_1'),(Y_{21},\dots,Y_{2T_0},\boldsymbol{z}_2'),\cdots,(Y_{n1},\dots,Y_{nT_0},\boldsymbol{z}_n')\}$. Among other things, this means that the values of the pre-intervention outcome and of the chosen covariates for the intervened unit should not be outliers relative to the donor pool. Otherwise, we would not be able to compute a counterfactual as a convex linear combination of the observations from the control group, rendering a synthetic control not comparable to the intervened unit. This resembles the "common support" condition sought in the impact evaluation literature (see, *inter alia*, Imbens and Wooldridge, 2009), and should serve as guidance to the researcher in choosing the dimensions in $\boldsymbol{z}_i$ along which the intervened country is to be compared to the synthetic control.

Yet, it is often the case that no weights $\boldsymbol{w}$ exist such that (3) and (6) are satisfied, and in practice $\boldsymbol{w}$ chosen such that these conditions hold approximately. To elaborate, define a $(k_z + k_y) \times 1$ vector $\boldsymbol{x}_0$ that collects the k_z values in $\boldsymbol{z}_0$ and also $k_y \leq T_0$ linear combinations of the pre-intervention outcome values $Y_{01},\dots,Y_{0T_0}$.[5] Similarly, let $\boldsymbol{X}$ be a $(k_z + k_y) \times n$ matrix such that its i-th column is defined analogously to $\boldsymbol{x}_0$ but contains data of country i from the donor pool. Thus, $\boldsymbol{Xw}$ is a $(k_z + k_y) \times 1$ vector with the synthetic control values. In order to measure the discrepancy between $\boldsymbol{x}_0$ and $\boldsymbol{Xw}$, consider the distance function

$$Q_1(\boldsymbol{w},\boldsymbol{V}) = (\boldsymbol{x}_0 - \boldsymbol{Xw})'\boldsymbol{V}(\boldsymbol{x}_0 - \boldsymbol{Xw}), \tag{7}$$

where $\boldsymbol{V}$ is a $(k_z + k_y) \times (k_z + k_y)$ symmetric, positive semidefinite matrix. Then, for a given matrix $\boldsymbol{V}$, the weights of the synthetic controls can be chosen as the solution of the quadratic program: minimize $Q(\boldsymbol{w},\boldsymbol{V})$, with respect to $\boldsymbol{w}$, subject to the restrictions $0 \leq w_i \leq 1$ for $i = 1,2,\dots,n$ and $w_1 + w_2 + \cdots + w_n = 1$.[6]

Although the inferential procedures described later are valid for any $\boldsymbol{V}$, the choice of $\boldsymbol{V}$ would affect the solution for $\boldsymbol{w}$, called $\boldsymbol{w}(\boldsymbol{V})$, and thus the properties of the synthetic control. Often $\boldsymbol{V}$ is taken to be a diagonal matrix, so its (j,j)-th entry reflects the relative importance assigned to the j-th variable in $\boldsymbol{x}$, when measuring the discrepancy between $\boldsymbol{x}_0$ and its counterfactual $\boldsymbol{Xw}$. As stated in Abadie and Gardeazabal (2003), the choice of $\boldsymbol{V}$ could be subjective, reflecting our prior knowledge of the relative predictive power of the variables in $\boldsymbol{x}$. On the other hand, the choice of $\boldsymbol{V}$ can also be datadriven. The predominant approach (see, *inter alia*, Abadie et al., 2010; Cavallo et al., 2013) is to choose $\boldsymbol{V}$ such that the mean squared prediction error (MSE) of the outcome variable in the pre-intervention period, which is estimated by

$$Q_2(\boldsymbol{V}) = \frac{1}{T_0}\sum_{t=1}^{T_0}\left(Y_{0t} - \sum_{i=1}^{n} w_i(\boldsymbol{V})Y_{it}\right)^2, \tag{8}$$

[5] The use of linear combinations of the pre-intervention outcomes, instead of the values themselves, is for computational simplicity. When n and T_0 are relatively large, the computational burden of the minimization problems below can be substantially reduced through this device. A common practice would be to choose *some* of the pre-intervention outcome values and check if (6) holds approximately for the resulting weights.

[6] Abadie et al. (2014) draw some parallels between this approach and standard regression, where $\boldsymbol{w}$ is estimated from the linear projection of $\boldsymbol{x}_0$ onto $\boldsymbol{X}$, without constraining the elements of $\boldsymbol{w}$ to be between 0 and 1. Thus, regression weights may take on negative values or values greater than one, and the resulting synthetic control would actually be an *extrapolation* of $Y_{1t},Y_{2t},\dots,Y_{nt}$, beyond the support of the comparison units. This renders Y_{0t}^{w} incomparable to Y_{0t}^{I}, even though the regression may produce a perfect (pre-intervention) fit. The restrictions $0 \leq w_i \leq 1$ for $i = 1,2,\dots,n$ prevent such extrapolation biases.

is minimized. The values of the minima of $Q_1(\cdot)$ and $Q_2(\cdot)$ can be used to assess the quality of the approximation of conditions (3) and (6), and may serve as guidance about whether the characteristics of the intervened unit are similar enough to those of the synthetic control. In some cases, the fit can be very poor even after the double minimization implied by the process of computing the synthetic control. In such situations, Abadie et al. (2010) recommend simply to refrain from using the synthetic control method.

3.3 Inference

Apart from the point estimation of the intervention effect, it is also desirable to compute measures of uncertainty around it. The estimation procedure, however, is highly nonlinear and, to date, there are no analytical approximations to such measures. An approach to assessing the sampling variability around a_t, borrowed from the impact evaluation literature (see Imbens and Wooldridge, 2009), would be to approximate the empirical distribution of a_t by resampling the members of the original donor pool, and then by repeating the estimation procedure using the resampled data.[7]

As discussed in Abadie et al. (2014), if the selection of comparison units in the donor pool was the result of probabilistic sampling, the percentiles of the empirical distributions across samples would form standard confidence intervals for the intervention effect. However, if we refrain from placing such probabilistic interpretation on the donor pool, these percentiles still serve to conduct a useful sensitivity analysis. In particular, our confidence that a baseline estimated intervention effect found for the intervened unit truly reflects a feature of the data, would be reinforced by observing such an effect across perturbations to the structure of the donor pool.

A complementary mode of inference, favored *inter alia* by Abadie and Gardeazabal (2003) and Abadie et al. (2014), is to conduct falsification or refutability exercises or "placebo runs". The idea is to repeat the estimation procedure on countries from the donor pool that, by construction, have not been exposed to the intervention. Then, if a particular synthetic control estimate truly reflects the impact of the intervention, no such impact should appear in the corresponding placebo run, where the intervention is artificially reassigned to a control unit.

4 Empirical exploration

In this section, we report the main results of this study. We first describe the data used in our empirical application, and then present the estimations and sensitivity analyses for the UN intervention effects in Kosovo and East Timor. The intervention effect is found to be negative in the case of Kosovo, and positive in the case of East Timor. These findings are robust to various deviations from the baseline results, remain after resampling the members of the donor pool, and are confirmed by sensible falsification exercises.

4.1 Data and definitions

In our empirical work, several decisions were strongly influenced by data availability, which is especially limited in the case of Kosovo since pre-intervention data (i.e., in the 1990s) are scarce. Even though data availability is considerably wider in the case of East Timor, in order to give a homogenous treatment to both countries under study, we decided to use for this country the same sources of information and variables as for Kosovo.

