This is an Accepted Manuscript of an article published by Taylor & Francis in Defence and Peace Economics, on 2019 avalaible at https://doi.org 10.1080/10242694.2018.1446241

The economic consequences of the Libyan Spring: A Synthetic Control Analysis*

Cruz A. Echevarría† and Javier García-Enríquez‡

Current version: 12 February 2018

Abstract

In 2011 a wave of revolutionary movements, the so-called Arab Spring, spread in the Middle East and North Africa. Libya was one of the most affected countries, ending Gaddafi's dictatorship after an international intervention and a civil war. This paper assesses the effects that this revolution had on Libyan economy. The analysis is made by means of the *synthetic control method* introduced by Abadie and Gardeazabal (2003). Our estimates for the 2011-2014 period show *i*) a cumulative loss in the growth rate of per capita real GDP of 64.15%; *ii*) a cumulative loss in per capita real GDP of 56,548 dollars; and *iii*) a cumulative loss in the aggregate real GDP of 350.5 billion dollars.

JEL classification: C23; Z3; F43; F47; O5

Keywords: Case Study; Synthetic Control Method; Treatment Effect; Arab Spring; Libya.

1 Introduction

The Arab Spring uprisings began in mid-December 2010 in Tunisia, forcing the president Zine El Abidine Ben Ali to abandon the country on 14 January 2011. This overthrow set a precedent for other countries in the region, and protests spread quickly in the Middle East and North Africa to other countries, although the degree of intensity varied across countries. While in some cases such as Egypt, Libya, Yemen and the aforementioned Tunisia their autocratic regimes were overthrown, other countries such as Jordan and Morocco also witnessed profound political changes, although their rulers remained in power. In some other countries, the Arab

*We are deeply indebted to Javier Gardeazabal for his comments and suggestions on an earlier draft. The usual disclaimer applies. This work was supported by the Spanish Ministry of Economy and Competitiveness and the European Fund of Regional Development under Grants ECO2015-64467-R and ECO2016-76884-P (MINECO/FEDER); and the Basque Government under Grants DEUI-IT-793-13 and DEUI-IT-783-13.

†Corresponding author. E-mail address: cruz.echevarria@ehu.eus Departamento de Fundamentos del Análisis Económico II. University of the Basque Country UPV/EHU. Avda. Lehendakari Aguirre 83, 48015 Bilbao, Spain.

‡E-mail address: javier.garcia@ehu.eus Departamento de Economía Aplicada III (Econometría y Estadística). University of the Basque Country UPV/EHU. Avda. Lehendakari Aguirre 83, 48015 Bilbao, Spain.

Spring did not go much beyond street protests and mild reforms as in Algeria, Lebanon, Oman, Qatar, Bahrain or Kuwait. The Syrian case can be considered as a special one as the uprising there became a violent armed conflict that continues at the time of writing [see Khan (2014), Hodler (2012), The Economist (2013)].

As is the case in most social movements, the Arab Spring can also be analyzed from an economic perspective in terms of both causes and consequences. Focusing on the economic consequences, the studies that have addressed this issue have mostly followed a purely descriptive approach, thereby forgoing a more rigorous, quantitative estimation.[1] This is the gap that this paper attempts to fill. More precisely, we aim to estimate the cost in terms of the gross domestic product lost as a result of the Arab Spring and the resulting revolution in one particular country, Libya, between 2011 and 2014.

Libya is a North African country that gained its independence from Italy in 1951. After a military coup King Idris I was overthrown in 1969 and Muammar Gaddafi became president. Under the Gaddafi's regime press freedom, trade unions and political opposition were banned. Moreover, the regime fought against dissidents both inside Libya and around the world. From an economic viewpoint Libya has the largest proven oil reserves in Africa and is among the ten largest oil producing countries, with 48,363 million barrels. Libya also has substantial natural gas reserves. In 2010 oil contributed 54% of total GDP and 83% of government revenues, being the main income source. This made Libya the richest North African country, with a GDP per capita of 29173 PPP US dollars in 2010. Moreover, its Human Development Index was 0.76, placing it within the high human development category (position 67th out of 188), and life expectancy at birth was 71.74. Libya also had a welfare system that allowed access to free education, free healthcare and financial aid for housing.[2]

Unlike the successful uprisings in Tunisia and Egypt (or even in Yemen), the Libyan rebellion resulted in a wholesale revolution, focusing on human rights abuses, political corruption and finally demanding the end of Muammar Gaddafi's rule. It was also more prolonged and violent, leading to a civil war and an international military intervention by NATO. On 20 August 2011 Tripoli was taken by the rebels and the Gaddafi regime collapsed. Gaddafi was captured and

[1] One exception is Groizard *et al.* (2016) which study foreign tourists' demand for travel to countries experiencing Arab Spring episodes by using a gravity model of bilateral tourism flows for Egypt, Syria, Tunisia and Yemen.

[2] The data come from different sources. OPEC Annual Statistical Bulletin 2016 for oil barrels reserves, http://www.opec.org/opec_web/en/data_graphs/330.htm; Federal Reserve Bank of St. Louis for the oil-revenue share of total government revenues, https://fred.stlouisfed.org/series/LBYGGRXOGDPXOPT; World Bank for oil rents as GDP share, http://databank.worldbank.org/data/reports.aspx?source=2&series=NY.GDP.PETR.RT.ZS&country=; World Bank for GDP, https://data.worldbank.org/indicator/NY.GDP.PCAP.PP.KD; United Nations Development Program, 2015, for the Human Development Index, http://hdr.undp.org/en/data; and World Bank for life expectancy, https://data.worldbank.org/indicator/SP.DYN.LE00.IN.

killed on 20 October 2011. The rebels, organized under the National Transitional Council, declared the liberation of Libya and the official end of the war on 23 October 2011.

After Gaddafi's fall local power centers emerged and different political parties, tribes and militias have been fighting to exert influence. At least two political bodies claim to be the government of Libya: one in Tripoli and another in Tobruk. Moreover, some parts of Libya are outside of either government's control, with various Islamist, rebel, and tribal militias administering some cities and areas. Under United Nations sponsorship an agreement to form a unified interim government was signed on December 2015. However, the repeated clashes around the country demonstrate the interim Government's inability to impose its authority.

Luckily, the economic damage due to the conflict was relatively small. Since the East was under rebel control from early on in the conflict, damage there was limited. Misrata and other towns did experience significant shelling during the fighting, but the loss of key infrastructure and manufacturing (minimal in Libya in any event) was not extensive. The hold-out towns of Sirte and Bani Walid were more heavily damaged in the final weeks of the war. In part, the low levels of physical damage are due to the fact that NATO planners went to great lengths to ensure that Libya's hydrocarbon industry was not seriously disrupted by military operations.

Even though NATO avoided targeting Libyan infrastructure, oil production dropped precipitously as a result of the fighting on the ground (from 1.78 million barrels per day in December 2010 to 98,000 barrels in July 2011), so economic activity contracted sharply during the war, with GDP for 2011 falling by 62% from the 2010 level. Fortunately, immediately after the war, oil production surged back, increasing to 1.6 million barrels per day a year after Gaddafi's death. As a consequence GDP increased by 105% in 2012. The capture of oil facilities by militias in 2013 brought oil production to 268,000 barrels per day by November 2013, affecting 2013 and 2014 GDP, which fell by 13% and 24% respectively.[3]

In order to estimate the cost of the revolutionary process in economic terms, here we make use of the Synthetic Control Method.[4] The SCM was first introduced in a seminal paper by Abadie and Gardeazabal (2003), where the economic cost in terms of GDP loss as a result of terrorism in the Basque Country was assessed. Since then, the same methodology with slight variations has been extensively applied in numerous comparative case studies. These are used by researchers to assess the effect of treatments (events or policies) on aggregate units (e.g. regions or countries) by estimating the evolution of outcomes for a unit affected by a particular treatment, and compare it with the evolution of a control group comprising a combination of control units, rather than one single control unit [Gardeazabal (2012)]. For

[3]The data come from two different sources: US Energy Information Administration, "International Energy Statistics" for oil barrels production, https://www.eia.gov/beta/international/data/browser; and World Bank for GDP, https://data.worldbank.org/indicator/NY.GDP.PCAP.PP.KD.

