Erasmus University Rotterdam
Erasmus School of Economics
Bachelor Thesis Econometrics and Operations Research, Economics and Business Economics

# The true costs of Brexit: an application of the synthetic control method

Prya Stikvoort (495139)

| | |
|---|---|
| Supervisor: | Prof. dr. Philip Hans Franses |
| Second assessor: | Dr. Phyllis Wan |
| Date final version: | 1st July 2024 |

**Abstract**

This paper aims to assess the true costs of Brexit by estimating its effect on UK GDP per capita using the synthetic control method as proposed in Abadie (2021). This method is employed in various settings to provide robustness of the results, namely by considering two different main donor pools, backdating the intervention to obtain insights about the timing of Brexit, and by performing the estimation for 500 randomly selected donor pools. Furthermore, permutation methods are used for inference on the findings. By performing the synthetic control method in different settings, this paper provides a comprehensive analysis of two important aspects of this method, namely the donor pool and the time of intervention. The main findings are that Brexit had a negative effect on UK GDP per capita and that the effects of Brexit likely started to unfold already in 2014. Although the magnitude of the effect differs across the various settings, the synthetic UK GDP per capita always lies above the real UK GDP per capita, therefore indicating a significant economic impact attributed to Brexit across all analyses.

# 1 Introduction

On June 23, 2016, the United Kingdom (UK) voted to leave the European Union (EU). This momentous decision is commonly referred to as Brexit and can be considered one of the most significant events in Europe over the past decade. The arguments for Brexit centered mainly on immigration and national autonomy (Arnorsson & Zoega, 2018) and this event can be seen as the biggest reversal of international economic integration in contemporary times (Dhingra & Sampson, 2022), thus having substantial economic consequences.

Even though opposition towards the UK's membership of the EU has fluctuated over the years, it has remained substantial ever since the UK joined in 1973 (Dennison & Carl, 2016). Already in 1975, the first national referendum was held to decide whether the UK should remain in the EU , then known as the European Community, with the majority voting to stay (Saunders, 2018). However, around 30 to 60 percent of the British population have always criticized EU membership. This scepticism towards the EU further increased due to a rapid rise in EU immigration, which began in the late 1990s, and the Eurozone debt crises (Dennison & Carl, 2016).

In 1993, the UK Independence Party (UKIP) was founded, a single-issue party promoting the withdrawal of the UK from the EU. Under Nigel Farage's leadership, UKIP adopted a strategic populist shift by rebranding its Euroscepticism within a broader populist framework, while simultaneously making immigration the party's top priority (Tournier-Sol, 2021). UKIP had a major influence on British politics and its growing electoral success played an important role in former prime minister and conservative David Cameron's pledge for a Brexit referendum which he proposed, and promised, during his 2013 Bloomberg speech (Tournier-Sol, 2021; Varga, 2023; Whiteley, Goodwin & Clarke, 2018). In the 2015 elections, which were won by the Conservatives, Cameron, who himself was pro-Remain, promised to call a referendum on Brexit (Varga, 2023). The referendum was eventually held in 2016 and lost by Cameron, leading him to resign.

Thus, the idea of Brexit started decades before the final decisive vote in 2016. Over the years it has occupied the British population, as well as other EU countries and the rest of the world. The dissatisfaction of the British on the UK's EU membership raises the question of whether the UK is in fact better off after Brexit. This paper explores that question by studying the effect of Brexit on UK GDP per capita using the synthetic control method (SCM) as described in Abadie (2021).

Originally proposed in Abadie and Gardeazabal (2003) and Abadie, Diamond and Hainmueller (2010), SCMs have experienced increasing popularity in empirical research and have been widely applied in economics and other disciplines. SCMs are a tool for comparative case studies and aim to estimate the effects of interventions implemented at an aggregate level affecting a small number of large units, on a certain aggregate outcome of interest (Abadie, 2021).

Comparative case studies rely on the concept that the effect of an intervention can be determined by comparing the changes in outcome variables between the treated unit and a group of similar, untreated units, which serve as the comparison units. To achieve this, the outcomes of the variable of interest of the intervention and comparison units must be driven by common factors that cause strong co-movement. SCMs aim to find a combination of the comparison units, which are to be selected from a donor pool, such that the trajectory of the outcome variable of this combination optimally resembles that of the unit under investigation (Abadie, 2021).

In this study, the intervention is Brexit, the treated unit is the UK, and the outcome of interest is GDP per capita. The aim is to produce a 'synthetic UK' which is made from a selection of similar countries chosen from the donor pool. As the comparison units are unaffected by the intervention, the synthetic control allows us to estimate the trajectory of the outcome variable in the absence of the intervention. Consequently, the effect of the intervention on the treatment unit can be evaluated by comparing the two trajectories, i.e., the actual trajectory and the synthetic trajectory, graphically. Moreover, its statistical significance can be assessed by using permutation methods (Abadie, 2021). Note that both of these options are carried out in this paper.

The units in the donor pool, i.e., the potential comparison units, must be similar to the exposed unit but unaffected by the intervention (Abadie, 2021). In this research, it is expected that (some) EU countries are good candidates based on the similarity criterion. Previous studies that employed SCMs in the context of Brexit, see for example Farid (2020), have also included EU countries in the donor pool. However, it is debatable to what extent EU countries are unaffected by Brexit. It is not unlikely that Brexit had spillover effects on other EU countries. Large spillover effects could lead to large bias if the comparison units affected by these effects are assigned high weights in the synthetic control. However, removing units that are initially ideal candidates for the donor pool because of their similarity to the treated unit, could substantially deteriorate the pre-intervention fit (Di Stefano & Mellace, 2020). Thus, a trade-off between bias and pre-intervention fit exists. This research therefore considers two main donor pools, one of which is a set of 31 OECD countries, while the other donor pool excludes EU countries from this set.

Furthermore, to provide robustness of the results, an extension to the main model is employed in which a synthetic control is estimated for 500 different donor pools. These donor pools each consist of 20 units randomly selected from the 31 available countries. To draw a credible conclusion regarding the effect of the intervention, one would hope the majority of the 500 different synthetic control estimators reveal the same effect. Moreover, recall that Brexit involves a complex timeline in which anticipation effects are likely present. This makes it difficult to assign a specific point in time to this event, as it did not happen all of a sudden. Therefore, the time of intervention is also backdated from 2016 until 2012, and the results of the associated synthetic control estimators are evaluated and compared against each other, which may also provide insights into when the effects of Brexit started to unfold.

This research contributes to the existing literature by focusing on the importance of the donor pool and examining the effect of using multiple donor pools. Moreover, the timing effect is investigated by backdating the intervention. Furthermore, the application in this paper provides further insights into how Brexit affected the UK economy. This knowledge may be useful for economic and political agents and it may influence people's opinion on the desire of their country to leave the EU.

The outline of this paper is as follows. Section 2 describes the used data, Section 3 presents the underlying methodology of the models, and Section 4 provides the results, followed by the conclusion in the last section.

# 2 Data

I use annual data for the period 1995-2022. All data is obtained from the World Bank[1].

## 2.1 Donor pool

In this research, two main donor pools are considered. The first donor pool, referred to as the complete donor pool, consists of a sample of 31 OECD countries because this ensures similar levels of income and economic development between the treated unit and the comparison units. The countries in the complete donor pool are listed in Table 2, along with the averages of the predictor variables which are discussed in more detail in the next subsection. The complete donor pool includes EU countries as well and they are expected to be ideal comparison units because of their similarity in terms of trade relationships, economic policies, and geographic proximity. However, as discussed in the introduction, these countries are possibly affected by Brexit which may lead to biased results.

When including potentially affected units in the donor pool, it is important to be aware of the direction of the potential bias of the resulting estimator (Abadie, 2021). This depends on whether we suspect the intervention to affect some comparison units either negatively or positively, which would respectively result in underestimation or overestimation of the synthetic trajectory. In the case of Brexit, one could expect this event to negatively affect GDP per capita of some EU countries due to, for example, its effect on trade. This can potentially result in an underestimation of the trajectory of GDP per capita in the absence of the intervention, which would indicate a smaller intervention effect. Therefore, a second donor pool is considered which is a subset of the first donor pool where EU countries are excluded. As this donor pool does not include units that are potentially affected, at least not to a significant extent, it is expected that

[1] `https://databank.worldbank.org/source/world-development-indicators`

the resulting synthetic trajectory of the outcome variable lies above the trajectory estimated by the complete donor pool, as it does not incorporate the negative effects of Brexit.