Including Kosovo and East Timor, the dataset consists of a panel of 115 countries (essentially emerging market economies) and 21 yearly observations, from 1990 to 2010. For the sake of ensuring comparable

[7] This resembles a non-parametric bootstrap. Abadie and Imbens (2008) show that bootstrapping is not valid for estimators based on a small numbers of units in the control group. As stated in Imbens and Wooldridge (2009), even though it is likely that the problems that invalidate the bootstrap disappear if the number of units in the control group increases (as would be the case for the synthetic control method), it is still unclear whether the bootstrap is an effective method for, say, constructing confidence intervals. A related method, known to be valid, is subsampling or "m out of n" bootstrap, advanced in Politis and Romano (1994). In our empirical exploration, we follow the subsampling approach by choosing m randomly.

units in the donor pool, and hence reducing interpolation biases, OECD countries (with the exception of the former Yugoslav republic of Slovenia) are excluded since most of them are middle-to-high income, consolidated economies. Similarly, large countries like Brazil, China, India, Russia and South Africa are also excluded because of their sheer size, which makes them not comparable to the small intervened countries under analysis. Countries with a population under 100,000 inhabitants have also been removed from the sample.

On the other hand, countries in the donor pool should not have been affected by a conflict or similar event during the post-intervention period. Thus, we remove countries that suffered, either directly or indirectly, from violent conflicts in the 2000s: Bangladesh, Burundi, Central African Republic, Chad, Colombia, Egypt, Eritrea, Georgia, Guinea, Haiti, Honduras, Iraq, Kenya, Liberia, Libya, Mali, Nepal, Nigeria, Pakistan, Sierra Leone, Somalia, Sri Lanka, Syria, and Yemen. Of course, Kosovo and East Timor are removed from the donor pool.

However, all countries in the "neighborhood" of the intervened countries are included in the donor pool. For Kosovo, Albania and the former Yugoslav countries (Bosnia and Herzegovina, Croatia, Macedonia, Montenegro, Serbia, and Slovenia). For East Timor, the Southeast Asian countries (Brunei, Cambodia, Indonesia, Laos, Malaysia, Myanmar, Philippines, Singapore, Thailand, and Vietnam). We have included them to formally evaluate whether neighbor countries, which is the "natural" comparison group in comparative studies, actually serve as good controls.

Data for Kosovo and East Timor are essentially available since 1990, so we consider a panel centered at the intervention year in 2000, with a 10-year window for the pre-intervention (1990 to 1999) and the post-intervention (2001 to 2010) periods. This choice for the length of the post-intervention period, also inspired by Abadie et al. (2014) in their study for the Unification of Germany, seems reasonable since we expect major changes in legal, political and economic systems to be noticeable in the medium-to-long term, several years after the intervention date.

The outcome variable is a broad measure of economic performance: the logarithm of per capita Gross National Income (GNI) in U.S. dollars, which, for brevity, will be referred to as "per capita GNI".[8] Since our purpose is to estimate the effect of the interventions on the medium-to-long-run prospects of the economy, we disregard short-term and cyclical variations in the data. To this end, Y is taken to be the trend for per capita GNI, as measured by the celebrated Hodrick and Prescott (1997) filter using a smoothing parameter of 6.25, following the analysis in Ravn and Uhlig (2002) for annual data.

As predictors in z, we consider the pre-intervention averages of variables that are persistent (likely to change little through time so they may be treated, for all that matters, as time-invariant), and that can be regarded as basic predictors of future economic growth. A first set of variables are of economic nature and are provided along with per capita GNI, by UNdata (2015). These are investment (gross capital formation) and total sectoral value added, measured as percentage of GDP. The composition of GDP, which roughly describes the "type" of economy, is the following: (i) agriculture (including hunting, forestry and fishing); (ii) construction; (iii) manufacture; (iv) mining (including other extractive industries and utilities); (v) transportation (including storage and communications); and (vi) services (including wholesale, retail, restaurants and hotels).

A second set of variables, from the World Bank (2015), includes basic demographic characteristics: (i) the logarithm of population density (habitants per square km of land area); (ii) the fertility rate (births per woman); (iii) life expectancy at birth (in years); and (iv) the dependency ratio (population younger than 15 or older than 64 to the working-age population).

[8] We have to content ourselves by working with nominal per capita GNI, as data on price deflators are unavailable for Kosovo and East Timor in the 1990s. We attempt, nonetheless, to make adjustments for inflation (π) and depreciation (δ) on the average nominal growth rates of the post-intervention period (g^n), to obtain an approximation of the real growth rates (g^r): $g^r = g^n + \delta - \pi$ (in this period, the official currency in Kosovo was the Euro; and in East Timor, the US dollar). Such adjustments would also be valid to approximate real intervention effects *only* under the (unverifiable) assumption that the synthetic control mimicked the pre-intervention path of the *real* as well as the nominal per capita GNI. Data on inflation and depreciation for the period 2003 to 2010 are readily available from the International Financial Statistics database of the International Monetary Fund.

Finally, together with z, vector x also includes three points of the pre-intervention outcome variable: per capita GNI in 1990 (at the beginning), in 1995 (in the middle) and in 1999 (at the end). Apart from the mechanic requirement of approximating the pre-intervention path of the outcome variable, these initial levels of per capita GNI are included because they should be relevant predictors of future developments of per capita GNI, as suggested by the widely accepted notion of convergence in economic growth theory.

The data for potential predictors in x enter our analysis in standardized form; i.e. with a sample average of zero and a sample standard deviation of one across all countries in the panel. Besides easing the visualization of results, this standardization prevents a given variable in x to be artificially dominant in the analysis below (especially in the choice of matrix V) due to differences in scale.

Figure 2 shows the sample distributions of the standardized variables in x as box plots. The boxes are limited by the 25th and 75th percentiles, whereas the whiskers are bounded by the 10th and 90th percentiles. For the purpose of this study, a value beyond these whiskers is considered an outlier. The position of the values of Kosovo (K) and East Timor (T) within the empirical distributions are also marked. A glance at these distributions reveals that the observations of the intervened countries can be regarded as "typical" from the cross-section of countries in most of the available dimensions. One remarkable exception occurs in the case of construction as a percentage of GDP, where both countries appear to have abnormally high pre-intervention values. Hence, this variable is excluded from the analysis. Similarly, the dependency ratio is excluded as a predictor in the case of Kosovo, whereas manufacture as a percentage of GDP is removed from the predictor set for East Timor.

4.2 Results for Kosovo

Base results

Figure 3(a) shows the path of per capita GNI for Kosovo and its synthetic control. The latter follows very closely the evolution of the former in the pre-intervention period, especially during the second half of the 1990s. The comparison between both paths unveils a negative effect of the intervention. In the pre-intervention period, per capita GNI in Kosovo and in Synthetic Kosovo decreased at a similar average annual rate of about 10 percent. In the post-intervention period, per capita GNI grew at an average annual rate of 10.0 percent in Kosovo, and of 12.8 percent in Synthetic Kosovo. Thus, unlike the positive rough estimates discussed in Figure 1, once the comparison is based on an adequate control, we find that the intervention was followed by a lower GNI per capita growth of 2.8 percent per annum.[9]

The intervention effects display interesting dynamics over time. Initially, between 2001 and 2002, the intervention effect is virtually nil. Then, most of the post-intervention negative effect takes place from 2003 to 2007, where Synthetic Kosovo outgrows actual Kosovo. A partial reversion of the negative effect is observed from 2008, the year Kosovo declared its independence, with actual Kosovo recovering and even outperforming the synthetic control.

These results are robust to relevant variants. Figure 3(b) further compares the outcome from Kosovo to that of alternative counterfactuals obtained by reestimating the synthetic control after removing the predictors in x one at a time. The negligible impact on the estimations of removing a predictor is remarkable. This is so because we have used a relatively large set of predictors in the baseline estimation (the comparisons between the intervened country and its potential controls are made among 12 dimensions), along with the fact that no single predictor seems to be dominant in determining the post-intervention path of the synthetic control.