[4]Matta *et al.* (2016) follow a similar approach where Tunisia and Egypt are the respective case studies.

instance, and among a rapidly increasing number of works, Horiuchi and Mayerson (2015) study the influence of the 2000 Palestinian Intifada upon Israel's economy. Pinotti (2015) estimates the economic costs of organized crime in Southern Italy. Cavallo *et al.* (2013) study the effect of catastrophic natural disasters on economic growth. Billmeier and Nannicini (2013) study the effect of economic liberalization on real GDP per capita in a worldwide sample of countries. Colonescu (2017) analyzes the effects of adopting the euro as a common currency in terms of GDP, inflation rate, and public debt. Abadie *et al.* (2015) estimate the economic impact of the 1990 German reunification on West Germany. Gardeazabal and Vega-Bayo (2017) assess the effect of the political and economic integration of Hong Kong. Grier and Maynard (2016) study the impact of Hugo Chavez's regime on the Venezuelan economy. Abadie *et al.* (2010) estimate the effect of the tobacco control program implemented in California in 1988 on per-capita cigarette sales. Bohn *et al.* (2014) assess the effect of Arizona's 2007 Legal Arizona Workers Act (LAWA) on the proportion of the state's population of Hispanic non-citizens.

After comparing the observed growth rates for per capita real GDP for Libya and the synthetic Libya since 2011 to 2014, -49.0% and 15.16% respectively, our main results can be stated as follows. First, the cumulative loss in terms of the growth rate in per capita real GDP as a consequence of the revolution over that whole period was equal to 64.15%. Second, this differential in the growth rates has had implications in terms of forgone output: *i*) the cumulative loss of per capita real GDP amounted to 56,548 dollars as a consequence of the revolution; and, *ii*) the value of the cumulative loss of aggregate real GDP was of 350.5 billion dollars.

The rest of the paper is organized as follows. Section 2 briefly describes the SCM. Section 3 discusses the alternative computational methods available to estimate the SCM. Section 4 describes the data and the variables used in our estimation exercise. Section 5 describes the results. Section 6 is devoted to a robustness check analysis. Finally, Section 7 concludes.

2 Synthetic Control Method

The SCM can be briefly described as follows.[5] Assume that a balanced panel data set consisting of $J+1$ *units* (countries, regions, cities,...) and $t = 1, 2, ..., T$ *periods* (years, months,...) is observed. At some point in time $t = T_0 + 1$, where $1 \leq T_0 < T$, one unit starts experiencing some type of uninterrupted *treatment* (event, shock, law,...) until $t = T$.[6] This allows one to split up the set of T periods between the *pre-treatment periods*, $t = 1, 2, .., T_0$, and the *post-treatment* periods, $t = T_0 + 1, T_0 + 2, ..., T$. Similarly, two types of units emerge: the

[5]This heavily draws on Abadie *et al.* (2010)-(2015).

[6]If the number of treated units were more than one, the method could be applied independently to each of the treated units or to an aggregate of all treated units.

treated unit, which without loss of generality can be referred to as $j = 1$, and those in the set of the potential control units, $j = 2, 3, ..., J + 1$, which make up the *donor pool*. The SCM is aimed at assessing the effect of the treatment on some post-intervention outcome(s), Y_{1t}, (e.g. GDP, consumption of tobacco, ethnic composition of the labor force,...). The problem is that the counterfactual is obviously unobserved, so that the performance of the treated unit for the post-treatment periods under the hypothesis of the absence of treatment is unknown. Therefore, comparison with the observed performance of the treated unit under treatment might seem unfeasible. More formally, the treatment effect for unit 1 (the treated unit) in period t is given by

$$\alpha_{1t} \equiv Y_{1t}^1 - Y_{1t}^0, \tag{1}$$

for $t = T_0 + 1, T_0 + 2, ..., T$, and where Y_{1t}^1 denotes the (observed) potential outcome under treatment, and Y_{1t}^0 denotes the (unobserved) potential outcome under the hypothesis of no treatment. The way to circumvent this difficulty was first suggested in Abadie and Gardeazabal (2003), which consisted of building up a *synthetic* treated unit: a weighted average conveniently obtained among the untreated units in such a manner that it mimics as closely as possible the performance of the treated unit for a vector of pre-treatment characteristics during the pre-treatment periods. The payoff is an estimate of Y_{1t}^0 which allows one to obtain an estimate for α_{1t}.

More formally, optimal weights are found by solving the following problem

$$\begin{array}{ll} \min_{\{w_j\}_{j=2}^{J+1}} & (\mathbf{X}_1 - \mathbf{X}_1^s)' V (\mathbf{X}_1 - \mathbf{X}_1^s) \\ \text{s. t.} & \left\{ \begin{array}{l} \mathbf{X}_1^s = \sum_{j=2}^{J+1} \mathbf{X}_j w_j, \\ w_j \geq 0, \text{ for } j = 2, 3, .., J + 1, \\ \sum_{j=2}^{J+1} w_j = 1, \end{array} \right. \end{array} \tag{2}$$

where $\mathbf{X}_1 - \mathbf{X}_1^s$ is the difference between the P-dimension pre-treatment characteristic vector of the treated unit, $\mathbf{X}_1$ (the *predictors* of the outcome variable) and those of the synthetic control, $\mathbf{X}_1^s \equiv \sum_{j=2}^{J+1} \mathbf{X}_j w_j$, w_j trivially denoting the weight assigned to control unit j. V is a diagonal, positive semidefinite matrix whose m-th element represents a weight that reflects the relative importance that the researcher assigns to the m-th variable in vector $\mathbf{X}$ as a predictor of the outcome variable, Y.[7]

Once optimal weights, w_j^*, have been obtained, the effect of the intervention on the treated unit for period t is estimated as

$$\hat{\alpha}_{1t} \equiv Y_{1t}^1 - \sum_{j=2}^{J+1} w_j^* Y_{jt}^0, \tag{3}$$

[7] Note that the non-negative elements of matrix V can always be normalized to add up to, say, 1 as this does not affect the optimal choice for the vector of weights $\mathbf{w}$.

for $t = T_0 + 1, T_0 + 2, ..., T$, and where (by construction and if the donor pool has been correctly specified) no Y_{jt}^0 is affected by the treatment or intervention experienced by unit 1.

A measure of the goodness of fit of the synthetic unit to the observed treated unit, and the one that we follow in this paper as a criterion to rank the alternative computational methods implemented, is the pre-treatment mean squared prediction error (Pre-MSPE), which is defined as

$$Pre\text{-}MSPE \equiv \sum_{t=1}^{T_0} \left(Y_{1t} - \sum_{j=2}^{J+1} Y_{jt} w_j^* \right)^2 / T_0. \tag{4}$$

Conversely, the post-treatment mean squared prediction error (Post-MSPE), defined as

$$Post\text{-}MSPE \equiv \sum_{t=T_0+1}^{T} \left(Y_{jt} - \sum_{j=2}^{J+1} Y_{jt} w_j^* \right)^2 / (T - T_0), \tag{5}$$

provides us with an approximate measure of the effect (in absolute terms) of the treatment effect, so that the ratio of the latter to the former is interpreted as a natural assessment of the quantitative effect of the treatment, and whose relevance will become apparent in the discussion in Sections 5 and 6.