### 2.2 Predictors

Similar to the application in Abadie, Diamond and Hainmueller (2015), where the effect of the German reunification on per capita GDP in West Germany is estimated, the predictors considered in this paper are trade openness, inflation rate, industry rate, and investment rate. Note that also pre-intervention values of the outcome variable are included as a predictor. Unfortunately, due to data availability limitations, the variable schooling used in Abadie et al. (2015) is not included in the predictors. In addition to the predictors in Abadie et al. (2015), also the variables current account balance and capital account balance are used, as the study by Divya and Devi (2014) reveals these are factors that significantly predict GDP. Note that a description of each of the variables is provided in Table 1. Table 2 shows the averages of the predictor variables for each of the countries in the complete donor pool. From Table 2, we can observe that there is quite some variation in the variable of interest, GDP per capita. Countries that have a GDP per capita similar to that of the UK are mostly EU countries. Non-EU countries that are relatively close in terms of GDP per capita are Iceland and the United States. Furthermore, from the table, it seems like the variables trade openness, inflation rate, industry rate, and capital account balance show more variation across countries, while investment rate and current account balance are relatively close across countries.

**Table 1:** The used predictor variables and their descriptions

| Variable | Description |
|---|---|
| GDP per capita | Gross domestic product divided by midyear population in current U.S. dollars (USD) |
| Trade openness | The sum of exports and imports of goods and services measured as a share of GDP |
| Inflation rate | Inflation as measured by the CPI |
| Industry rate | Value added of industry (including construction) as a percentage of GDP |
| Investment rate | Gross capital formation as a percentage of GDP |
| Current account balance | The sum of net exports of goods and services, net primary income, and net secondary income, as a percentage of GDP |
| Capital account balance | The sum of acquisitions and disposals of non-produced nonfinancial assets, such as land sold to embassies and sales of leases and licenses, as well as capital transfers, including government debt forgiveness, as a percentage of GDP |

*Notes:* All the data on these predictors is retrieved from the World Bank. The time period that is used for these variables in the optimisation is from 2007 until 2015. Recall that the initial time of intervention is 2016.

**Table 2:** The countries in the complete donor pool and their predictor means over the initial pre-intervention period from 1995 until 2015

**(a)**

| Country | GDP pc | Trade | Inflation | Industry | Investment | Current a/c. | Capital a/c. |
|---|---|---|---|---|---|---|---|
| Australia | 38.1 | 40.9 | -3.8e-04 | -4.6 | 26.4 | 25.2 | 2.7 |
| Austria* | 38.4 | 90.8 | -1.3e-03 | 2.6 | 24.6 | 27.1 | 1.8 |
| Belgium* | 35.9 | 142.1 | -1.2e-03 | 1.4 | 23.0 | 22.7 | 1.9 |
| Canada | 35.7 | 69.4 | -4.8e-05 | -0.5 | 22.2 | 27.6 | 1.9 |
| Chile | 8.9 | 64.8 | 2.6e-03 | -1.2 | 25.4 | 31.3 | 3.9 |
| Costa Rica | 6.0 | 79.2 | 8.1e-04 | -4.3 | 20.2 | 24.0 | 10.0 |
| Denmark* | 47.1 | 88.4 | 1.2e-04 | 3.6 | 21.3 | 21.7 | 2.0 |
| Finland* | 37.8 | 73.2 | 9.0e-04 | 3.3 | 22.8 | 28.1 | 1.5 |
| France* | 33.7 | 53.9 | 4.6e-04 | 0.4 | 22.0 | 19.7 | 1.5 |
| Germany* | 35.6 | 68.8 | -1.7e-04 | 3.4 | 21.5 | 27.2 | 1.5 |
| Greece* | 20.0 | 52.4 | 1.3e-02 | -6.6 | 21.2 | 17.5 | 3.2 |
| Iceland | 43.0 | 80.5 | -5.3e-04 | -6.2 | 21.8 | 21.9 | 4.5 |
| Ireland* | 42.3 | 166.8 | -4.7e-03 | -1.2 | 23.5 | 29.1 | 2.2 |
| Italy* | 29.7 | 50.1 | 1.2e-03 | -0.1 | 20.0 | 23.4 | 2.2 |
| Korea, Rep. | 18.6 | 74.9 | 4.4e-05 | 2.3 | 32.5 | 34.2 | 3.1 |
| Latvia* | 8.5 | 98.1 | 1.4e-02 | -6.2 | 27.0 | 22.3 | 5.9 |
| United Kingdom | 36.9 | 54.7 | -2.0e-04 | -2.5 | 17.6 | 20.6 | 2.1 |

**(b)**

| Country | GDP pc | Trade | Inflation | Industry | Investment | Current a/c. | Capital a/c. |
|---|---|---|---|---|---|---|---|
| Lithuania* | 8.6 | 113.6 | 1.5e-02 | -6.0 | 22.0 | 27.6 | 5.6 |
| Luxembourg* | 82.8 | 273.8 | -2.1e-03 | 6.6 | 20.2 | 14.5 | 2.0 |
| Mexico | 8.5 | 53.8 | -2.5e-04 | -1.4 | 22.2 | 33.3 | 9.4 |
| Netherlands* | 40.7 | 126.5 | -2.6e-03 | 5.8 | 21.2 | 21.0 | 2.0 |
| New Zealand* | 26.7 | 58.9 | 7.1e-03 | -3.9 | 22.7 | 22.9 | 2.2 |
| Norway | 65.7 | 70.0 | -4.5e-04 | 10.8 | 24.3 | 34.7 | 2.0 |
| Poland* | 8.6 | 71.1 | 8.7e-03 | -3.9 | 21.7 | 29.6 | 6.1 |
| Portugal* | 17.6 | 67.6 | 1.6e-02 | -6.5 | 22.7 | 22.0 | 2.3 |
| Slovak Republic* | 11.7 | 141.3 | 7.6e-03 | -4.0 | 27.1 | 30.2 | 4.9 |
| Slovenia* | 17.8 | 118 | -4.7e-04 | -0.5 | 25.6 | 29.3 | 4.9 |
| Spain* | 23.9 | 55.5 | 6.7e-03 | -3.2 | 24.3 | 25.6 | 2.5 |
| Sweden* | 43.5 | 80.4 | -1.5e-03 | 5.1 | 22.5 | 24.7 | 1.1 |
| Switzerland | 61.3 | 104.0 | -1.0e-02 | 8.5 | 26.8 | 25.9 | 0.6 |
| Turkiye | 7.3 | 48.2 | -5.8e-05 | -3.2 | 25.7 | 27.1 | 32.4 |
| United States | 42.9 | 26.0 | -2.3e-04 | -3.4 | 21.6 | 20.7 | 2.3 |
| United Kingdom | 36.9 | 54.7 | -2.0e-04 | -2.5 | 17.6 | 20.6 | 2.1 |

*Notes:* All data is retrieved from the World Bank. EU countries are marked with an asterisk (*). GDP per capita is measured in current U.S. dollars (thousands), while trade openness, inflation rate, industry rate, investment rate, current account balance, and capital account balance are expressed as percentages of GDP. See Table 1 for a detailed description of each variable. Note that the UK, which is displayed in the last row, is not part of the donor pool, it is included in the table for comparative purposes.

# 3 Methodology

This study implements the same methodology as in Abadie (2021) and is presented below.

## 3.1 Notation

The data is obtained for $J + 1 = 32$ units where it is assumed that unit $j = 1$ is the treated unit (in this case the UK) and the rest of the units are forming the first donor pool consisting of $J = 31$ units. The second donor pool consists of 12 units after removing the EU countries. The time span is $T = 28$ periods.

The time of intervention $T_0$ in this study is intially the year 2016. Note that Brexit was announced in June 2016 but the UK formally withdrew from the EU only on 31 January 2020. However, 2020 is not chosen to be the time of intervention because of potential anticipation effects. If forward-looking economic agents react ahead of the intervention, or if some aspects of the intervention are implemented prior to its official enactment, synthetic control estimators may be biased (Abadie, 2021). In this case, Abadie (2021) advises backdating the intervention in the data set to a period before any potential anticipation effects such that the full extent of the effect of the intervention can be estimated.