Figure 3(a) also shows the path of alternative synthetic controls, based on suboptimal choices on w. The first one, dubbed "Regional synthetic", is the result of estimating the counterfactual from a donor pool of exclusively neighbor countries. Even though it does a decent job in replicating the path of the pre-intervention outcome in the first half of the 1990s, well before the intervention took place, its performance is considerably

[9] In the post-intervention period, average depreciation and inflation amounted, respectively, to −3.9 and 1.5 percent in Kosovo; and to 4.1 and 12.2 percent in synthetic Kosovo. Thus, in terms of the real per capita GNI growth rate the intervention effect would be close to zero: [10.0 − 3.9 − 1.5] − [12.8 + 4.1 − 12.2] = −0.1 percent.

less satisfactory in the late 1990s. The second one, "Control average", uses equal weights for all units in the full donor pool. The results highlight the inappropriateness of these alternative controls, as post-intervention outcomes are, in all likelihood, also reflecting persistent pre-intervention discrepancies.

To gain a deeper understanding of the workings of the synthetic control method, Table 1 presents, in panel (a), the composition of several synthetic controls; in panel (b), their corresponding measures of goodness of fit; and in panel (c), the comparison between the pre-intervention values of the predictors $\boldsymbol{x}$. The first three columns focus on the neighborhood of Kosovo. In case (1), a synthetic control is obtained as the simple average of Albania, Serbia, Macedonia, and Montenegro; in case (2) the synthetic control is optimized using these four countries as the donor pool; and in case (3) Albania and all former Yugoslav countries are considered.

Even though the optimized synthetic controls (2) and (3) fare much better than the simple average, as both the pre-intervention root MSE (RMSE, the square root of Q_2) and the distance function Q_1 are substantially smaller, their performance is still poor: the minimized pre-intervention RMSE is sizeable, and there remain certain dimensions in vector $\boldsymbol{x}$ for which the synthetic control significantly differ from the actual; namely, the share of transportation to GDP and the population density. Thus, despite what conventional wisdom might indicate, the neighbor countries do not comprise a set of fully comparable units to Kosovo.

Case (4) illustrates rather eloquently the potential gains of seeking comparable units in the complete donor pool, beyond the geographical or historical neighborhood of the intervened country. It shows how an optimized average between Serbia, which did not feature in cases (2) or (3), and Angola, a seemingly unrelated country, reduces considerably (by a third) the pre-intervention RMSE. Cases (5), (6), and (7) show alternative combinations of countries (including the African Ethiopia and São Tomé and Príncipe, and the former Soviet republics of Kyrgyzstan and Ukraine) that further improve the pre-intervention RMSE. Finally, case (8) shows the composition of Synthetic Kosovo, whose path is displayed in Figure 3(a). Even though it may have a slightly higher pre-intervention RMSE, it has a smaller Q_1, and a fully comparable set of predictors $\boldsymbol{x}$.

It is worth mentioning that, in our exploration, it was generally difficult to obtain a clearly dominant choice of $\boldsymbol{w}$. Even though important findings such as the predominance of Serbia would remain, slight variations in the countries included in the donor pool or other parameters of the optimization setup, would render different compositions for the synthetic control. For instance, consider cases (5) and (7), where Angola and São Tomé and Príncipe are replaced by Ethiopia, or cases (6) and (7), where Kyrgyzstan is replaced by a combination of Ethiopia and Ukraine.

This phenomenon is akin to overfitting, given the large size of the donor pool. Different solutions to a relatively complex optimization problem would correspond to close local minima, where the exclusion of a given unit from the donor pool can be compensated by the inclusion of the linear combinations of other units, as the examples above illustrate. Yet, again as in overfitting, even though the choice of $\boldsymbol{w}$ may be sensitive, the resulting predictions $\boldsymbol{X}\boldsymbol{w}$ and the path of the synthetic controls are not. Differences between, say, 2.5 and 3.5 in the pre-intervention RMSE are inconsequential, and are even difficult to detect visually.

Resampling

Figure 3(c) shows the evolution of the intervention effect, the deviation of Kosovo's per capita GNI from the outcome of the synthetic control, along with the cross-sectional percentiles calculated by considering, through resampling, synthetic controls optimized along various donor pools. The sampling consists of two stages: first a uniform random number m between 2 and n is chosen, and then m units are taken randomly without replacement from the original donor pool. This process is repeated 5,000 times.

Our sensitivity analysis indicate that the UNMIK can be associated to a robust "statistically negative" effect in Kosovo's economic performance, since its post-intervention outcomes are outperformed by that of counterfactual units that shared Kosovo's pre-intervention characteristics, but were not affected by the intervention.

Placebo runs

To conclude this section, we now present some falsification or placebo runs to inquire whether our results for Kosovo reflect the effect of the UNMIK or can be regarded as artifacts of the synthetic control algorithms. To this end, we repeat the analysis for selected countries that did not experience the intervention (at least directly), and compare the magnitude of the estimated intervention effect to that of Kosovo. In order to confirm that the estimated effect of the UNMIK in Kosovo is a salient feature of the data, we should not be able to find a similar negative intervention effect in the placebo runs.

We refine the resampling schedule used in the sensitivity analysis of the placebo runs. As before, a number m between 2 and n is chosen randomly, and then m units are sampled from the donor pool. To prevent the intervention effect in the placebo runs from reflecting persistent differences recorded at the time of the intervention, or at least to reduce this bias, we discard draws that produce a pre-intervention MSE 10 times higher, or more, than the baseline, full-sample counterpart. The process is repeated until 5,000 valid draws are obtained. This refinement attempts to automatize the practical guidance of not placing too much confidence in results based on runs with a poor pre-intervention fit.

Figure 4 presents the estimated intervention effects, along with their resampled "confidence" bands, for the neighbor countries of Kosovo: Albania and former Yugoslavia. It is important to recall that Serbia is the most important single unit of the synthetic control, whereas Albania, Bosnia and Herzegovina, and Montenegro appear as important units composing "regional" synthetic controls. Thus, even though are placebo runs are limited to a handful of countries, they have been carefully selected such that each of them can be regarded individually as a relevant comparison unit vis-a-vis Kosovo.

A noteworthy characteristic of these runs is that, despite the refinement described above, their sensitivity bands are much wider than those corresponding to Kosovo (shown in Figure 3), meaning that if any intervention effect exists, it may be difficult to detect and its estimation would be particularly sensitive to the structure of the donor pool. There is also a tendency to obtain positive point estimates, unlike the negative effect previously found for Kosovo. In the cases of Albania, Croatia, and to a lesser extent Montenegro, this positive intervention effect appears to be "significant". Thus, if this effect is an indirect consequence of the UNMIK in the neighbor countries, it reinforces the idea that the UNMIK effect in Kosovo is "significantly" negative. Recall that none of these countries are part of the synthetic control finally used for Kosovo, and therefore such positive effects would not influence in any way our baseline results for Kosovo.

All in all, the placebo experiments support the previous conclusion of a negative UNMIK effect.

4.3 Results for East Timor

Base results

Figure 5(a) shows the path for per capita GNI in East Timor and its synthetic control. Both follow a remarkably similar path throughout the pre-intervention period, growing at an average annual rate of about 7.25 percent. On the other hand, in the post-intervention period East Timor grows at the remarkable average rate of 19.5 percent, while the estimated growth rate for Synthetic East Timor is close to 11.5 percent. Thus, the synthetic control method suggests than in this case the intervention is followed by a higher GNI per capita growth rate of 8 percent per annum.[10]

The positive intervention emerges gradually: between 2001 and 2004, the intervention effect is very close to zero; after 2005, actual East Timor systematically outperforms its synthetic control. Moreover, these results are robust to variants in the predictor variables $\boldsymbol{x}$. Figure 5(b) shows the outcome of East Timor and of alternative counterfactuals obtained by reestimating the synthetic control after removing one predictor at a time. As in the case of Kosovo, the impact on the estimation of removing predictors is negligible.