The natural question that arises at this point is whether the difference in Eq. (**??**) is really driven by the treatment at issue. To this end Abadie *et al.* (2010) carry out what they call 'in-space placebo studies', extending the concept introduced in Abadie and Gardeazabal (2003). Assume that the effect were estimated for *every single unit* $j = 2, 3, ..., J + 1$ in the donor pool, thereby obtaining a distribution of effect estimates. If the estimated effect $\hat{\alpha}_{1t}$ is truly due to the treatment under study, one would expect that the probability of finding an effect of larger magnitude among those obtained in the placebo analysis would be low because, by construction, none of the units in the donor pool was treated. Similarly, one could conduct 'in-time placebo studies': focusing on the treated unit, but this time assuming different treatment periods, one should obtain a low probability of finding a treatment effect of greater magnitude than that estimated for the true treatment period [see Abadie *et al.* (2015)].

3 Alternative computing methods

The optimal weights obtained upon solving the problem in Eq. (**??**), $\mathbf{w}^*(V) = \{w_j^*(V)\}_{j=2}^{J+1}$, will depend of course on the choice for matrix V. Several alternatives have been considered in the literature.[8] The first one would represent a subjective choice by letting the researcher

[8] See Firpo and Possebom (2016), Ferman and Pinto (2017), Bohn *et al.* (2014).

make a reasonable choice reflecting his/her previous knowledge about the relative importance of each particular predictor among those in $\mathbf{X}$ for the outcome $\mathbf{Y}$. The second one amounts to a nested, double minimization, where matrix V and control weights $\mathbf{w}^*$ are jointly obtained in such a way that $\mathbf{w}^*(V)$ solves the problem in Eq. (**??**) and V minimizes the square distance of the outcomes between the treated and the synthetic units,

$$\begin{array}{ll} \min_{\{V_p\}_{p=1}^{P}} & (\mathbf{Y}_1 - \mathbf{Y}_1^s)' (\mathbf{Y}_1 - \mathbf{Y}_1^s) \\ \text{s. t.} & \left\{ \begin{array}{l} \mathbf{Y}_1^s = \sum_{j=2}^{J+1} \mathbf{Y}_j w_j(V), \\ V = \left[\begin{array}{cccc} V_1 & 0 & ... & 0 \\ 0 & V_2 & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & V_P \end{array} \right] \\ V_p \geq 0, \text{ for } p = 1, 2, .., P, \\ \sum_{p=1}^{P} V_1 = 1, \end{array} \right. \end{array} . \tag{6}$$

[see Abadie and Gardeazabal (2003), Appendix B]. A third alternative, usually referred to as cross-validation, consists of dividing the pre-treatment years into a *training period* (from $t = 1$ to $t = T^*$) and a *validation period* (the remaining periods, i.e. from $t = T^* + 1$ to $t = T_0$). After using predictors measured in the training period, matrix V is selected, such that the resulting synthetic control minimizes the mean square prediction error over the validation period.[9] Then, the matrix V selected in the previous step and the predictor data measured in the validation period are used to obtain the optimal $\mathbf{w}$. This method has been followed, among others, in Abadie and Gardeazabal (2003), Abadie *et al.* (2015) and Pinotti (2015). As a final alternative, another data-driven method can be used to set matrix V: after regressing the outcome $\mathbf{Y}$ on the set of predictors $\mathbf{X}$, the elements of matrix V are obtained by comparing the corresponding OLS coefficients (in modulus or squared) over the sum of all coefficients (in modulus or squared, respectively). In the sequel we will refer to this alternative as the *regression-based* method.

The above discussion leads to a natural question: how robust optimal weights, w_j^*, are to alternative methods of computing predictor weights? Leaving aside the arguable option of making a subjective choice based on the researcher's previous knowledge about the relative importance of predictors, here we will compare the performance of the three other alternative computational methods pointed out above in our particular case study.

The first two methods are implemented with R using the **synth** package, the nested minimization being the default option [see https://cran.r-project.org/web/packages/Synth/index.html]. In both cases, two alternative available procedures can be used to minimize the

[9] This is defined as $(T_0 - T^*)^{-1} \sum_{t=T^*+1}^{T_0} \left(Y_{1t} - \sum_{j=2}^{J+1} Y_{jt} w_j^* \right)^2$. Intuitively, the cross-validation technique selects the matrix V that minimizes out-of-sample prediction errors.

Pre-MSPE: the **ipop** function [see https://www.rdocumentation.org/packages/kernlab/versions/0.9-25/topics/ipop] and the **LowRankQP** function [see https://www.rdocumentation.org/packages/LowRankQP/versions/1.0.2]. The the **synth** package is also implemented for Stata© and Matlab©. At the time of writing, the three of them can be downloaded from Jens Hainmueller's web page at https://web.stanford.edu/~jhain/software.htm. Regarding the joint minimization method, the Stata© package **synth** claims to allow for the option **nested** [see http://fmwww.bc.edu/RePEc/bocode/s/synth.html]. At the time of writing, however, it results in an error message. As for the Matlab© package, it represents the least elaborate alternative: as is the case with all numerical optimization procedures, a guessed solution as the starting values for the minimization problem is required, but the user is expected to be provide them, which may (and if fact does) lead to completely different solutions for the weights. Package **synth** in R, in turn, tries two initial guess-work solutions for the elements of diagonal matrix V on its own (equally weighted and regression-based). And, after trying the two, that leading to the lowest Pre-MSPE is selected to compute the optimal weights, w_j's [see https://github.com/cran/Synth/blob/master/R/synth.R]. The third method is run with Stata© and its package **synth**.

4 Data and variables

We use annual country-level data: a balanced panel data consisting of 14 countries and 25 yearly observations for each. Details follow.

In a first step, the countries in the donor pool have been restricted so that they can be considered as members of a "club" with some common key features. Otherwise, the resulting synthetic treated country unit might be economically meaningless despite possibly resulting in a substantial improvement in the pre-treatment mean squared prediction error, our goodness of fit measure. Given the role played by oil in the Libyan economy, a natural criterion to select the donor pool is to consider only those countries in which oil represents a sizable fraction of their respective economies. Thus, according to *World Bank Open Data*, in a first stage we selected the list of countries for which oil rents represent at least 5% of GDP on average for the period 1990-2010: Republic of Congo, Angola, Iraq, Kuwait, Equatorial Guinea, Libya, Gabon, Saudi Arabia, Nigeria, Republic of Chad, Oman, Qatar, Iran, Brunei, Venezuela, United Arab Emirates, Algeria, Bahrain, Trinidad and Tobago, Ecuador, Papua New Guinea, Sudan, Norway, Cameroon, Suriname, Vietnam and Malaysia [see https://data.worldbank.org/indicator/NY.GDP.PETR.RT.ZS].

In a second stage, some of these countries have been dropped for diverse reasons. In an attempt to strengthen the homogeneity of the units under consideration, we have eliminated those within the sample which are neither African nor Islamic majority countries: Venezuela, Ecuador, Papua New Guinea, Norway, Suriname, Vietnam and Trinidad and Tobago. Otherwise, the treatment experienced by these two countries would influence the performance of the synthetic Libya, thereby precluding a sensible causality inference. And some other countries were removed as they were involved in armed conflicts. Otherwise, including them in the donor pool would prevent us from obtaining a reliable synthetic unit. Thus, Iraq, Algeria, Republic of Congo, Angola, Republic of Chad and Sudan were left out. Therefore, the result is a reduced, but homogeneous, donor pool with 13 members: Bahrain, Brunei, Cameroon, Equatorial Guinea, Gabon, Iran, Kuwait, Malaysia, Oman, Nigeria, Qatar, Saudi Arabia and United Arab Emirates (UAE in the sequel).

Regarding the sample period, this goes from 1990 to 2014, which is further split between the pre-treatment and the post-treatment periods, 1990-2010 and 2011-2014 respectively. Ideally, a long enough pre-treatment period would be desirable to the extent that it led to a better fit between the treated and the synthetic units. Extending the pre-treatment, however, poses two problems. First, the availability of long, reliable economic time series data is an issue when departing from OECD or developed economies. And, second, expanding the pre-treatment period beyond some given year might not be recommendable if the treated unit and the control units had experienced diverging or non-homogeneous trends in earlier periods. Thus, the initial purpose of setting 1980-2010 as the pre-treatment period has been abandoned as it results in a rather poor fit (as measured by the pre-treatment mean square prediction error) between Libya and its synthetic.