As mentioned in the introduction, even before 2016, many events took place that drew attention to Brexit. It is therefore likely that anticipation effects are present even before to the vote in 2016. For that purpose, I will also perform the analysis for different $T_0$, namely by backdating from 2016 until 2012. Note that in the absence of anticipation effects, the results of a credible synthetic control should not be affected by backdating the intervention (Abadie, 2021).

The outcome of interest (GDP per capita) for unit $j$ at time $t$ is represented by $Y_{jt}$ and its set of $k = 7$ predictors is given by $X_{1j}, ..., X_{1k}$. Moreover, the $k \times 1$ vector $\mathbf{X}_1$ contains the values of the predictors for unit $j = 1$ (i.e., the treated unit) and the $k \times J$ matrix $\mathbf{X}_0 = [\mathbf{X}_2 ... \mathbf{X}_{j+1}]$ contains the values of the predictors for the $J$ untreated units. Furthermore, $Y_{jt}^N$ is defined as the potential response of unit $j$ at time $t$ without intervention. For unit $j = 1$ (the treated unit) and a post-intervention period $t > T_0$, $Y_{1t}^I$ is defined as the potential response under the intervention. The effect of the intervention for the treated unit in period $t > T_0$ then is:

$$\tau_{1t} = Y_{1t}^I - Y_{1t}^N. \tag{1}$$

Note that $Y_{1t}^I$ is observed because unit $j = 1$ underwent the intervention. Crucial in SCMs is estimating $Y_{1t}^N$, i.e., the outcome of interest in the absence of the intervention.

## 3.2 Estimation

A synthetic control is a weighted average of the units in the donor pool. Given a set of weights $\mathbf{W} = (w_2, ..., w_{J+1})'$, the synthetic control estimators of $Y_{1t}^N$ and $\tau_{1t}$ are, respectively (Abadie, 2021):

$$\hat{Y}_{1t}^N = \sum_{j=2}^{J+1} w_j Y_{jt}, \tag{2}$$

and

$$\hat{\tau}_{1t} = Y_{1t} - \hat{Y}_{1t}^{N}. \tag{3}$$

Note that the weights are restricted to be nonnegative and to sum to one to avoid extrapolation. Given a set of nonnegative constants $v_1, ..., v_k$ which reflect the relative importance of each predictor variable when measuring the discrepancy between $\mathbf{X}_1$ and $\mathbf{X}_0\mathbf{W}$, Abadie and Gardeazabal (2003) and Abadie et al. (2010) propose to choose $\mathbf{W}^* = (w_2^*, ..., w_{J+1}^*)'$ that minimises

$$||\mathbf{X}_1 - \mathbf{X}_0\mathbf{W}|| = \left(\sum_{h=1}^{k} v_h \left(X_{h1} - w_2 X_{h2} - ... - w_{J+1} X_{hJ+1}\right)^2\right)^{\frac{1}{2}}. \tag{4}$$

This results in a synthetic control that best mimics the pre-intervention values for the treated unit of predictors of the outcome variable. For given $v_1, ..., v_k$, minimising the above function can be achieved using constrained optimisation. However, $\mathbf{V}$ still needs to be chosen and there are multiple ways to do this, see Abadie (2021), Abadie and Gardeazabal (2003) and Abadie et al. (2010), and Abadie et al. (2015).

This paper follows the approach of Abadie and Gardeazabal (2003) and Abadie et al. (2010) who choose $\mathbf{V}$ such that the synthetic control $\mathbf{W}(\mathbf{V})$ minimises the mean squared prediction error (MSPE) of this synthetic control with respect to $Y_{1t}^N$. The MSPE here is defined as:

$$\sum_{t \in \mathcal{T}_0} \left(Y_{1t} - w_2(\mathbf{V}) Y_{2t} - ... - w_{J+1}(\mathbf{V}) Y_{J+1t}\right)^2, \tag{5}$$

for some set $\mathcal{T}_0 \subseteq \{1, 2, ..., T_0\}$ of pre-intervention periods.

### 3.3 Inference

This section discusses the methodology of the inferential procedures that are applied in this paper, one of which is based on permutation methods, while another involves random donor pool selection.

#### 3.3.1 Permutation methods

Abadie et al. (2010) provide a method of inference for SCMs that is based on permutation methods. The aim is to estimate 'placebo effects' for each unit in the donor pool by iteratively reassigning the intervention to each untreated unit. The permutation distribution is then given by pooling the placebo effects with the effect estimated for the treated unit. It follows that the effect of the intervention is significant if its magnitude is extreme relative to the permutation distribution.

It may be the case that units in the donor pool suffer poor pre-intervention fit when estimating synthetic controls in the above procedure. Therefore, Abadie et al. (2010) draw inference using the permutation distribution of the ratio between the post-intervention and pre-intervention root mean squared predictor error (RMSPE). This ratio is also used in this paper and is defined as follows:

$$r_j = \frac{R_j(T_0+1, T)}{R_j(1, T_0)}, \tag{6}$$

where $R_j(t_1, t_2)$, for $0 \leq t_1 \leq t_2 \leq T$ and $j = \{1, ..., J+1\}$, is the RMSPE of the synthetic control estimator for unit $j$ and time periods $t_1, ..., t_2$ (see Abadie (2021) for the exact expression). A $p-$value based on the permutation distribution of $r_j$ is given by (see Abadie (2021))

$$p = \frac{1}{J+1} \sum_{j=1}^{J+1} I_+(r_j - r_1), \tag{7}$$

where $I_+(\cdot)$ is an indicator function that is equal to one for nonnegative elements and zero otherwise.

### 3.3.2 Random donor pool selection

In the context of the synthetic control method, the donor pool is a crucial aspect that can significantly influence the results. The selection of units in the donor pool can affect the robustness of the findings. It is possible that one finds a significant treatment effect but that this applies only to a specific donor pool. To mitigate this issue and provide more support for my findings, I estimate synthetic controls for 500 random donor pools, where each donor pool contains 20 units randomly drawn from the 31 available units. Consequently, I assess the consistency of the effect of the intervention by examining the gaps in GDP per capita between the real and synthetic trajectory for all donor pools. Note that some donor pools might suffer a poor pre-intervention fit, which makes the synthetic trajectory of the outcome variable in the post-intervention period unreliable. Therefore, for each donor pool, its synthetic control is only included in the evaluation if the resulting MSPE is at most two times the minimum MSPE of all estimators as produced by the 500 donor pools.

## 4 Results

This section demonstrates the results from the synthetic control method[2] applied in the case of Brexit, for various settings as discussed in Section 3.

First, I construct a synthetic UK where the donor pool includes EU countries and where the time of intervention is 2016. Recall that the weights $v_1, ..., v_k$ ($k = 7$) reflect the relative importance of each predictor. They are provided in Table 3 and are chosen such that MSPE of the synthetic control with respect to the outcome variable (GDP per capita) is minimised (see Section 3.2). The weights indicate that the most important predictor is trade openness, which makes sense given that it was expected that there would be strong co-movement in this variable between the UK and other EU countries as they are all part of the same closely integrated EU trade environment. Consequently, trade openness is followed by, from highest to lowest weight, current account balance, GDP per capita, investment rate, capital account balance, inflation rate, and industry rate.

[2]The applications of the synthetic control method are performed with the Synth package in R, see Jens Hainmueller aut cre, Alexis Diamond aut and Alberto Abadie aut (2011).

**Table 3:** The estimated weights $v_1, ..., v_7$ assigned to each predictor (complete donor pool)

| GDP per capita | Trade openness | Inflation rate | Industry rate | Investment rate | Current account balance | Capital account balance |
|---|---|---|---|---|---|---|
| 0.09 | 0.61 | 0.00 | 0.14 | 0.07 | 0.01 | 0.08 |

*Notes:* The countries that are contained in the complete donor pool are listed in Table 2.

The synthetic control weights are shown in Table 4. Based on these weights, it follows that the synthetic control of the UK is a weighted average of Greece, Iceland, and the United States. The rest of the countries attain a weight of nearly zero.