[10] In the post-intervention period, average depreciation and inflation amounted, respectively, to 0.0 and 5.2 percent in East Timor; and to −0.4 and 7.1 percent in synthetic East Timor. Thus, in terms of the real per capita GNI growth rate the intervention effect would be higher: [19.5 + 0.0 − 5.2] − [11.5 − 0.4 − 7.1] = 10.3 percent.

It is interesting to note that, unlike the case of Kosovo, this point estimate is similar to the one described in Figure 1. The reason is that the neighbor countries appear to be more comparable for East Timor. Figure 5(a) shows the evolution of the "Regional" synthetic control which, at least visually, fares as well as Synthetic East Timor in replicating the pre-intervention path of the outcome variable. Using this control, the intervention effect is estimated as an increase in GNI per capita growth of 6.25 percent per year, even closer to the initial estimate of Figure 1. In spite of this good performance, the goodness of fit statistics of Synthetic East Timor turn out to be better.

Table 2 presents the composition of several synthetic controls. The first three cases focus on the neighborhood of East Timor, and case (3) corresponds to the regional synthetic control in Figure 5(a). The resulting distances Q_1 and Q_2 are rather small, suggesting that this counterfactual may be used as a basis for comparison. However, the synthetic control presents some important pre-intervention differences in the predictor variables with East Timor: namely, the share of mining in GDP and the fertility rate are considerably smaller, whereas life expectancy is somehow higher. Thus, as shown in the remaining columns of Table 2, it may be possible to obtain an even more comparable synthetic control by exploring the complete donor pool, something eventually achieved by the synthetic control in column (8). Relative to the regional control of column (3), the influence of neighbors Cambodia, Indonesia and Malaysia can be mimicked by a linear combination of Papua New Guinea (also a neighbor country), Guatemala in Latin America and the small African Kingdom of Swaziland.

Resampling

Finally, we can see in Figure 5(c) the difference between East Timor and its synthetic control, with the 95 percent "confidence bands". We observe that the positive effect of the UNTAET is due to the significant expansion experienced by the economy from 2005 onwards. This finding is in line with the prevailing view that the UN intervention in East Timor was a successful experience.

Placebo runs

Figure 6 shows the intervention effects and their resampled bands estimated for a selected group of Southeast Asian countries. In general, we fail to find a robust positive intervention effect in these cases. Among them, only Vietman contributes to the computation of synthetic East Timor.

A remarkable exception is Indonesia, for which our analysis reveals a robust significant impact of the UN intervention. Even though it is beyond the scope of the paper to provide a fully-fledged explanation for this result, a reasonable conjecture is that the end of an armed conflict where Indonesia was involved meant that resources initially allocated to the conflict could have been reallocated to more efficient or productive activities. It is also important to recall that Indonesia receives a weight of zero when computing the synthetic control, so this positive intervention effect does not influence the reported results for East Timor. Even if Indonesia did belong to the synthetic control, we observe that the effect of the intervention is much larger for East Timor, especially from 2008 onwards. Thus, if Indonesia were used in the computation of synthetic East Timor, the estimated UNTAET would be smaller, but would nevertheless remain positive and robust.

All in all, the placebo experiments support the previous conclusion of a positive UNTAET effect.

5 Closing remarks

A basic premise in intervention analysis is that we can infer on the effect of an intervention by comparing the outcomes of an intervened unit to those of a control that has not been affected by the intervention. A widespread practice in comparative case studies, based on a qualitative and somehow subjective approach to the estimation problem, is to consider neighbor countries as good candidates to compose a proximity control group. Under this view, we find positive effects, and even of comparable magnitudes, of the UN interventions in both cases under analysis.

The synthetic control approach, on the other hand, leaves the determination of a suitable control to a quantitative, data-driven, and objective procedure whereby the control is obtained as the solution of an optimization problem seeking to minimize the discrepancies between the intervened unit and its control, both in the pre-intervention path of the outcome variable and in pre-intervention covariates that serve as predictors of future outcomes. Following this reasoning, we find that in the case of Kosovo the neighbor countries (Albania and the former Yugoslav countries) maintain important differences with Kosovo: not only were they richer by the time of the UN intervention, but also tend to have a much lower population density and a different breakdown of their economic activities as a share of GDP. Likewise, in the case of East Timor, the neighbor countries (Southeast Asian economies) were also richer by the time of the UN intervention, and are much less dependent on mining activities, among other differences. In both cases, the synthetic controls, which are computed from a control group wider than the neighborhood of each country, are comparable to the intervened countries in all these dimensions.

Thus, once the intervention effect in assessed using comparable units, we conclude that after the UNMIK was established, Kosovo was outperformed in economic terms by its synthetic control (i.e., a negative intervention effect); in contrast, after the UNTAET was established, East Timor outperformed its synthetic control (i.e., a positive intervention effect). These results highlight the importance of country- and conflict-specific factors in determining whether a particular foreign intervention is successful, at least in terms of future economic growth. In addition, the results provide quantitative evidence to the widespread notion that, again in terms of economic growth, the UN intervention in Kosovo has failed to fulfill its initial goals, whereas the UN intervention in East Timor delivered the intended outcome.

Data availability posed important restrictions on the scope of our analysis. We believe that, as further data become available, an interesting research agenda would be to analyze the effects of UN international administrations on alternative outcomes that may reflect better the well-being of the affected populations, or the degree of violence and social unrest experienced by these societies after the intervention. Similarly, it would be fruitful to replicate our analysis using even better definitions of the units under study; for instance, by considering Serbian and Indonesian provinces as a controls, rather than neighbor countries. The synthetic control method provides a powerful analytical tool to pursue such research agenda.

References

Abadie, A., Diamond, A., and Hainmueller, J. (2010). Synthetic control methods for comparative case studies: Estimating the effect of California's tobacco control program. *Journal of the American Statistical Association*, 105(490):493–505.

Abadie, A., Diamond, A., and Hainmueller, J. (2014). Comparative politics and the synthetic control method. *American Journal of Political Science*, 59(2):495–510.

Abadie, A. and Gardeazabal, J. (2003). The economic costs of conflict: A case study of the Basque Country. *American Economic Review*, 93(1):113–132.

Abadie, A. and Imbens, G. W. (2008). On the failure of the bootstrap for matching estimators. *Econometrica*, 76(6):1537–1557.

Ando, M. (2015). Dreams of urbanization: Quantitative case studies on the local impacts of nuclear power facilities using the synthetic control method. *Journal of Urban Economics*, 85:68–85.

Baccini, L., Li, Q., and Mirkina, I. (2014). Corporate tax cut and foreign direct investment. *Journal of Policy Analysis and Management*, 33(4):977–1006.

Barone, G. and Mocetti, S. (2014). Natural disasters, growth and institutions: A tale of two earthquakes. *Journal of Urban Journal of Urban Economics*, 84:52–66.

Bassock, D., Fitzpatrick, M., and Loeb, S. (2014). Does state preschool crowd-out private provision? The impact of universal preschool on the childcare sector in Oklahoma and Georgia. *Journal of Urban Journal of Urban Economics*, 83:18–33.

Benvenisti, E. (1993). *The International Law of Occupation*. Princeton University Press.

Bertrand, M., Duflo, E., and Mullainathan, S. (2004). How much should we trust differences-in-differences estimates? *Quarterly Journal of Economics*, 119(1):249–275.

Billmeier, A. and Nannicini, T. (2013). Assessing economic liberalization episodes: A synthetic control approach. *Review of Economics and Statistics*, 96(5):983–1001.

Bohn, S., Lofstrom, M., and Raphael, S. (2014). Did the 2007 Legal Arizona Workers Act reduce the state's unauthorized immigrand population? *Review of Economics and Statistics*, 96(2):258–269.