We set the (log of) per capita real GDP as our outcome variable. Regarding the countries in the donor pool, the series has been obtained from Penn World Table, Version 9.0, available at http://www.rug.nl/ggdc/productivity/pwt/ [See Feenstra *et al.* (2015)]. More precisely, we have used the expenditure-side real GDP at chained PPPs (in mil. 2011US$) which allows us to make sound international comparisons. Unfortunately, this database does not contain data for Libya, the treated country, and as a result we have had to construct it. The procedure has been as follows. First, observations for the period 1990-1999 have been obtained from Maddison Project, available at http://www.ggdc.net/maddison/maddison-project/data.htm [see Bolt and Van Zanden (2014)]. Second, observations for the period 1999-2014 have been obtained from the *World Bank Open Data* [see https://data.worldbank.org/indicator/NY.GDP.PCAP.PP.KD]. And, third, these last two series have been conveniently chained taking year 1999 as the overlapping period.

Concerning the predictors, our choice of covariates follows three basic principles: *i*) economic meaning, *ii*) availability for all units, and *iii*) predictive power, i.e. a potential covariate is accepted as such in so far as it increases the quality of the synthetic unit, the quality being measured in terms of reduced Pre-MSPE. First, we do not introduce the whole series of lagged values (*i.e.* pre-treatment values) of the outcome variable as a means to improve the quality of the synthetic unit as opposed to some other works in the literature [see Billmeier and Nannicini (2013), Bohn *et al.* (2014), Gobillon and Magnac (2016), Hinrichs (2012) and Gardeazabal and Vega-Bayo (2016) among others]. As pointed out by Kaul *et al.* (2017), using all outcome lags as separate predictors make all other predictors irrelevant, regardless of how important all other predictors are for accurately predicting post-treatment values of the outcome, potentially threatening the estimator's unbiasedness. Along such lines, we only introduce five lagged values of the (log of) per capita real GDP corresponding to 1990, 1995, 2000, 2005 and 2010. As a matter of fact, it turns out that, after including the growth factors discussed below, introducing all the lagged values results in a higher Pre-MSPE.

Second, we use the average of the aggregate investment to GDP ratio for four periods: 1990-1994, 1995-1999, 2000-2004 and 2005-2010 as predictors. The source of data for Libya is *International Monetary Fund*, *World Economic Outlook Database*, *October 2017 Edition* [see https://www.imf.org/external/pubs/ft/weo/2017/02/weodata/index.aspx]. The data source used for the control units is Penn World Table, Version 9.0 [see above]. Third, we also include the average share of imports of goods and services to GDP ratio for two periods: 2001-2005 and 2006-2010 as predictors. Data have been obtained from the *World Bank Open Data* [see https://data.worldbank.org/indicator/NE.IMP.GNFS.ZS]. And, fourth, we finally include life expectancy at birth [see Aghion *et al.* (2011)]. In particular we consider the average values for four periods: 1990-1994, 1995-1999, 2000-2004, and 2005-2010. Data have been obtained from the *World Bank Open Data* [see https://data.worldbank.org/indicator/SP.DYN.LE00.IN

In the process of selecting appropriate predictors, some potential candidates have been discarded, whether because (to the best of our knowledge) data are not available for all the countries (e.g. stock of capital per capita or total factor productivity) or because, as it turned out, the Pre-MSPE rose (e.g. share of exports of GDP, share of oil rents of GDP, human capital measured as average years of educational attainment per capita, averages of the growth rate of per capita real GDP for several periods, and average consumption to GDP ratio).

5 Results

The time path for (the log of) per capita real GDP for Libya and the population-weighted average for *all* the countries in the donor pool is shown in Figure 1. The respective growth rates (given by the *slopes* of the respective time paths) share some similarities in the pre-treatment years until 1997. After that year, patterns start diverging. Additionally, the *levels* (which represent our outcome variable) are quite different, so that a better alternative to build up the synthetic Libya must be found.

[Insert Figure 1 around here]

Following the SCM above outlined, therefore, a more accurate synthetic Libya is obtained. Our main results are presented in Tables 1, 2 and 3, and in Figures 2 and 3.

In Table 1 we report the values of Pre-MSPE, Post-MSPE and also the ratio of the latter to the former as a natural assessment of the quantitative effect of the treatment, whose relevance will become apparent in the discussion in Sections 5 and 6. Additionally, it also shows the composition of the synthetic Libya, i.e. the control countries and the corresponding weights for a total of five alternative computations.

We first tried the joint minimization of the Pre-MSPE with respect to both the elements of the diagonal matrix V and the weights w, using the **synth** package in R. The results are shown for the two alternative available functions used to minimize the Pre-MSPE discussed in Section 3. Thus, using the **ipop** function, we obtained that Pre-MSPE $=$ 0.005823; while using instead the **LowRankQP** function, we obtained a slightly higher value, Pre-MSPE $=$ 0.007232.

We next tried the cross validation option, once again using the **synth** package in R. Not surprisingly, given the reduced length of the pre-treatment period which should be further divided and seen as the sum of two components (training 1990-2000, plus validation 2001-2010 -our choice), the resulting Pre-MSPE was higher than the one obtained in our previous results. As is the case with the nested minimization, the value of the Pre-MSPE depends on the particular function implemented in R for the minimization. Using the **ipop** function, we found a Pre-MSPE value of 0.743554; whereas using the **LowRankQP** function instead, we obtained Pre-MSPE $=$ 0.865060.

Finally, we run the exercise with Stata© package **synth** which implements the regression-based method referred to above. In this case, we obtained a value of Pre-MSPE $=$ 0.007610.

Two remarks are in order. First, the lowest value for Pre-MSPE is obtained with R package, under the nested minimization option and the ipop function. And, given our criterion to select the best alternative among the available computing methods as that attaining the lowest Pre-MSPE as it implies a better goodness of fit, this is the way in which all our further results have been computed. This, in turn, implies that synthetic Libya is best obtained as a weighted average of Cameroon, Saudi Arabia and UAE. The corresponding weights, w_j^*'s, are 19.5%, 39.8% and 40.7%, respectively, where decimal positions have been rounded so that weights below 0.001 are set equal to zero.[10] The output file also reports the weights without rounding, and the reported values of the outcome variable are internally computed with the optimal weights calculated with double precision.[11]

Second, control countries and their optimal weights are not robust to alternative methods. Depending on the computational method used, not only the weights but also the control units which make the synthetic treated unit differ. For instance, in three of the five numerical experiments, the synthetic unit is made up of three control units, while the other two require four control countries; and in these cases, control units do not coincide.

[Insert Table 1 around here]

Table 2 shows the predictor and the outcome variables for *i*) Libya, *ii*) the synthetic Libya, and *iii*) their corresponding (population-weighted) means for the 13 control countries. As can be seen, as regards Libya and its synthetic, the fit for the five lagged values of per capita GDP as well as for the four averages of life expectancy at birth is quite remarkable, whereas the fit for the rest of the predictors (averages of investment and import shares of GDP) is not particularly noteworthy.[12] And, not surprisingly, the resulting matrix V assigns higher weights to the pre-intervention values than to the other covariates: the sum of the corresponding V_p's exceeds 96% [see Eq. (**??**)]. Considering the average among the control countries rather than the synthetic Libya, however, would imply a less precise fit, which is in line with the remark made above regarding Figure 1 [see Table 2, columns 4 and 8].

[Insert Table 2 around here]

[10]For the sake of space saving, countries assigned zero weights are not shown. The reader is referred to Section 4.

[11]Stata©, however, only shows rounded (higher than or equal to 0.001) weights and reports the values of the outcome variable for the synthetic unit using only those weights. A simple calculation made with the help of a worksheet is enough to reveal this.