**Table 4:** Synthetic weights for the UK when using the complete donor pool

| Country | Synthetic control weight | Country | Synthetic control weight |
|---|---|---|---|
| Australia | 0 | Lithuania | 0 |
| Austria | 0 | Luxembourg | 0 |
| Belgium | 0 | Mexico | 0 |
| Canada | 0 | Netherlands | 0 |
| Chile | 0 | New Zealand | 0 |
| Costa Rica | 0 | Norway | 0 |
| Denmark | 0 | Poland | 0 |
| Finland | 0 | Portugal | 0 |
| France | 0 | Slovak Republic | 0 |
| Germany | 0 | Slovenia | 0 |
| Greece | 0.26 | Spain | 0 |
| Iceland | 0.36 | Sweden | 0 |
| Ireland | 0 | Switzerland | 0 |
| Italy | 0 | Turkiye | 0 |
| Korea, Rep. | 0 | United States | 0.38 |
| Latvia | 0 | | |

Table 5 shows the pre-Brexit characteristics of the UK, its synthetic version, and the average of all the countries in the donor pool. It can be observed that overall, the synthetic control matches the UK more closely than the average of the 31 OECD countries. Some of the variables have values that are quite similar for the UK and for the synthetic UK.

Figure 1 visualises the results of the synthetic control estimation, with the trajectory of GDP per capita of the UK and its synthetic control in Figure 1a, and the gap in GDP per capita between the UK and its synthetic control in Figure 1b. We can observe that the synthetic UK resembles the actual UK GDP per capita reasonably well during the pre-Brexit period, except for the period around 2010 and 2015. It should be noted that no clear 'splitting' point can be observed at the time of the intervention. The gap in GDP per capita seems to increase substantially between 2015 and 2016 (see Figure 1b). Therefore, even though there are some years during the pre-intervention period in which the synthetic control does not match the actual UK, there seems to be an effect of which its exact nature remains unclear. That is, it is not clear if 2016 is the right time of intervention as the two trajectories split at an earlier point in time. However, the results suggest that overall the used donor pool is reasonably able to

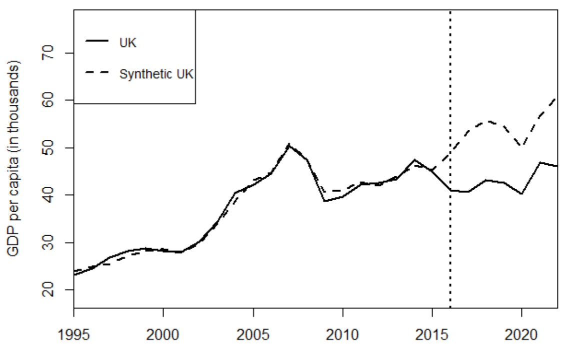
**Table 5:** Predictor means before the intervention (complete donor pool as listed in Table 2)

| | UK | Synthetic UK | Donor pool |
|---|---|---|---|
| GDP per capita | 44.1 | 44.4 | 40.2 |
| Trade openness | 58.7 | 59.8 | 96.5 |
| Inflation rate | 2.3 | 3.2 | 2.5 |
| Industry rate | 18.7 | 19.0 | 24.4 |
| Investment rate | 16.8 | 18.8 | 22.8 |
| Current account balance | -3.6 | -4.8 | -0.2 |
| Capital account balance | -0.0005 | 0.003 | 0.003 |

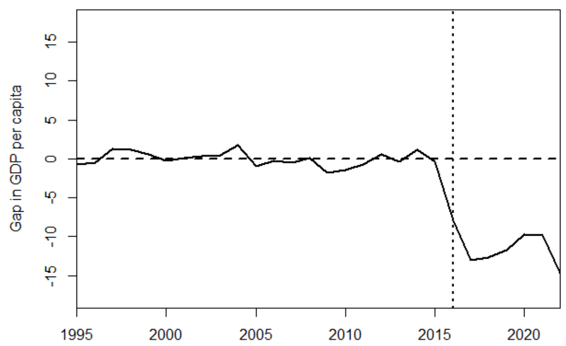
*Notes:* GDP per capita is in current US dollars (thousands), inflation rate is in percent, and trade openness, industry rate, investment rate, current account balance, and capital account balance are reported as percentages of GDP. The time of intervention is 2016 and all variables are measured over the period 2007 until 2015.

produce a combination of OECD countries that tracks the UK GDP per capita. Lastly, observe that the resulting synthetic UK GDP per capita lies above the actual UK GDP per capita over the post-Brexit period, which suggests Brexit had a negative effect.

**Figure 1:** Synthetic control estimation using the complete donor pool (as listed in Table 2)




**(a)** Trajectory of GDP per capita: the UK versus the synthetic UK




**(b)** Gap in GDP per capita between the UK and the synthetic UK

## 4.1 Donor pool that excludes EU countries

As mentioned before, I also estimate a synthetic control for the UK using a donor pool that does not include EU countries because these countries are possibly affected by Brexit which could lead to biased results. The countries that form the EU-excluded donor pool can be found in Table 7, along with their synthetic weights. Performing the estimation in this particular setting yields a synthetic control that is a weighted average of Costa Rica, Iceland, and the United States. As expected, the fit of the model deteriorates when excluding EU countries from the donor pool which can be concluded from the achieved MSPE, as defined in Equation 5. The synthetic control estimator for the complete donor pool attains an MSPE of 0.78, while the MSPE from the model with the EU-excluded donor pool is equal to 1.18.

The model fit has thus worsened, as can also be seen in Table 8, which shows the predictor means before the intervention for the UK, its synthetic control, and the average of all the units in the donor pool. In terms of pre-intervention characteristics, the synthetic control is still more similar to the real UK as compared to the average of the donor pool, which indicates the persistent efficacy of the synthetic control method. However, when comparing the EU-excluded synthetic control estimator to that of the complete donor pool, the values of the predictors are not as close. The reason for the worsened fit is that the removed units (the EU countries) have similar characteristics to the UK and furthermore often experience the same shocks. The EU countries should therefore ideally be included in the donor pool and removing them results in less similarity between the units in the donor pool and the treated unit, which makes it harder to estimate a synthetic control that closely tracks the real UK GDP per capita.

Furthermore, we can observe that the importance of the predictors also changes substantially (see Table 6). While trade openness was the most important predictor in the model with the complete donor pool, it now only achieves a weight of 0.01.

Figure 2 illustrates the results of the estimation when using an EU-excluded donor pool. The plots show that despite the decline in model fit, until 2014, the synthetic control is still able to closely track the actual UK GDP per capita. Moreover, we can again observe an effect as there are notable gaps between the two trajectories after 2014.

**Table 6:** The estimated weights $v_1, ..., v_7$ assigned to each predictor (EU-excluded donor pool)

| GDP per capita | Trade openness | Inflation rate | Industry rate | Investment rate | Current account balance | Capital account balance |
|---|---|---|---|---|---|---|
| 0.84 | 0.01 | 0.00 | 0.00 | 0.05 | 0.09 | 0.00 |

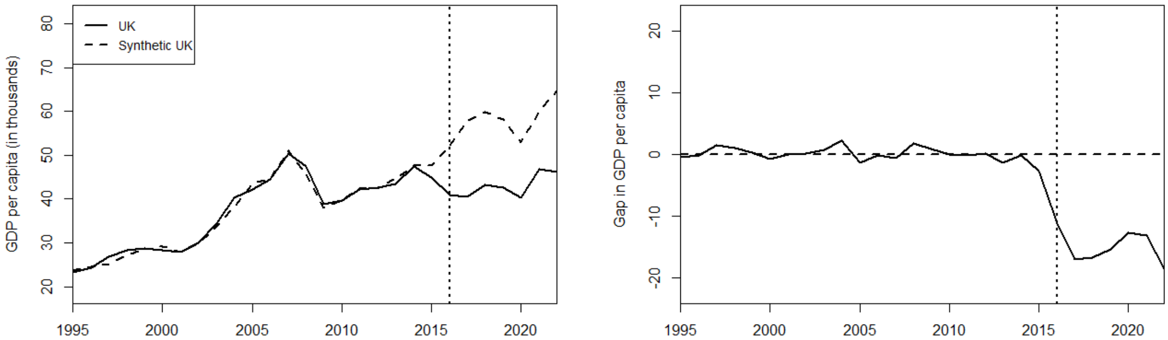
*Notes:* The countries that are contained in the EU-excluded donor pool are listed in Table 7.