Bove, V. and Nisticò, R. (2014). Coups d'état and defense spending: A counterfactual analysis. *Public Choice*, 161(3-4):321–344.

Cavallo, E., Galiani, S., Noy, I., and Pantano, J. (2013). Catastrophic natural disasters and economic growth. *Review of Economics and Statistics*, 95(5):1549–1561.

Collier, P., Chauvet, L., and Hegre, H. (2008). The security challenge in conflict-prone countries. *Copenhagen Consensus Center*.

Collier, P., Elliot, V., Hegre, H., Hoefler, A., Reynal-Querol, M., and Sambanis, N. (2003). *Breaking the Conflict Trap: Civil War and Development Policy*. World Bank Group and Oxford University Press.

DeAngelo, G. and Hansen, B. (2014). Life and death in the fast lane: Police enforcement and traffic fatalities. *American Economic Journal: Economic Policy*, 6(2):231–257.

Dorsett, R. (2013). The effect of the Troubles on GDP in Northern Ireland. *European Journal of Political Economy*, 29:119–133.

Fox, G. H. (2012). Transformative occupation and the unilateralist impulse. *International Review of the Red Cross*, 94(885):237–266.

Gurmendi, A. (2015). Your country, my rules. Can military ooccupation create successful transitions? *Georgetown Journal of International Law*, 46(4):979–1007.

Hodrick, R. J. and Prescott, E. C. (1997). Postwar U.S. business cycles: An empirical investigation. *Journal of Money, Credit and Banking*, 29(1):1–16.

Imbens, G. W. and Wooldridge, J. M. (2009). Recent developments in the econometrics of program evaluation. *Journal of Economic Literature*, 47(1):5–86.

Jimenez, R. and Mercado, J. (2014). Energy intensity: A decomposition and counterfactual exercise for Latin American countries. *Energy Economics*, 42:161–171.

Jinjarak, Y., Noy, I., and Zeng, H. (2013). Capital controls in Brazil - Stemming a tide with a signal? *Journal of Banking and Finance*, 37(8):2938–2952.

Kane, T. (2012). Development and US troop deployments. *Foreign Policy Analysis*, 8(3):255–273.

Kane, T. and Jones, G. (2012). U.S. troops and foreign economic growth. *Defence and Peace Economics*, 23(3):225–249.

Kennedy, R. (2013). Fading colours? A synthetic comparative case study of the impact of 'colour revolutions'. *Comparative Politics*, 31(2):153–163.

Kiesel, K. and Villas-Boas, S. B. (2013). Can information costs affect consumer choice? Nutritional labels in a supermarket experiment. *International Journal of Industrial Organization*, 42(3):151–160.

King, I. and Mason, W. (2006). *Peace at Any Price: How the World Failed Kosovo*. Crises in World Politics. Cornell University Press.

Klasik, D. (2013). The ACT of enrollment: The college enrollment effects of state-required college entrance exam testing. *Educational Researcher*, 42(3):151–160.

Lee, M. J. (2005). *Micro-Econometrics for Policy, Program, and Treatment Effects*. Advanced Text Series in Econometrics. Oxford University Press.

Lee, S. (2014). The effect of supermajority vote requirements for tax increase in California: A synthetic control method approach. *State Politics and Policy Quarterly*, 14(4):414–436.

Martin, I. and Mayer-Rieckh, A. (2005). The United Nations and East Timor: From self-determination to state-building. *International Peacekeeping*, 12(1):125–145.

Ofstad, O. (2012). Reconciliation and conflict resolution in East Timor. Working Paper (April), Oxford Institute for Ethics, Law and Armed Conflict.

Politis, D. N. and Romano, J. P. (1994). Large sample confidence regions based on subsamples under minimal assumptions. *The Annals of Statistics*, 22(4):2031–2050.

Ravn, M. O. and Uhlig, H. (2002). On adjusting the HP-filter for the frequency of observations. *Review of Economics and Statistics*, 84(2):371–380.

Roberts, A. (2006). Transformative military occupation: Applying the laws of war and human rights. *American Journal of International Law*, 100(580):580–622.

Sanso-Navarro, M. (2011). The effects on American foreign direct investment in the United Kingdom from not adopting the euro. *Journal of common market studies*, 49(2):463–483.

Saunders, J., Lundberg, R., Braga, A. A., Ridgeway, G., and Miles, J. (2015). A synthetic control approach to evaluating place-based crime interventions. *Journal of Quantitative Criminology*, 31(3):413–434.

Strohmeyer, H. (2001). Collapse and reconstruction of a judicial system: The united nations missions in kosovo and east timor. *American Journal of International Law*, 95(1):46–63.

Triponel, A. and Pearson, S. (2010). What do you think should happen? Public participation in transitional justice. *Pace International Law Review*, 22(1):103–144.

UNdata (2015). *UN Data portal*. United Nations Statistics Division (data.un.org).

Vickers, M. (1998). *Between Serb and Albanian: A History of Kosovo*. Columbia University Press.

Weinstein, J. (2005). Autonomous recovery and international intervention in comparative perspective. Working Paper 57, Center for Global Development.

Wilde, R. (2001). From Danzig to East Timor and beyond: The role of international territorial administration. *American Journal of International Law*, 95(3):583–606.

Wolfrum, R. (2005). International administration in post-conflict situations by the United Nations and other international actors. *Max Planck Yearbook of United Nations Law*, 9:649–696.

World Bank (2015). *World Development Indicators*. World Bank Group.

Table 1. *Results for Kosovo*

	Actual	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
(a) Synthetic weights									
Albania		0.25	0.22	0.00					0.00
Serbia		0.25	0.00	0.00	0.60	0.52	0.65	0.71	0.63
Bosnia and Herzegovina				0.40					0.00
Macedonia		0.25	0.00	0.00					0.00
Montenegro		0.25	0.78	0.60					0.00
Angola					0.40	0.21			0.07
Ethiopia								0.21	0.07
Sao Tome and Principe						0.16			0.05
Kyrgyzstan							0.35		0.12
Ukraine						0.11		0.08	0.06
(b) Goodness of fit									
Pre-intervention RMSE (Q_2)		24.2	14.5	17.1	5.0	3.5	2.8	2.5	2.9
Distance (Q_1)		0.82	0.69	0.29	0.32	0.20	1.13	0.80	0.11
(c) Pre-intervention standardized values of X									
Gross capital formation (% GDP)	0.05	−0.32	0.46	0.92	−0.29	−0.36	−0.64	−0.64	−0.50
Agriculture (% GDP)	−0.70	0.23	−0.20	0.09	0.47	0.28	0.24	0.04	0.06
Manufacture (% GDP)	−0.18	0.02	−0.63	−0.55	0.21	0.31	0.64	0.43	0.58
Mining (% GDP)	−0.03	−0.02	−0.05	−0.13	0.08	0.01	−0.44	−0.14	−0.12
Transportation (% GDP)	−0.68	0.16	0.93*	0.98*	0.33	−0.11	0.52	−0.15	0.04
Services (% GDP)	−0.32	−0.02	0.37	−0.14	−0.42	−0.29	−0.58	−0.76	−0.70
Life expectancy (years)	0.44	0.39	1.01	0.27	−0.06	0.09	0.46	0.23	0.28
Population density (logs)	0.86	−0.33	−0.97*	−1.20*	−0.26	−0.16	−0.25	−0.33	0.14
Fertility rate (children per women)	−0.57	−0.31	−0.38	−0.40	−0.01	−0.12	−0.49	−0.18	−0.65
Per capita GNI in 1990 (logs)	0.51	−0.06	0.18	0.08	0.13	0.37	0.18	0.33	0.47
Per capita GNI in 1995 (logs)	−0.02	0.05	−0.27	−0.60	0.10	0.36	0.19	0.35	−0.05
Per capita GNI in 1999 (logs)	−0.34	0.34	0.29	−0.35	0.17	0.44	0.19	0.35	−0.34