[12]Abadie, Diamond, and Hainmueller (2015), p. 498, claim that "*if the number of preintervention periods in the data is large, matching on preintervention outcomes (...) helps control for unobserved factors and for the heterogeneity of the effect of the observed and unobserved factors on the outcome of interest*". As remarked in Abadie and Gardeazabal (2003), p. 128, however, including *only* the pre-treatment values to construct the synthetic unit "*could be less appropriate to construct counterfactual per capita GDP paths, since it does not take into account information about known determinants of economic growth (...)*"

Figure 2 shows the time paths for the log of per capita gross domestic product for both Libya and the synthetic Libya. The fit between Libya and its synthetic during the pre-treatment period, 1990-2010, looks quite acceptable, and that part of the figure is nothing but the graphical consequence of the numerical results shown above in Table 2. Starting in 2011, however, a substantial gap between the predicted trend for the synthetic Libya and that for the observed Libya appears as a result of the revolution and the civil war that it brought about. The growth rate of per capita real GDP dramatically fell during the whole treatment period here considered, 2011-2014, even becoming negative (except in 2012). Figure 3 shows the time path of the treatment effect, that is to say, the difference in Eq. (**??**) in terms of the gap in (log) per capita real GDP.

[Insert Figures 2 and 3 around here]

Figures 2 and 3 show the qualitative implications quite neatly. But the quantitative implications are shown in Table 3 in terms of both forgone growth and forgone output. The observed growth rate of the per capita real GDP between 2011 and 2014 was -49%, while that for the synthetic Libya in the same period was 15.2%. In other words, the cumulative loss in terms of the growth rate of per capita real GDP during the 2011-2014 period was 64.2% [see columns 2-4 for further detail]. This gap in the growth rates had consequences in terms of output: as a result of the political and economic instability caused by the revolution, the sum of per capita real GDP losses for the 2011-2014 period was 56,548 dollars (as measured in PPP 2011 international dollars). And, in aggregate terms, the sum of real GDPs for that period was 350.5 billion dollars lower (measured in the same units). See columns 5 and 6.

These figures make sense if one bears in mind that the oil industry is, in terms of GDP share, the main economic sector in Libya. In particular, in 2010 oil represented 54% of total GDP and 83% of government revenues. However, as a consequence of the war, oil production dramatically dropped (from 1.78 million barrels per day in December 2010 to only 98,000 barrels per day in July 2011), so that economic activity, as measured by GDP, sharply contracted by 62%. After the war, oil production surged back, increasing to 1.6 million barrels per day by October 2012, thereby leading to a 105% increment of GDP in 2012. The capture of oil facilities by militias in 2013 brought oil production down to 268,000 barrels per day by November 2013, negatively affecting the GDP in 2013 and 2014 with two successive drops of 13% and 24% respectively. Thus, we may conclude that by the year 2014 Libya had not yet recovered its pre-revolution economic situation, but rather continued to be stuck considerably below its pre-revolution economic position.

[Insert Table 3 around here]

6 Robustness tests

One might always wonder whether the results presented in the previous section are really caused by the treatment under consideration, or they are merely a statistical fluke. The usual way to clear the doubt consists of conducting some placebo analysis [see, for instance, Abadie *et al.* (2010) and Abadie *et al.* (2015)]. We first tried *in-space* placebo analysis. In this case, this means that Libya was moved to the donor pool and then we applied the SCM to every other country in the pool which, as explained above, was composed of nations not exposed to a revolution similar in nature to that in Libya. Therefore, the exercise allows one to obtain a distribution of the effects on the countries which did not experience the treatment: our results would be validated if the probability of finding other significant treatment effects were low enough.

The results are shown in Figure 4.*a*. Each gray line depicts the estimated gap in (log) per capita real GDP for each of the 13 runs of the test, while the thicker line denotes the estimated gap for Libya. The fit for the pre-treatment period for Libya is reasonably precise, although it is also apparent that our outcome variable during the pre-treatment years, 1990–2010, is not as well approximated for some of the countries in the donor pool as convex combinations of the rest of the countries in the pool. Thus, while the Pre-MSPE for Libya, equals 0.005823 [see Table 1], the median and the mean for the other countries are 0.011441 and 0.2045902 respectively. And this naturally reduces confidence in the results of post-treatment analysis and placebo tests.

Following the same procedure as in Abadie *et al.* (2010), we exclude in turn those control countries whose performance during the pre-treatment period is relatively worse than that of Libya in terms of Pre-MSPE. Thus, Figure 4.*b* shows the result of dropping those countries with Pre-MSPE twenty times or more higher than Libya, so that 5 countries are eliminated: this implies that there is a reduction *i*) in the number of countries with substantial deviations from a zero gap before 2011, and also *ii*) in the number of countries with an abnormal non-zero gap in the post-treatment years. Reducing the threshold level, of course, strengthens this intuitive perception. For instance, Figure 4.*c* and Figure 4.*d* show the results of excluding control units with Pre-MSPE five times or more and Pre-MSPE one and a half times or more higher than Libya's respectively, so that 1 additional country is dropped in each of the last two cases.

[Insert Figures 4 around here]

As a way to objectively assess the significance of the effect (gap) for Libya, we build up the whole distribution of the ratios of post- to pre-treatment mean squared prediction error, R. If the SCM has correctly identified the synthetic Libya and the treatment effect, one will expect that such a ratio will be significantly higher for Libya than for any of the countries in the donor pool. Figure 5 shows the histogram for the distribution of R for the 14 countries (including, therefore, Libya and the 13 control units) in our sample. The ratio R for Libya equals 82.5914 [see Table 1], which is not surpassed by any other country; and the resulting *p-value* [or probability of finding a ratio R higher than or equal to Libya's], $p(R)$, is 0.0714, where

$$p(R) \equiv \sum_{k=1}^{J+1} \mathbf{1}[R_k \geq R]/(J+1),$$

$\mathbf{1}(*)$ denoting the indicator function of event $*$. Therefore, based on this test, one can conclude that the SCM has correctly identified the treatment effect on the Libyan economy.

[Insert Figure 5 around here]

As a complement to the previous robustness check, we also run an *in-time* placebo analysis. If the SCM had been implemented and the results were sound, no significant effect should be detected under the hypothesis that the treatment had occurred in some previous period to which it effectively took place. Thus, for instance, Figure 6 shows the result for a hypothetical treatment in year 2005. Our result is clearly reinforced: according to this test, it seems that the SCM has correctly identified the consequences in terms of per capita real GDP of the Arab spring for Libya.

[Insert Figure 6 around here]

For completeness, we also run a final robustness check. Following Abadie *et al.* (2015), we re-estimate the baseline model three times to construct a synthetic Libya leaving out in each iteration one of the three countries that received a positive weight (namely, Cameroon, Saudi Arabia and UAE). As acknowledged by the authors, by dropping such control units the goodness of fit is sacrificed to some degree, but in turn this sensitivity analysis allows us to verify to what extent the results are driven by any particular control country. The outcome of the exercise is displayed in Figure 7.