**Table 7:** Synthetic weights for the UK when using an EU-excluded donor pool

| Country | Synthetic control weight | Country | Synthetic control weight |
|---|---|---|---|
| Australia | 0 | Mexico | 0 |
| Canada | 0 | New Zealand | 0 |
| Chile | 0 | Norway | 0 |
| Costa Rica | 0.16 | Switzerland | 0 |
| Iceland | 0.46 | Turkiye | 0 |
| Korea, Rep. | 0 | United States | 0.38 |

From the plots, it can be observed that the effect of Brexit seems to be somewhat larger. Note that this is in line with the expectations regarding the direction of the potential bias of the resulting estimator. As mentioned in Section 2.1, when (some of) the comparison units are affected negatively by the intervention, this leads to an underestimation of the synthetic trajectory of the outcome variable. In the case of Brexit, it is possible that this intervention had a negative impact on GDP per capita for some EU countries. As a result, the synthetic UK underestimates GDP per capita because it captures this negative effect. Hence, the difference

**Figure 2:** Synthetic control estimation using a donor pool that excludes EU countries (as listed in Table 7)




**(a)** Trajectory of GDP per capita: the UK versus the synthetic UK

**(b)** Gap in GDP per capita between the UK and the synthetic UK

**Table 8:** Predictor means before the intervention (EU-excluded donor pool as listed in Table 7)

| | UK | Synthetic UK | Donor pool |
|---|---|---|---|
| GDP per capita | 44.1 | 44.4 | 40.3 |
| Trade openness | 58.7 | 65.1 | 68.2 |
| Inflation rate | 2.3 | 4.3 | 3.4 |
| Industry rate | 18.7 | 20.7 | 26.9 |
| Investment rate | 16.8 | 19.6 | 24.5 |
| Current account balance | -3.6 | -4.1 | -0.7 |
| Capital account balance | -0.0005 | -0.0004 | -0.0003 |

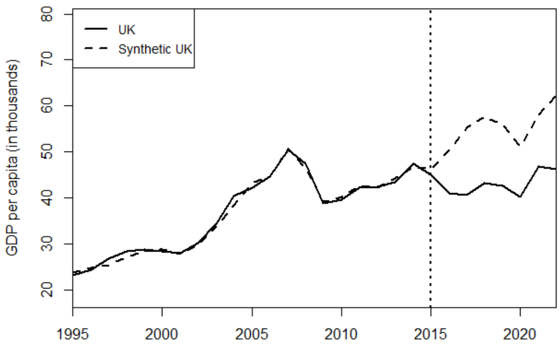
*Notes:* GDP per capita is in current US dollars (thousands), inflation rate is in percent, and trade openness, industry rate, investment rate, current account balance, and capital account balance are reported as percentages of GDP. The time of intervention is 2016 and all variables are measured over the period 2007 until 2015.

between the real and synthetic trajectory is smaller (indicating a smaller intervention effect) and this can also be observed when comparing Figure 1 and Figure 2. However, considering the decrease in model fit, the reliability of the synthetic control produced by the donor pool that excludes EU countries remains up for discussion.

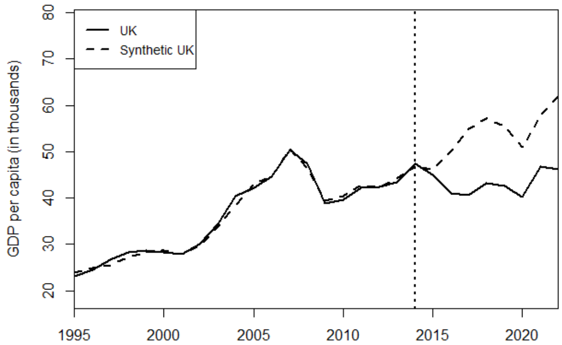
## 4.2 Backdating the intervention

The previous results do not show a clear splitting point of the real and synthetic trajectory of GDP per capita in 2016. This raises the question of whether 2016 is the appropriate time of intervention. To obtain more insights into the timing effect, I estimate synthetic controls by backdating the intervention each year from 2016 to 2012 using the complete donor pool. Figure 3 shows the resulting plots of GDP per capita of the UK versus its synthetic version after backdating Brexit.

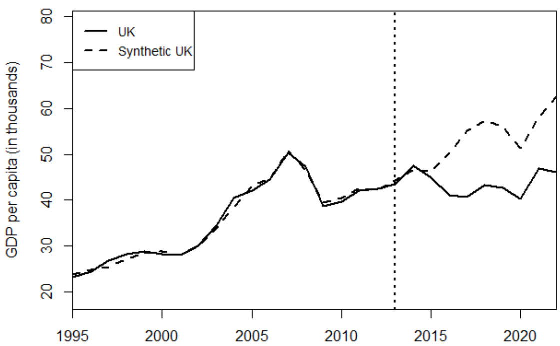
**Figure 3:** Synthetic control estimation when backdating the intervention (complete donor pool as listed in Table 2)



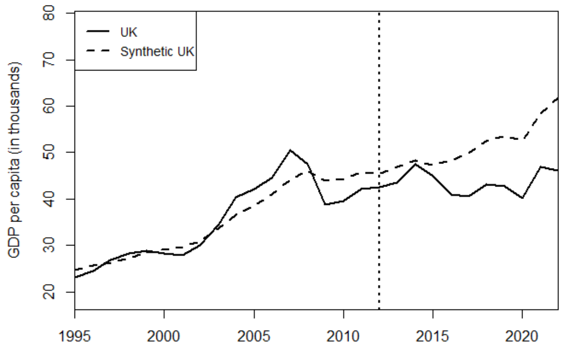

**(a)** Real and synthetic trajectory of UK GDP per capita when the intervention is backdated to 2015




**(b)** Real and synthetic trajectory of UK GDP per capita when the intervention is backdated to 2014




**(c)** Real and synthetic trajectory of UK GDP per capita when the intervention is backdated to 2013



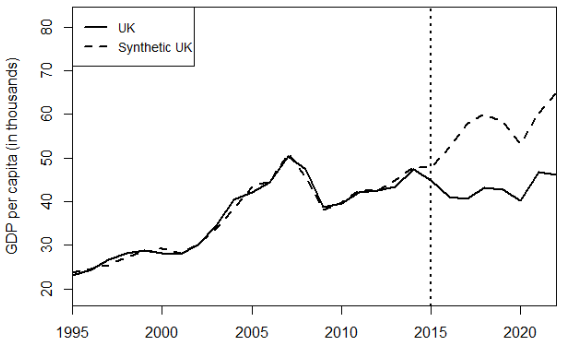

**(d)** Real and synthetic trajectory of UK GDP per capita when the intervention is backdated to 2012

In all the plots the splitting point seems to be in 2014, which is particularly evident in the plots of 2015, 2014, and 2013. However, it should be noted that in the plot of 2012, there are substantial gaps between the UK and its synthetic control in the pre-treatment period which indicates that in the case of the complete donor pool, backdating until 2012 is too far back in time.

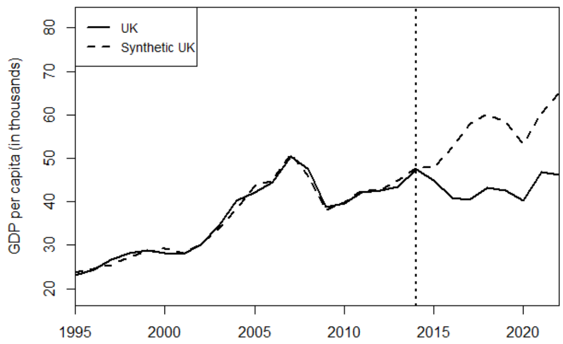
When using an EU-excluded donor pool, we can make the same observation that the splitting point is in 2014, in this case for each time of intervention: 2016, 2015, 2014, 2013, and 2012, see Figure 4. Recall that the plot of 2016 was shown in Figure 2b in Section 4.1. Additionally, the pre-intervention fit when backdating to 2012 is much better for the EU-excluded donor pool compared to the complete donor pool.