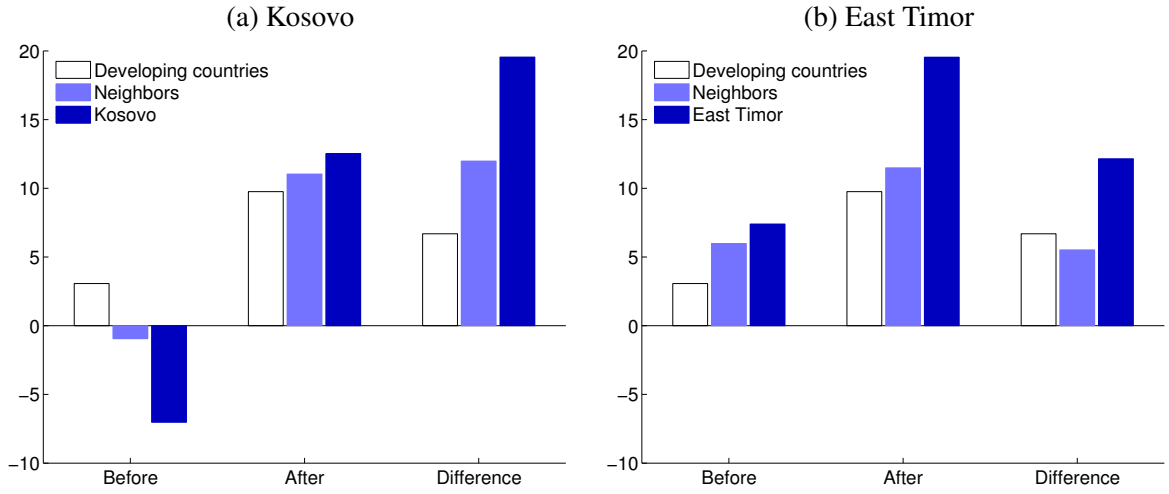
Notes: Panel (a): A missing (blank) value means that the particular country was not included in the donor pool, whereas a value of 0.00 means that the country was included in the donor pool, but "optimally" received a zero weight. The results in column (8) use the complete donor pool (115 countries) and all unreported countries received a weight of zero. Panel (b): the units of the RMSE are deviations with respect to the synthetic control, whereas the units of Q_1 are pure standardized values. Panel (c): "*" indicates that the absolute difference between the actual and the corresponding syntethic control is "large"; in particular, greater than 1.28, taking as a benchmark the standard normal distribution for which $\Pr(|z| \leq 1.28) = 0.80$.

Table 2. *Results for East Timor*

	Actual	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
(a) Synthetic weights									
Brunei Darussalam				0.00					0.00
Cambodia				0.35					0.00
Indonesia		0.20	0.38	0.05					0.00
Lao				0.00					0.00
Malaysia		0.20	0.00	0.15					0.00
Myanmar				0.00	0.59				0.00
Papua New Guinea							0.42		0.14
Philippines		0.20	0.25	0.00					0.00
Singapore				0.00					0.00
Thailand		0.20	0.00	0.18					0.00
Vietnam		0.20	0.38	0.28		0.46	0.37	0.35	0.39
Guatemala							0.21	0.46	0.22
Guinea-Bissau						0.11		0.19	0.10
Swaziland						0.44			0.15
(b) Goodness of fit									
Pre-intervention RMSE (Q_2)		71.1	1.5	0.7	1.7	1.1	1.7	1.8	0.8
Distance (Q_1)		0.62	1.46	0.03	0.95	0.44	1.04	0.81	0.02
(c) Pre-intervention standardized values of X									
Gross capital formation (% GDP)	0.46	0.07	−0.73	−0.14	−0.25	−0.37	−0.24	−0.47	−0.30
Agriculture (% GDP)	0.34	−0.15	0.41	0.29	0.28	−0.27	0.34	−0.20	0.46
Mining (% GDP)	1.03	−0.31*	−0.28*	−0.32*	−0.59*	−0.32*	0.03	−0.41*	−0.03
Transportation (% GDP)	−0.73	0.14	−0.65	−0.00	0.29	0.10	−0.31	−0.19	−0.86
Services (% GDP)	0.59	−0.13	−0.33	0.31	0.21	−0.45	−0.37	−0.07	0.27
Dependency Ratio	0.49	−0.26	0.08	0.32	−0.55	−0.02	−0.29	0.07	0.51
Life expectancy	−1.17	0.37*	0.21*	0.05	0.35*	0.50*	0.30*	0.44*	−0.03
Population density (logs)	0.05	0.24	0.66	−0.11	0.21	0.43	−0.11	0.36	0.32
Fertility rate (children per women)	1.40	−0.36*	−0.28*	−0.10*	−0.56*	−0.58*	−0.44*	−0.34*	0.23
Per capita GNI in 1990 (logs)	−1.16	0.11	−0.55	−0.47	−0.12	−0.10	−0.08	0.00	−1.16
Per capita GNI in 1995 (logs)	−0.64	0.12	−0.58	−0.41	−0.11	0.03	−0.08	0.10	−0.64
Per capita GNI in 1999 (logs)	−0.63	0.13	−0.62	−0.36	−0.11	0.23	−0.18	0.24	−0.62

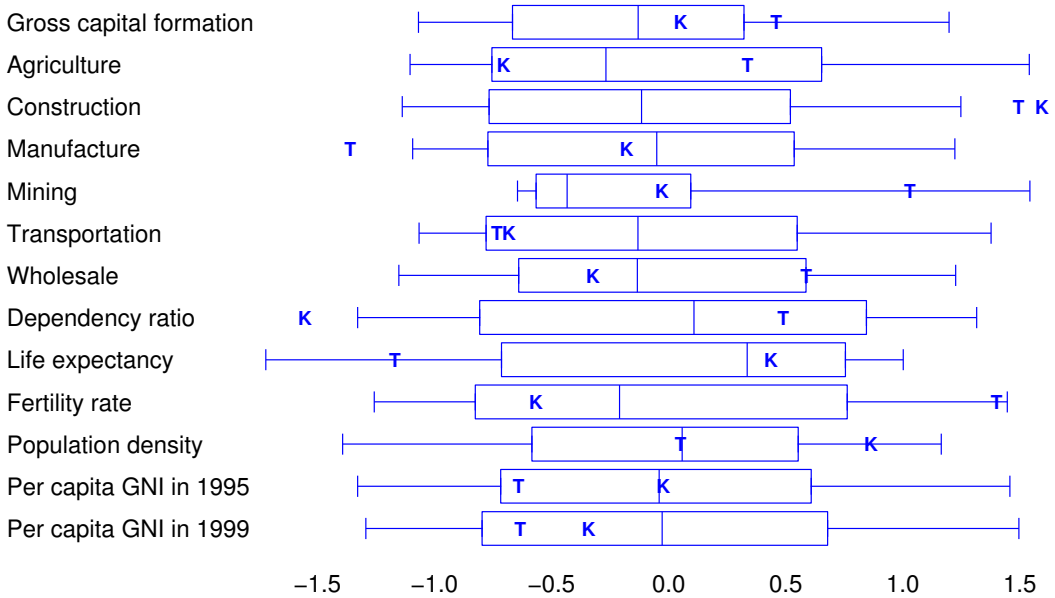
Notes: See notes to Table 1.