[Insert Figure 7 around here]

The leave-one-out synthetic control that shows the smallest and the largest effects of the revolution are those that exclude Saudi Arabia and Cameroon [Table 4], thereby providing a sort of confidence interval. The result is that the effects differ from those in Table 3. And the consequence is immediate: the robustness of the results obtained under the baseline model is conditioned by the reduced number of control countries which are assigned positive weights.[13]

[Insert Table 4 around here]

7 Conclusions

In 2011, beginning in Tunisia, a revolutionary movement quickly spread through the Middle East and North Africa. These popular uprisings, known as the Arab Spring, demanded democratic reforms in countries that in general suffered from dictatorships. In this paper, making use of the SCM, we have sought to assess the economic cost of the revolutionary process and the subsequent civil war over the period 2011-2014 in the particular case of Libya. The SCM consists of building a *synthetic* treated unit (Libya in our case) as a weighted average obtained among the untreated units or controls in such a manner that it mimics as closely as possible the performance of the treated unit for a vector of pre-treatment characteristics during the pre-treatment periods. The outcome that we try to emulate is the (log of) per capita GDP, and the predictors that we use to achieve it consist of some lags in the outcome and some averages of the respective shares in GDP of gross investment, imports and the life expectancy at birth. The fit between Libya and the synthetic Libya during the pre-treatment period appears quite accurate. Starting from 2011, however, a substantial gap between the

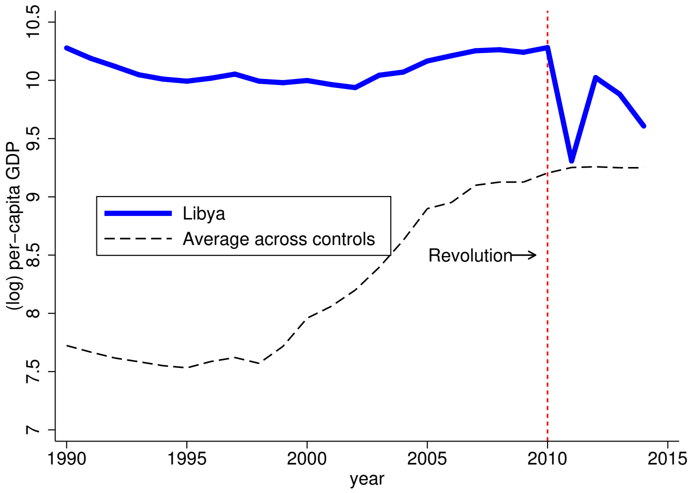
[13] Exclusion of individual control countries in a previous version of this paper (where the donor pool consisted of a set of 117 control units, 7 of which having positive weights) had a much reduced effect on the resulting synthetic Libya.

predicted trend for the synthetic Libya and Libya appears. A summary of the quantitative implications is as follows: the observed growth rate of the per capita real GDP between 2011 and 2014 was -49.0%, while that for the synthetic Libya in the same period was 15.2%. To put it differently, the loss in terms of the growth rate of per capita real GDP in that period was 64.2%. This gap in the growth rates had consequences in terms of output: in the absence of revolution, the cumulative per capita real GDP in the period 2011-2014 would have been 56,548 dollars higher and, in aggregate terms, the cumulative real GDP would have been 350.5 billion dollars higher (in PPP 2011 international dollars).

In order to know whether the results are really caused by the treatment under consideration some placebo analyses were conducted. We first tried the *in-space* placebo analyses, moving Libya to the donor pool and applying the SCM to every other country in the pool thereafter. Our results prove robust: the probability of finding a country with a higher effect than that found for Libya is only a 7.14%. We have also run an *in-time* placebo check under the hypothesis that the revolution had occurred in 2005, detecting no significant treatment effect. This further reinforces our conclusion that the SCM has correctly identified the consequences of the Arab spring for Libya in terms of real GDP. And, lastly, we have also conducted a leave-one-out analysis by re-estimating the base line model to construct alternative synthetic Libyas by omitting one of the controls with positive weight each time. This time, however, the low number of control countries with positive weights casts some doubts on the robustness of the results obtained under the benchmark model.

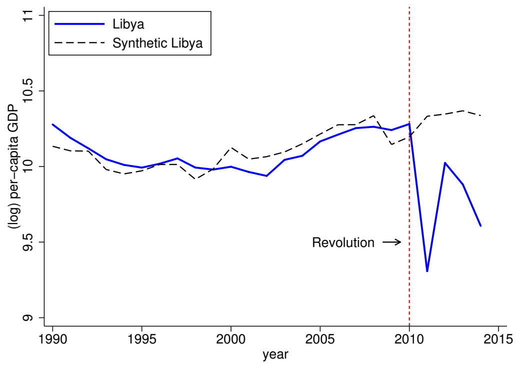
Figures

FIGURE 1
Trends in per capita real GDP: Libya vs. donors



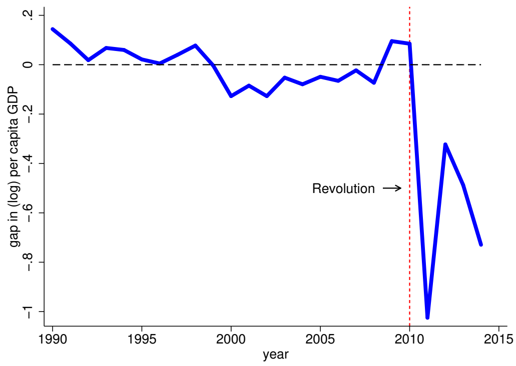
Key to Figure 1. Time paths for (the log of) per capita real GDP for Libya and the (population-weighted) average of the 13 countries in the donor pool.

FIGURE 2
Trends in per capita real GDP:
Libya vs. synthetic Libya



Key to Figure 2. Time paths of (the log of) per capita real GDP for Libya and the synthetic Libya for both the pre- and the post-treatment periods.

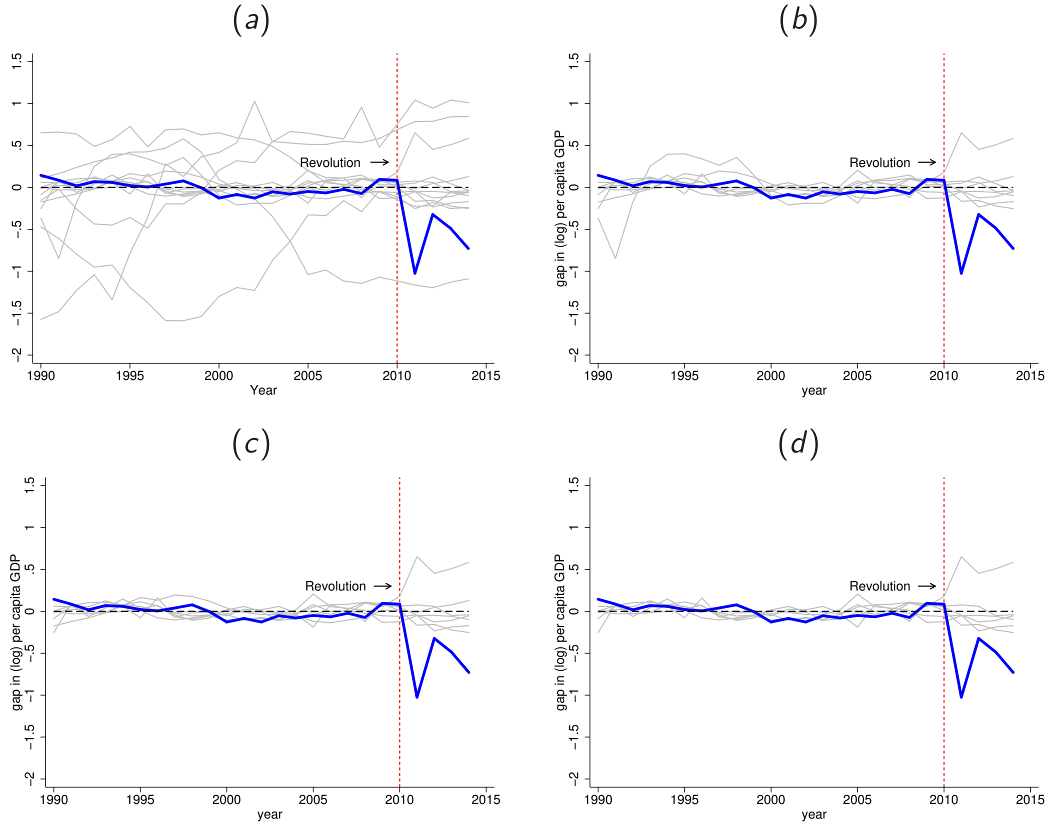
FIGURE 3
Trends in per capita real GDP gap:
Libya vs. synthetic Libya



Key to Figure 3. Time path for the gap in (the log of) per capita real GDP. See Eq. (**??**) for details.