Overall, for both donor pools, the synthetic control estimators with a backdated intervention attain lower MSPE's (see Table 9), indicating a better pre-intervention fit. For the years 2013, 2014, and 2015 the MSPE remains almost the same when using the complete donor pool. This

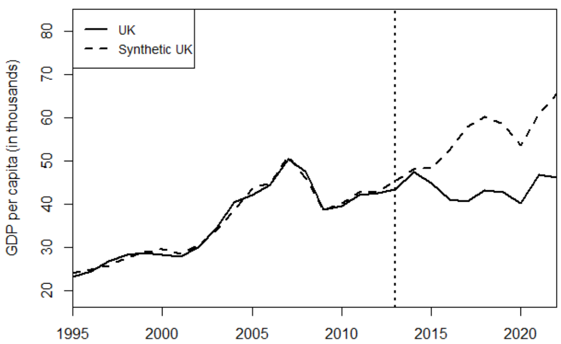
**Figure 4:** Synthetic control estimation when backdating the intervention (EU-excluded donor pool as listed in Table 7)



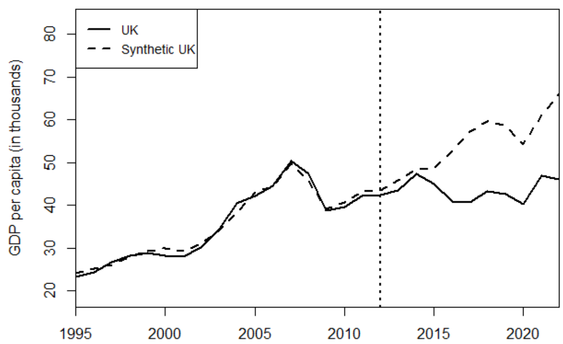

**(a)** Real and synthetic trajectory of UK GDP per capita when the intervention is backdated to 2015




**(b)** Real and synthetic trajectory of UK GDP per capita when the intervention is backdated to 2014




**(c)** Real and synthetic trajectory of UK GDP per capita when the intervention is backdated to 2013




**(d)** Real and synthetic trajectory of UK GDP per capita when the intervention is backdated to 2012

is also the case for the synthetic control estimators that are obtained using an EU-excluded donor pool, although the differences are somewhat larger. We again observe that the MSPE is higher when using a donor pool without EU countries (as expected), except for the year 2012. Note that the improvement in MSPE is also reflected in the graphs in Figure 3 (complete donor pool) and Figure 4 (EU-excluded donor pool). We can observe that indeed the synthetic and real trajectories match more closely.

Abadie (2021) mentions that in the absence of anticipation effects, backdating can be applied to assess the credibility of a synthetic control estimator. The authors provide two credibility criteria. First, after backdating the intervention, the synthetic control estimator should closely track the outcome variable of the treated unit from the period the intervention was backdated to until the actual intervention, i.e., there should be no effects during this period when artificially reassigning the time of intervention (see Abadie et al. (2015)). Second, even when backdating the intervention, which causes the procedure to not use any information on the timing of the

**Table 9:** MSPE's for different times of intervention

| Time of intervention | MSPE (complete donor pool) | MSPE (EU-excluded donor pool) |
|---|---|---|
| 2012 | 9.04 | 1.14 |
| 2013 | 0.65 | 0.80 |
| 2014 | 0.65 | 0.89 |
| 2015 | 0.63 | 0.85 |
| 2016 | 0.78 | 1.18 |

*Notes:* The countries that are contained in the complete donor pool are listed in Table 2. Refer to Table 7 for the countries in the EU-excluded donor pool.

actual intervention, the gap between the real and synthetic trajectories should appear at the time of the actual intervention (Abadie, 2021).

The results obtained so far suggest that 2014 might be a more appropriate time of intervention compared to 2016. In the introduction, it was mentioned that the first concrete event putting Brexit on the agenda was in 2013, when David Cameron promised a Brexit referendum by 2017 (Varga, 2023). Economic agents may therefore have anticipated Brexit well in advance of the 2016 vote, and these anticipation effects could have affected GDP per capita prior to the vote. Note that simply choosing 2014 as the time of Brexit, and therefore concluding that the effects on UK GDP per capita can be devoted solely to this event, is not the optimal approach. However, it should be noted that Brexit was arguably the most significant event during those years. Nevertheless, the unclear nature of its actual timing complicates the application of the synthetic control method, which requires a clear point in time at which the intervention happened.

All things considered, I continue the analysis using 2014 as the time of intervention, first considering the complete donor pool. The weights $v_1, ..., v_7$ in Table 10 indicate that the most important predictors, from highest to lowest weight, are GDP per capita, capital account balance, trade openness, industry rate, investment rate, inflation rate, and current account balance. Notice that the importance of the predictors changes substantially compared to the initial time of intervention 2016 (see Table 3 for the original weights). Initially, GDP per capita only attained a weight of 0.09, whereas after backdating the intervention to 2014 this variable becomes the most important predictor with a weight of 0.83. Conversely, the weight of the initially most important predictor trade openness decreases from 0.61 to 0.05.

Table 11 shows the synthetic control weights when performing the estimation using 2014 as the time of intervention. The resulting synthetic control is now a weighted average of Costa Rica, Greece, Iceland, Italy, and the United States.

**Table 10:** The estimated weights $v_1, ..., v_7$ assigned to each predictor (time of intervention 2014)

| GDP per capita | Trade openness | Inflation rate | Industry rate | Investment rate | Current account balance | Capital account balance |
|---|---|---|---|---|---|---|
| 0.83 | 0.05 | 0.00 | 0.02 | 0.02 | 0.00 | 0.08 |

*Notes:* The estimation is performed using the complete donor pool as listed in Table 2.

The estimation results for the EU-excluded donor pool are as follows. The most important

**Table 11:** Synthetic weights for the UK (time of intervention 2014)

| Country | Synthetic control weight | Country | Synthetic control weight |
|---|---|---|---|
| Australia | 0 | Lithuania | 0 |
| Austria | 0 | Luxembourg | 0 |
| Belgium | 0 | Mexico | 0 |
| Canada | 0 | Netherlands | 0 |
| Chile | 0 | New Zealand | 0 |
| Costa Rica | 0.08 | Norway | 0 |
| Denmark | 0 | Poland | 0 |
| Finland | 0 | Portugal | 0 |
| France | 0 | Slovak Republic | 0 |
| Germany | 0 | Slovenia | 0 |
| Greece | 0.10 | Spain | 0 |
| Iceland | 0.39 | Sweden | 0 |
| Ireland | 0 | Switzerland | 0 |
| Italy | 0.07 | Turkiye | 0 |
| Korea, Rep. | 0 | United States | 0.37 |
| Latvia | 0 | | |

*Notes:* This table corresponds to the complete donor pool.

predictors, from highest to lowest weight, are GDP per capita, capital account balance, trade openness, industry rate, investment rate, inflation rate, and current account balance, see Table 13 for the weights. Notice that these weights are identical to the weights from the estimation with the complete donor pool and time of intervention 2014, see Table 10. However, when we compare the weights from the initial time of intervention, as shown in Table 6, to the new weights, we can observe that these do not change much. Hence, backdating the intervention in the case of the EU-excluded donor pool does not have a large effect on the importance of the predictors.

Table 14 shows the synthetic weights when using the EU-excluded donor pool and time of intervention 2014. The resulting synthetic control produced from the EU-excluded donor pool is now a weighted average of Costa Rica, Iceland, and the United States.

Assuming 2014 is a reliable timing of Brexit, which should be considered a limitation of this research, we can draw the conclusion that the synthetic control estimators are reasonably credible for the following reasons. First, when backdating the intervention from 2014 to 2013, the synthetic UK GDP per capita still closely tracks the actual UK GDP per capita in the period 2013-2014. However, when backdating to 2012, this conclusion can only be drawn for the synthetic control estimator that is obtained through an EU-excluded donor pool. Second, when backdating to 2012, the gap in GDP per capita starts from 2014 (i.e., the 'actual' time of intervention). Hence, backdating the intervention does not affect the splitting point of the real and synthetic trajectories. Recall that this can be observed from Figure 3 (complete donor pool) and Figure 4 (EU-excluded donor pool) which show that the graphs remain mostly the same when backdating the intervention.