Figure 1. *Per capita GNI growth in intervened and control countries*



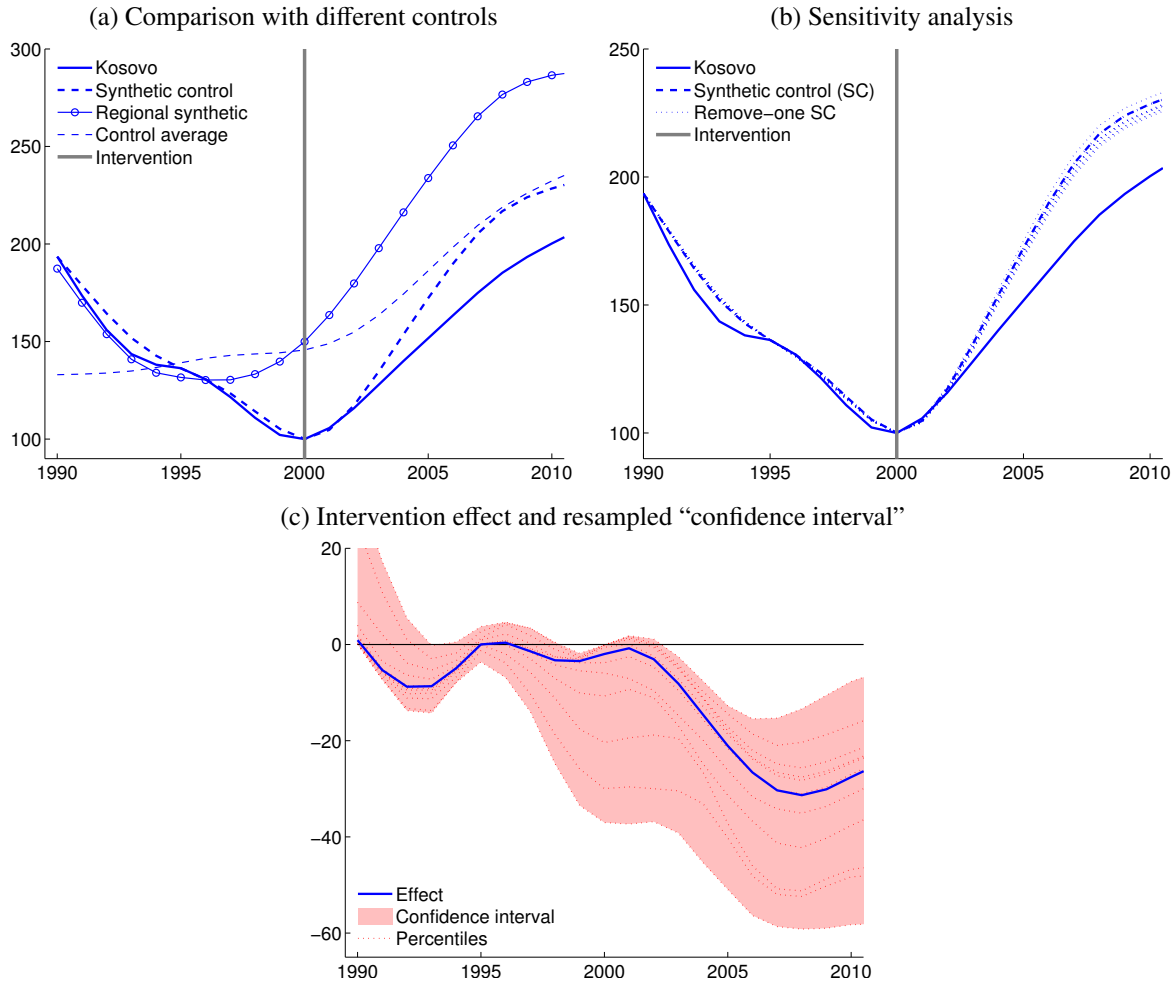
Notes: Data sources, variable definitions, and the list of neighbor countries are given in section 4.1. The figures are average growth rates of per capita GNI in US$, over the periods 1990-1999 ("Before") and 2001-2010 ("After").

Figure 2. *Cross-sectional distributions of growth predictors (pre-intervention values)*



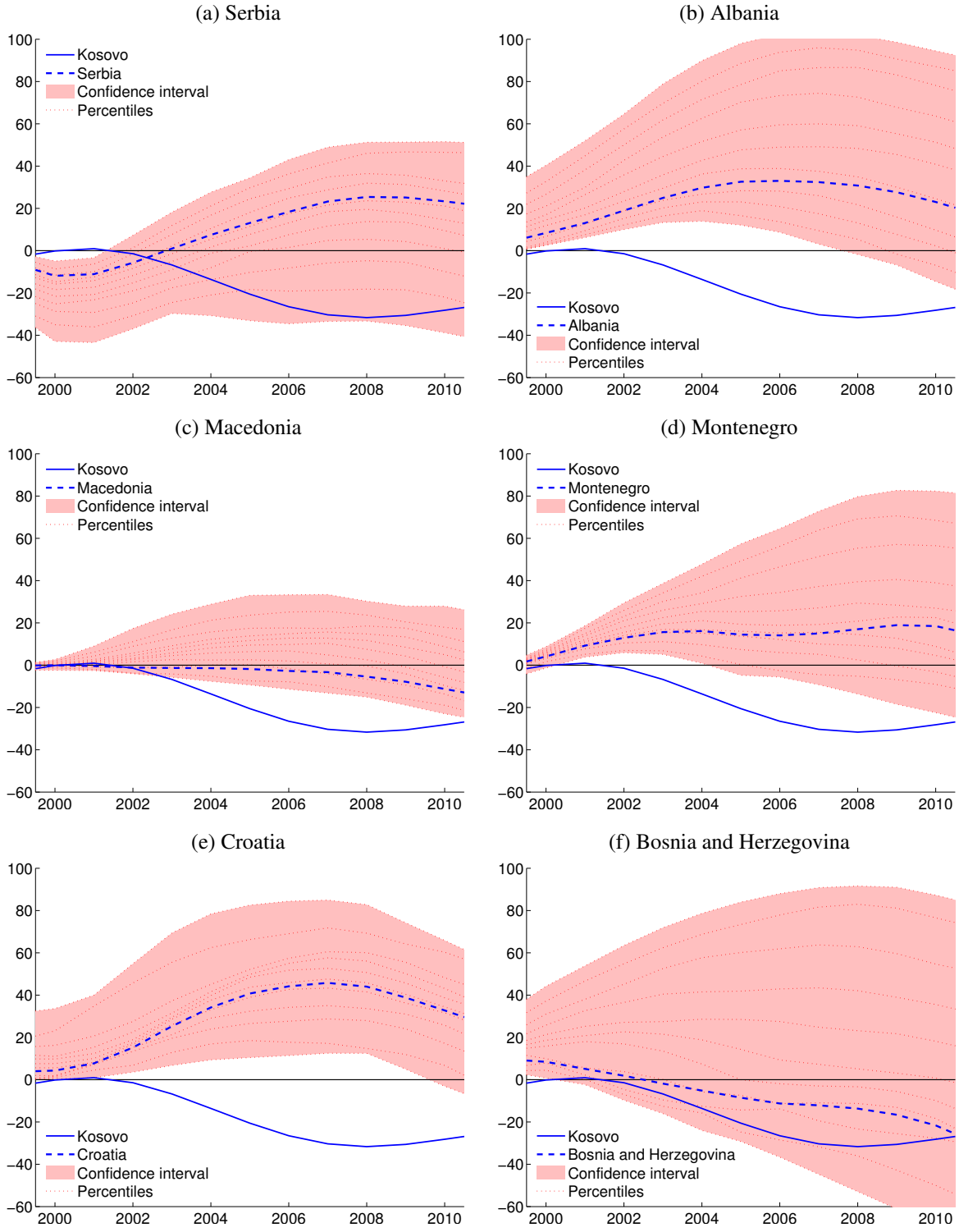
Notes: Standardized data (average of zero and standard deviation of one across all 117 countries in our sample). The limits of the boxes are the first and third quartiles, whereas the whiskers indicate 10th and 90th percentiles. The values of Kosovo [resp., East Timor] within the empirical distributions are marked with a "K" [resp., "T"].