FIGURE 4

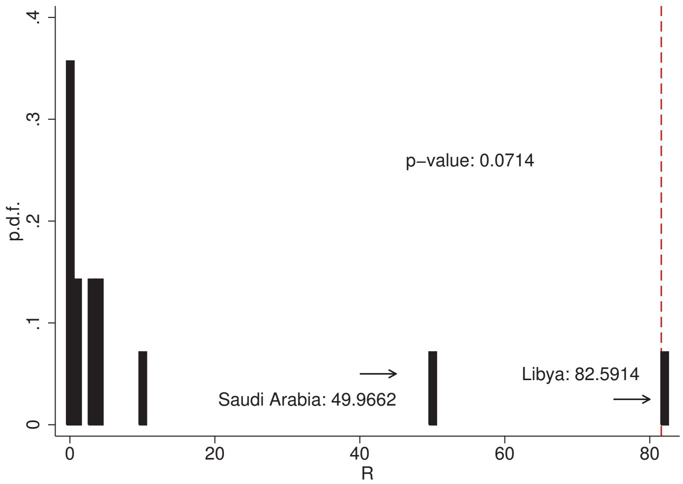
In-Space placebo test



Key to Figure 4. (log) per capita real GDP gap in Libya and placebo gaps. Panel a: all 13 control countries. Panel b: 8 control countries after dropping countries with Pre-MSPE 20 times higher or more than Libya. Panel c: 7 control countries after dropping countries with Pre-MSPE 5 times higher or more than Libya. Panel d: 6 control countries after dropping countries with Pre-MSPE 1.5 times higher or more than Libya.

FIGURE 5

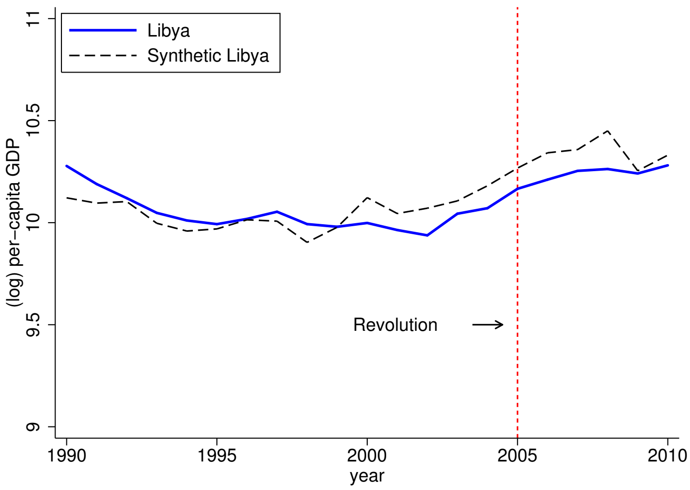
Distribution of R



Key to Figure 5. Ratio of post-revolution MSPE to pre-revolution MSPE: Libya and 13 control countries.

FIGURE 6

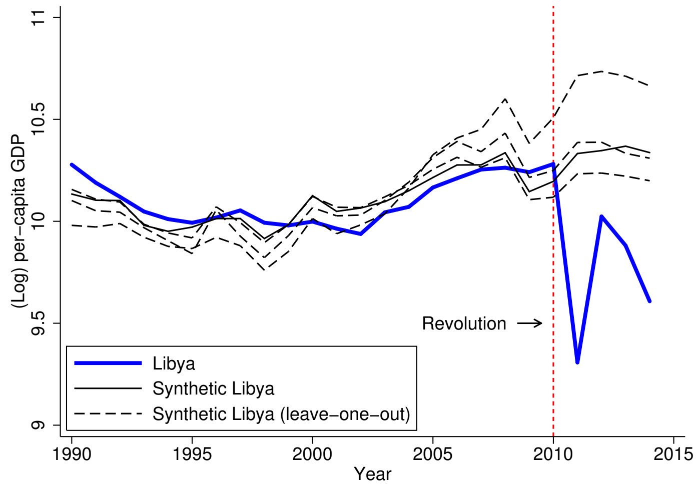
In-Time placebo test



Key to Figure 6. In-time placebo test under the counterfactual that the Libyan revolution had occurred in 2005.

FIGURE 7

Leave-One-Out distribution



Key to Figure 7. Time paths of (the log of) per capita real GDP for Libya, the synthetic Libya with all control countries with positive weights in Table 1, and the synthetic Libya excluding one such control countries at a time.

Tables

TABLE 1
Synthetic Libya: Alternative Computational Tools

(*a*) **Nested minimization: R (ipop)**

Country	*Weight*	*Country*	*Weight*	*Country*	*Weight*
Cameroon	0.195	Saudi Arabia	0.398	UAE	0.407

Pre-MSPE: 0.005823; *Post-MSPE:* 0.480949; *Ratio Pre- to Post-MSPE:* 82.5914

(*b*) **Nested minimization: R (LowRankQP)**

Country	*Weight*	*Country*	*Weight*	*Country*	*Weight*
Cameroon	0.153	Saudi Arabia	0.497	UAE	0.349

Pre-MSPE: 0.007232; *Post-MSPE:* 0.632938; *Ratio Pre- to Post-MSPE:* 87.5153

(*c*) **Cross Validation: R (ipop)**

Country	*Weight*	*Country*	*Weight*	*Country*	*Weight*	*Country*	*Weight*
Equatorial Guinea	0.325	Iran	0.196	Qatar	0.296	UAE	0.183

Pre-MSPE: 0.743554; *Post-MSPE:* 1.529246; *Ratio Pre- to Post-MSPE:* 1.434109

(*d*) **Cross Validation: R (LowRankQP)**

Country	*Weight*	*Country*	*Weight*	*Country*	*Weight*	*Country*	*Weight*
Equatorial Guinea	0.364	Malaysia	0.202	Qatar	0.349	UAE	0.085

Pre-MSPE: 0.865060; *Post-MSPE:* 1.725711; *Ratio Pre- to Post-MSPE:* 1.994904

(*e*) **Regression-based**: **Stata**

Country	*Weight*	*Country*	*Weight*	*Country*	*Weight*
Cameroon	0.157	Saudi Arabia	0.511	UAE	0.332

Pre-MSPE: 0.007610; *Post-MSPE:* 0.610567; *Ratio Pre- to Post-MSPE:* 80.2349

Key to Table 1. Countries in the donor pool with positive estimated weights for synthetic Libya for different computational methods. Panels (*a*) and (*b*) have been obtained under the joint minimization of Pre-MSPE with respect to weights (w_j's) and the elements of matrix V in Eqs. (**??**) and (**??**). Two alternative functions implemented in *R* have been used: ipop and LowRankQP. Panels (*c*) and (*d*) have been obtained by using the cross-validation method (for 1990-2000 and 2001-2010 as the training and the validation periods respectively) with *R*. Panel (*e*) shows the results obtained with Stata$^{©}$, the elements of V being computed with the regression-based method. Values for Pre-MSPE, Post-MSPE are also shown [see Eqs. (**??**) and (**??**)].

TABLE 2

Economic Growth Predictors Before the Revolution

	Libya	Synthetic Libya	Control Countries		Libya	Synthetic Libya	Control Countries
$\ln y_{1990}$	10.278	10.134	7.724	LEB$_i$	69.236	67.320	56.712
$\ln y_{1995}$	9.993	9.971	7.532	LEB$_{ii}$	70.215	68.249	57.602
$\ln y_{2000}$	9.999	10.126	7.958	LEB$_{iiia}$	70.905	69.504	58.693
$\ln y_{2005}$	10.166	10.215	8.899	LEB$_{iva}$	71.691	70.823	60.716
$\ln y_{2010}$	10.281	10.196	9.205	Imports$_{iiib}$	28.708	32.886	32.463
Invest.$_i$	20.545	30.205	17.962	Imports$_{ivb}$	34.345	45.612	31.056
Invest.$_{ii}$	8.741	26.455	17.770				
Invest.$_{iiia}$	36.225	24.422	16.728				
Invest.$_{iva}$	33.321	28.768	19.400				

Key to Table 2. $\ln y_t$: log of per capita real GDP for year t; Invest.$_k$: average investment to GDP ratio for period k; LEB$_k$: average life expectancy at birth for period k; Imports$_k$: average imports to GDP ratio for period k. Periods are denoted as follows i: 1990-1994; ii: 1995-1999; $iiia$: 2000-2004; $iiib$: 2001-2005; iva: 2005-2010; ivb: 2006-2010. All shares are expressed in per cent terms. Averages across control countries are population weighted.