Note that on average, the actual UK GDP per capita lies 9,840 USD below its synthetic counterpart over the post-treatment (i.e., from 2014 until 2022) period. In the pre-treatment

**Table 12:** Predictor means before the intervention (time of intervention 2014)

| | UK | Synthetic UK | Donor pool |
|---|---|---|---|
| GDP per capita | 43.5 | 43.6 | 40.2 |
| Trade openness | 58.8 | 61.2 | 95.3 |
| Inflation rate | 2.7 | 4.4 | 2.9 |
| Industry rate | 18.9 | 20.4 | 24.6 |
| Investment rate | 16.6 | 19.4 | 23.0 |
| Current account balance | -3.3 | -5.6 | -0.6 |
| Capital account balance | -0.0004 | 0.001 | 0.004 |

*Notes:* GDP per capita is in current US dollars (thousands), inflation rate is in percent, and trade openness, industry rate, investment rate, current account balance, and capital account balance are reported as percentages of GDP. The time of intervention is 2014 and all variables are measured over the period 2007 until 2013. Lastly, note that this table concerns the complete donor pool as listed in Table 2.

**Table 13:** The estimated weights $v_1, ..., v_7$ assigned to each predictor (time of intervention 2014)

| GDP per capita | Trade openness | Inflation rate | Industry rate | Investment rate | Current account balance | Capital account balance |
|---|---|---|---|---|---|---|
| 0.83 | 0.05 | 0.00 | 0.02 | 0.02 | 0.00 | 0.08 |

*Notes:* The estimation is performed using an EU-excluded donor pool as listed in Table 14.

period, the actual UK GDP per capita is on average 90 USD higher than its synthetic version (the average of the absolute gaps is equal to 666 USD). For the EU-excluded donor pool, the actual GDP per capita of the UK is 12,262 USD less than the GDP per capita of the synthetic UK over the post-treatment period, while in the pre-treatment period, the actual UK GDP per capita is on average 4 USD lower (with the average of the absolute gaps being equal to 741 USD).

These results indicate that the (negative) effect of Brexit on UK GDP per capita can be quantified by an amount between 9,840 and 12,262 USD. Note that, considering the direction of the potential bias of the estimator due to affected control units (EU countries), the amount of 9,840 USD likely is an underestimation of the real effect. This is also indicated by the amount of 12,262 USD which represents the difference in the outcome of interest when the synthetic control is produced using an EU-excluded donor pool, which is indeed higher than when using a donor pool with EU countries.

Table 16 shows the average gaps between the real UK GDP per capita and its synthetic counterpart over the post-intervention period for each different time of intervention. Note that these are calculated by subtracting the real UK GDP per capita from the synthetic UK GDP per capita such that a positive gap indicates a negative effect of Brexit. As we already observed from the plots in Figures 1 through 4, the synthetic UK always lies above the actual UK, hence the gaps in Table 16 are all positive. The gaps from the estimation based on the complete donor pool vary between 8,734 USD and 10,103 USD, while the EU-excluded donor pool yields synthetic controls that result in gaps between 11,982 USD and 12,667 USD. Note that for the EU-

**Table 14:** Synthetic weights for the UK (time of intervention 2014)

| Country | Synthetic control weight | Country | Synthetic control weight |
|---|---|---|---|
| Australia | 0 | Mexico | 0 |
| Canada | 0 | New Zealand | 0 |
| Chile | 0 | Norway | 0 |
| Costa Rica | 0.16 | Switzerland | 0 |
| Iceland | 0.45 | Turkiye | 0 |
| Korea, Rep. | 0 | United States | 0.39 |

*Notes:* This table corresponds to the EU-excluded donor pool.

**Table 15:** Predictor means before the intervention (EU-excluded donor pool)

| | UK | Synthetic UK | Donor pool |
|---|---|---|---|
| GDP per capita | 43.5 | 43.5 | 39.8 |
| Trade openness | 58.8 | 64.4 | 68.5 |
| Inflation rate | 2.7 | 5.1 | 3.7 |
| Industry rate | 18.9 | 21.0 | 27.2 |
| Investment rate | 16.6 | 19.6 | 24.5 |
| Current account balance | -3.3 | -5.5 | -1.0 |
| Capital account balance | 0.0004 | -0.0003 | 0.001 |

*Notes:* GDP per capita is in current US dollars (thousands), inflation rate is in percent, and trade openness, industry rate, investment rate, current account balance, and capital account balance are reported as percentages of GDP. The time of intervention is 2014 and all variables are measured over the period 2007 until 2013. Lastly, note that this table concerns the EU-excluded donor pool as listed in Table 14.

excluded donor pool, the gaps are more stable across the different times of variation compared to the complete donor pool. These results further indicate the negative effect of Brexit on UK GDP per capita. For both donor pools and each time of intervention, we find that the synthetic UK GDP per capita lies above the actual UK GDP per capita, with the gap ranging from 8,734 USD up to 12,667 USD.

Recall that in Section 4.1 it was shown that when using an EU-excluded donor pool, the post-intervention effect is somewhat larger. It was discussed that this is in line with the expectations regarding the potential bias of the estimator that is based on a donor pool that includes potentially affected units (in this case the complete donor pool as listed in Table 2). The observation that the effect is somewhat larger when using a donor pool that solely includes unaffected units persists after backdating the intervention. This can be observed from the rightmost column in Table 16, which corresponds to the resulting gaps from the EU-excluded donor pools for each different time of intervention. The gaps in this column are all larger compared to the gaps from the complete donor pool.

To conclude, this section revealed that backdating the intervention provides valuable insights, of which the most important one is that the appropriate time of intervention is 2014, rather than 2016 which the initial setup relied upon. It holds for both donor pools that for each time of intervention, the real and synthetic trajectories match reasonably closely and that the

**Table 16:** Average gap between real and synthetic UK GDP per capita over the post-Brexit period for different times of intervention

| | Complete donor pool (as in Table 2) | EU-excluded donor pool (as in Table 14) |
|---|---|---|
| 2012 | 8,845 | 12,667 |
| 2013 | 10,063 | 12,512 |
| 2014 | 9,840 | 12,262 |
| 2015 | 10,103 | 12,195 |
| 2016 | 8,734 | 11,982 |

*Notes:* The gaps are in USD and are computed by subtracting the real UK GDP per capita from its synthetic counterpart and taking the average over the post-intervention period.

splitting point between the real and synthetic UK GDP per capita appears in 2014, with the only exception being backdating the intervention to 2012 using the complete donor pool. Here, the fit deteriorates substantially such that the post-Brexit trajectories of UK GDP per capita are not entirely reliable, see the plot in Figure 3d. This indicates that backdating Brexit to 2012 might be too far back in time.

However, as discussed, 2014 is assumed to be appropriate as this time of intervention leads to reliable results. These results show that the negative effect on UK GDP per capita that follows after backdating Brexit to 2014 lies between 9,840 and 12,262 USD. It should be noted that we cannot say with certainty to what extent the effect on UK GDP per capita can be attributed to Brexit, as the estimation relies on an assumption about the timing of the intervention.

# 5 Permutation analysis

To obtain the permutation distribution of the ratio $r_j$ (i.e., the ratio between the post-intervention and pre-intervention RMSPE) defined in Equation 6, I iteratively reassign the treatment (in 2014) to each of the units in the complete donor pool. The $p$-value based on the found permutation distribution, as defined in Equation 7, is equal to 0, as all $r_j, j = 2, ..., J+1$ are smaller than $r_1 = 14.49$. This indicates that the effect of the treatment (i.e., the post-intervention impact) is statistically significant and not due to random chance.

To visualise the results of the permutation methods, Figure 5 shows the gaps in GDP per capita for each of the control units when the intervention is assigned to that unit, as well as the gap in GDP per capita for the actual treated unit (i.e., the UK) which is indicated by the black line. Note that the only country to which the intervention could not be applied is Ireland as the Synth package (Jens Hainmueller aut cre et al., 2011) was not able to construct a synthetic control for this unit. Moreover, only estimators that attain an MSPE that is at most five times the MSPE of the original treatment unit, i.e., the UK, are included in the plot. From the figure, we can observe that the UK substantially differentiates itself from the rest of the units. Overall, there is a low probability for a control unit to obtain a gap as large as the gap of the UK. This further indicates the significance of the intervention effect.