Figure 3. *Estimation results for Kosovo*



Notes: Panels (a) and (b) are in logarithm scale times 100, and a constant was subtracted from each series such that the actual value of Y_{0t} in year 2000 is normalized to 100. Panel (c) shows percent deviations of Y_{0t} from its synthetic control Y_{0t}^{w}. The percentiles come from a resampling exercise where the synthetic control is reestimated using a "pseudo" donor pool that consists in a random number of units from the original donor pool. We use 5,000 repetitions.

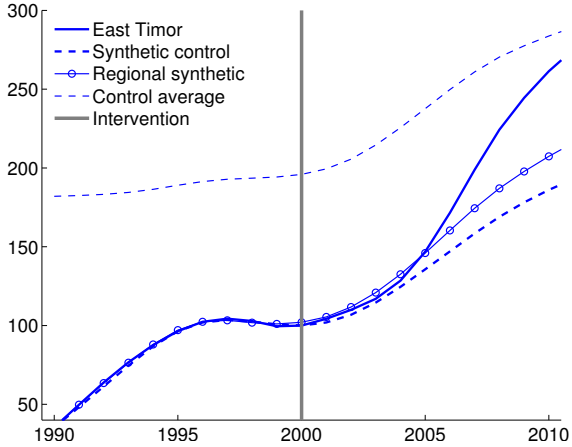
Figure 4. *Intervention effect from placebo runs for selected Kosovo neighbors*



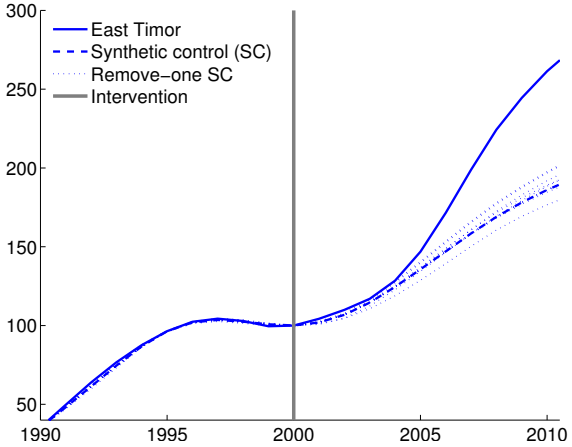
Notes: All panels show percent deviations of Y_{0t} from its synthetic control Y_{0t}^{w}. The percentiles come from a resampling exercise where the synthetic control is reestimated using a "pseudo" donor pool that consists in a random number of units from the original donor pool. Only draws such as the ratio of the pre-intervention MSE of the outcome to its full sample counterpart is less than or equal to 10 are kept, and the exercise stops when 5,000 such draws are obtained.

Figure 5. *Estimation results for East Timor*

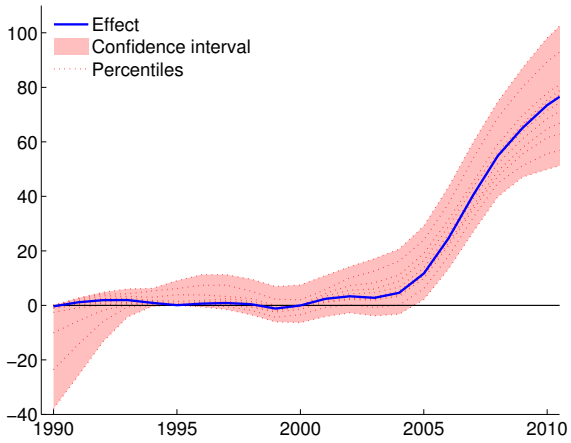
(a) Comparison with different controls



(b) Sensitivity analysis

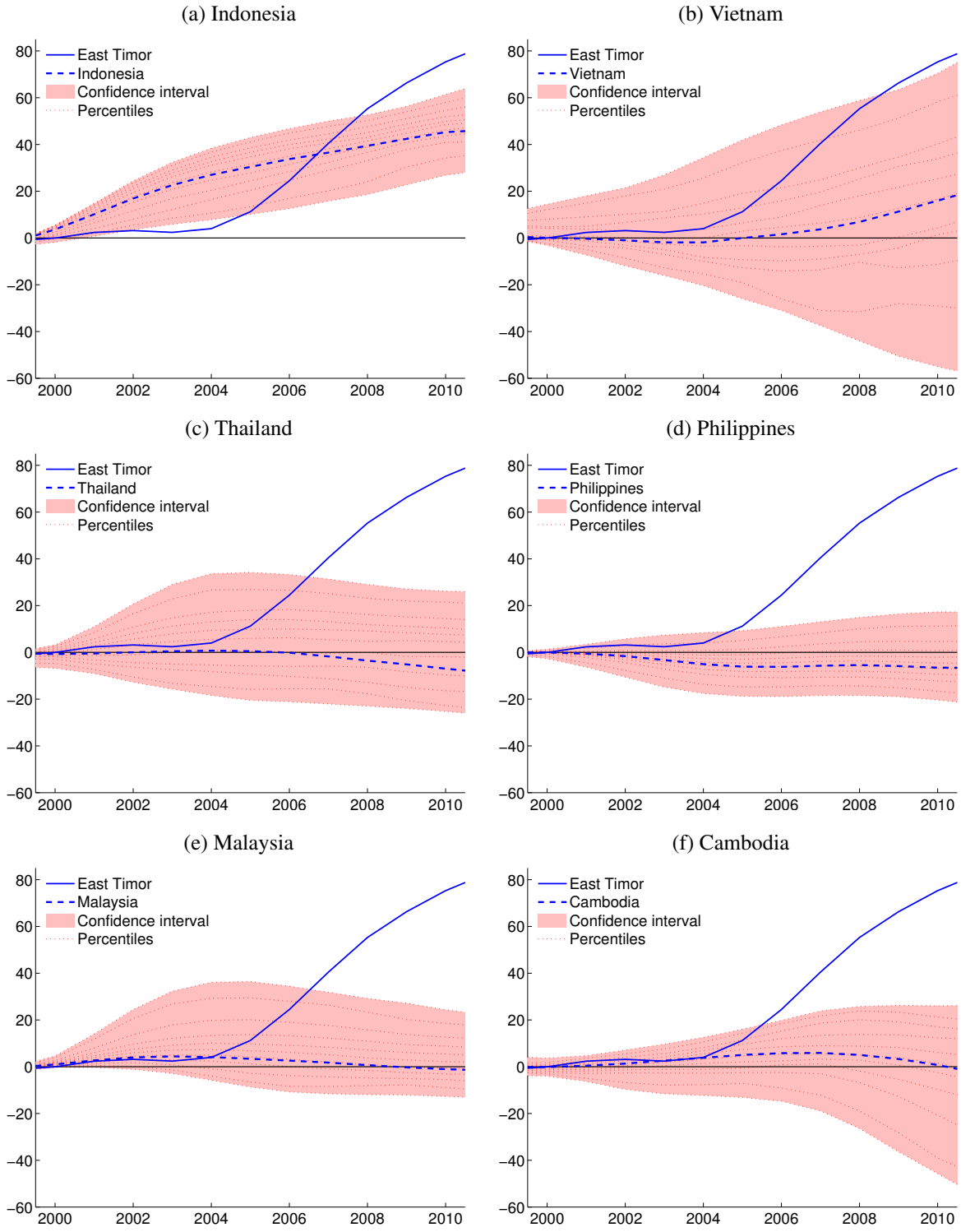


(c) Treatment effect and resampled confidence interval



Notes: See notes to Figure 3.

Figure 6. *Intervention effect from placebo runs for selected East Timor neighbors*



Notes: See notes to Figure 4.