TABLE 3
Growth and GDP effects

Period	g	g^s	Δg	$\Delta pcGDP$	ΔGDP (billions)
2011	-62.214	14.641	-76.856	$-19,693$	-122.0
2012	104.657	1.430	103.228	$-8,595$	-53.3
2013	-13.312	2.201	-15.513	$-12,284$	-76.1
2014	-23.915	-3.099	-20.816	$-15,975$	-99.1
2011-14	-48.995	15.157	-64.152	$-56,548$	-350.5

Key to Table 3. g: annual growth rate (%) of per capita real GDP for Libya; g^s: annual growth rate (%) of per capita real GDP for synthetic Libya; $\Delta g \equiv g - g^s$: difference in growth rates (%) between Libya and synthetic Libya; $\Delta pcGDP$: loss in per capita real GDP; ΔGDP: loss in real GDP. Both $\Delta pcGDP$ and ΔGDP are measured in PPP 2011 international dollars.

TABLE 4
Leave-one-out test

		Leaving out Cameroon *(maximum effect)*				Leaving out Saudi Arabia *(minimum effect)*			
Period	g	g^s	Δg	$\Delta pcGDP$	ΔGDP (billions)	g^s	Δg	$\Delta pcGDP$	ΔGDP (billions)
2011	-62.2	23.0	-85.2	$-33,991$	-210.5	12.2	-74.4	$-16,781$	-103.9
2012	104.7	2.1	102.6	$-23,382$	-144.9	0.4	104.3	$-5,348$	-33.2
2013	-13.3	-2.3	-11.0	$-25,330$	-156.9	-1.5	-11.8	$-7,932$	-49.1
2014	-23.9	-4.7	-19.3	$-27,916$	-173.2	-2.2	-21.7	$-11,998$	-74.4
2011-14	-49.0	17.0	-66.0	$-110,619$	-685.6	8.4	-57.4	$-42,059$	-260.7

Key to Table 4. See Key to Table 3.

References

Abadie, A. and J. Gardeazabal. 2003. "The Economic Costs of Conflict: A Case Study of the Basque Country." *The American Economic Review*, 93 (1): 113-132. https://doi.org/10.1257/000282803321455188

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(490): 493-505. https://doi.org/10.1198/jasa.2009.ap08746

Abadie, A., A. Diamond, and J. Hainmueller. 2015. "Comparative Politics and the Synthetic Control Method." *American Journal of Political Science*, 59(2): 495–510. https://doi.org/10.1111/ajps.12116

Aghion, P., P. Howitt and F. Murtin. 2011. "The relationship between health and growth: When Lucas meets Nelson-Phelps." *Review of Economics and Institutions*, 2(1): 1-241. https://doi.org/10.5202/rei.v2i1.1

Billmeier, A. and T. Nannicini. 2013. "Assessing Economic Liberalization Episodes: A Synthetic Control Approach." *The Review of Economics and Statistics*, 95(3): 983–1001. http://www.mitpressjournals.org/doi/suppl/10.1162/REST_a_00324.

Bohn, S., M. Lofstrom, and S. Raphael. 2014. "Did the 2007 Legal Arizona Workers Act Reduce the State's Unauthorized Immigrant Population?." *The Review of Economics and Statistics*, 96(2): 258–269. http://www.mitpressjournals.org/doi/suppl/10.1162/REST_a_00429.

Bolt, J. and J. L. van Zanden. 2014. "The Maddison Project: collaborative research on historical national accounts." The Economic History Review, 67(3): 627–651. http://dx.doi.org/10.1111/1468-0289.12032

Cavallo, E., S. Galiani, I. Noy, and J. Pantano. 2013. "Catastrophic Natural Disasters and Economic Growth." *The Review of Economics and Statistics*, 95(5): 1549–1561. https://doi.org/10.1162/REST_a_00413

Colonescu, C. 2017. "Macroeconomic Effects of the European Monetary Union: A Counterfactual Analysis." *Athens Journal of Business and Economics*, forthcoming.

Feenstra, Robert C., R. Inklaar and M. P. Timmer. 2015. "The Next Generation of the Penn World Table." *American Economic Review*, 105(10): 3150-3182. https://doi.org/10.1257/aer.20130954

Ferman, B and C. Pinto. 2017. "Revisiting the Synthetic Control Estimator." *mimeo* available at https://sites.google.com/site/brunoferman/research. [Last accessed April 2017.]

Firpo, S. and V. Possebom. 2016. "Synthetic Control Estimator: A Generalized Inference Procedure and Confidence Tests." *mimeo* available at https://www.aeaweb.org/conference/2017/preliminary/paper/Tys6f3QR. [Last accessed April 2017.]

Gardeazabal, J. 2012. "Methods for Measuring Aggregate Costs of Conflict." *The Oxford Handbook of the Economics of Peace and Conflict*, Ed. by Michelle R. Garfinkel and Stergios Skaperdas, Oxford University Press, U.S.A., Ch. 11.

Gardeazabal, J. and A. Vega-Bayo. 2017. "An Empirical Comparison Between the Synthetic Control Method and Hsiao *et al.*'s Panel Data Approach to Program Evaluation." *Journal of Applied Econometrics*, 32(5): 983–1002. https://doi.org/10.1002/jae.2557

Gobillon, L. and T. Magnac. 2016. "Regional Policy Evaluation: Interactive Fixed Effects and Synthetic Controls." *Review of Economics and Statistics*, 98(3): 535–551. https://doi.org/10.1162/REST_a_00537

Groizard, J. L., M. Ismael and M. Santana. 2016. "The economic consequences of political upheavals: the case of the Arab Spring and international tourism." *mimeo* available at http://preview.tinyurl.com/k5fnzka.

Grier, K. and N. Maynard. 2016. "The economic consequences of Hugo Chavez: A synthetic control analysis." *Journal of Economic Behavior & Organization* 125: 1-21. https://doi.org/10.1016/j.jebo.2015.12.011

Hinrichs, P. 2012. "The Effects of Affirmative Action Bans on College Enrollment, Educational Attainment, and the Demographic Composition of Universities." *Review of Economics and Statistics*, 94(3): 712–722. https://doi.org/10.1162/REST_a_00170

Hodler, R. 2012. "The Political Economics of the Arab Spring." *CESifo Working Paper* No. 4023, Category 2: Public Choice.

Horiuchi, Y. and A. Mayerson. 2015. "The Opportunity Cost of Conflict: Statistically Comparing Israel and Synthetic Israel." *Political Science Research and Methods*, 3(3): 609-618. https://doi.org/10.1017/psrm.2014.47

Kaul, A., S. Klößner, G. Pfeifer and M. Schieler. 2017. "Synthetic Control Methods: Never Use All Pre-Intervention Outcomes Together With Covariates". Working Paper. http://www.oekonometrie.uni-saarland.de/papers/SCM_Predictors.pdf. [Last accessed February 2018.]

Khan, M. 2014. "The Economic Consequences of the Arab Spring." Atlantic Council, Rafik Hariri Center for the Middle East, ISSUEBRIEF.

Matta, S., S. Appleton and M. Bleaney. 2016. "The Impact of the Arab Spring on the Tunisian Economy." *CREDIT Research Paper*, No. 15/09, University of Nottingham.

Pinotti, P. 2015. "The Economic Costs of Organised Crime: Evidence from Southern Italy." *The Economic Journal*, 125: F203-F232. https://doi.org/10.1111/ecoj.12235

The Economist. 2013. "*The Arab Spring: A Climate of Change.*" Special Report, The

Arab Spring.