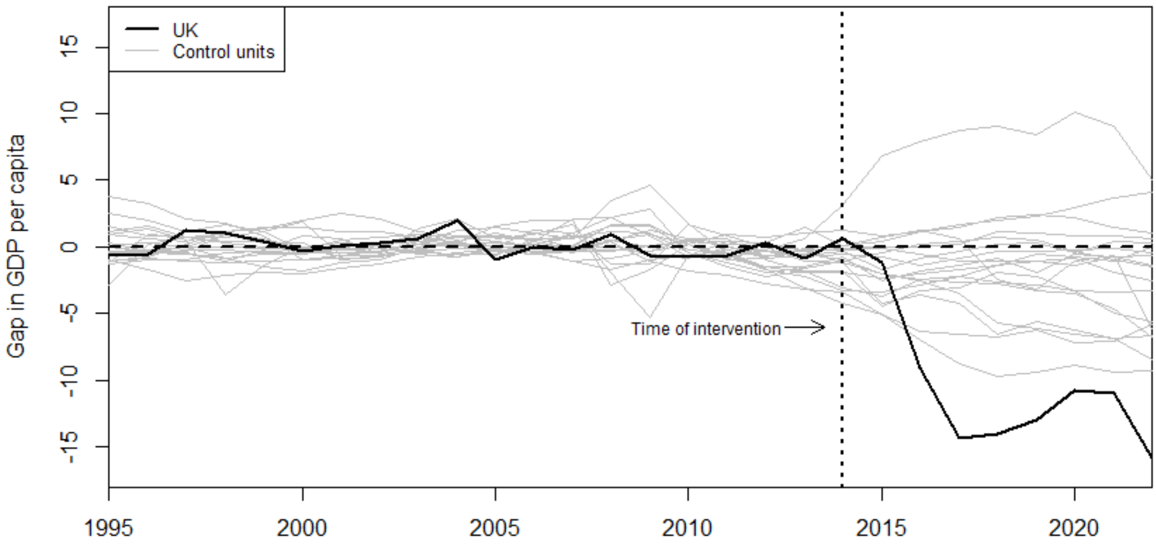



**Figure 5:** The gaps between the real and synthetic GDP per capita when the intervention is iteratively reassigned to each unit, except Ireland, in the complete donor pool (as listed in Table 2)

# 6 Random donor pools

Figure 6 shows the gaps for the randomly selected donor pools, where the time of intervention is 2014. Note that for visibility purposes, only donor pools that attain an MSPE of at most three times the minimum MSPE of all donor pools are included. We can observe that after the intervention, for all donor pools there is a substantial gap in GDP per capita. This indicates that the negative effect of the outcome variable is not susceptible to a specific donor pool which further supports the finding that Brexit had a negative effect on UK GDP per capita. Note however that over the 500 donor pools, the average post-treatment gap is equal to 6,017 USD. This is lower than the gaps from the fixed complete and EU-excluded donor pools, which are shown in Table 16.

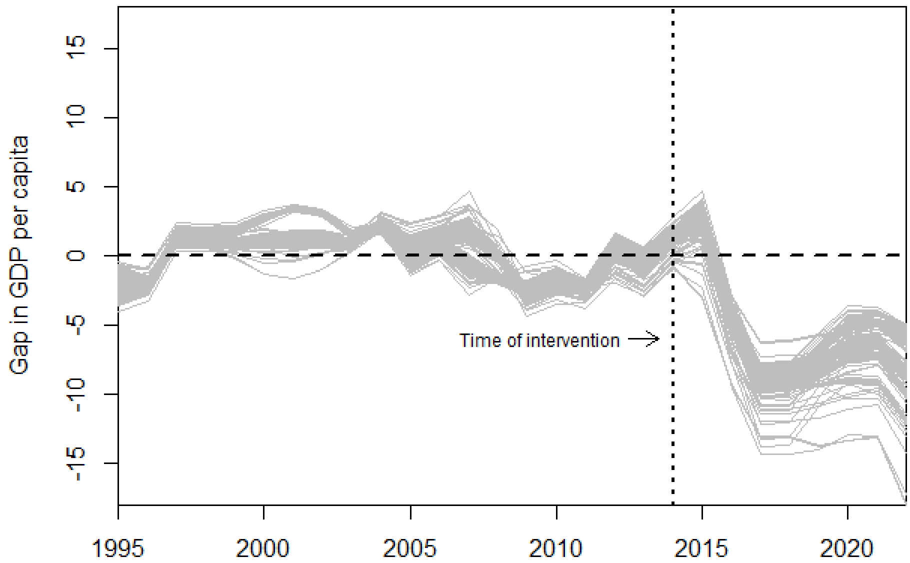



**Figure 6:** Gaps for different donor pools. Note that in this plot only donor pools that attain an MSPE of at most three times the minimum MSPE are included

## 7 Conclusion

This paper studied the effect of Brexit on UK GDP per capita by means of the synthetic control method as proposed in Abadie et al. (2015). The method is employed in various settings to ensure the robustness of the findings. First, the synthetic control is estimated using the complete donor pool and the year 2016 as time of intervention. The results show that the fit of the model is reasonably accurate but the synthetic control is only able to track actual UK GDP per capita until 2014. After 2014, the real and synthetic trajectories of UK GDP per capita split, with the real UK GDP per capita lying below its synthetic counterpart. This indicates a negative effect of Brexit on the outcome of interest, which is equal to 8,734 USD.

Additionally, an EU-excluded donor pool is considered to avoid the potential bias caused by affected control units. As expected, the pre-intervention fit of the model based on this donor pool is not as good as that of the complete donor pool. However, the difference is not substantial. When evaluating the post-intervention gap, we observe that at 11,982 USD, this is larger as compared to the complete donor pool. This is in line with the expectations regarding the potential bias of the estimator, i.e., the synthetic control underestimates the trajectory as it incorporates negative effects on the control units.

To investigate the timing effect, the intervention is backdated until 2012. The results mostly show a decreasing pattern in MSPE along with a negative intervention effect ranging from 9,840 to 12,262 USD. We observe that in each estimation, the trajectories of the real and synthetic UK GDP per capita mostly stay the same and that in all cases, the splitting point between these two trajectories appears to be in 2014. This suggests that 2014 might be a more appropriate time of intervention, which implies the effects of Brexit may have already started to unfold in 2014 due to anticipation in the economy. However, we cannot state with certainty to what extent the estimated negative effect on UK GDP per capita can truly be devoted to Brexit. Nevertheless, considering its extensive timeline of events (leading to anticipation effects), it is likely that Brexit is responsible for a significant portion.

Furthermore, permutation methods and random donor pools are considered to provide more robustness to the findings. The permutation analysis shows that the effect of the intervention in the UK is significant, as no substantial gaps between real and synthetic GDP per capita are attained when assigning the intervention to all the units in the donor pool. Moreover, when performing the synthetic control method using the UK as the treated unit and 2014 as the time of intervention for 500 different donor pools, the gaps remain substantial for each donor pool. This indicates that the intervention effect is not susceptible to the selection of the donor pool, which provides robustness to the finding that Brexit had a negative effect on UK GDP per capita. This effect is however lower compared to the fixed donor pools, at 6,017 USD.

## References

Abadie, A. (2021, June). Using synthetic controls: Feasibility, data requirements, and methodological aspects. *Journal of Economic Literature*, *59*(2), 391-425. Retrieved from `https://www.aeaweb.org/articles?id=10.1257/jel.20191450` doi: 10.1257/jel.20191450

Abadie, A., Diamond, A. & 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. Retrieved from `https://doi.org/10.1198/jasa.2009.ap08746` doi: 10.1198/jasa.2009.ap08746

Abadie, A., Diamond, A. & Hainmueller, J. (2015). Comparative politics and the synthetic control method. *American Journal of Political Science*, *59*(2), 495-510. doi: https://doi.org/10.1111/ajps.12116

Abadie, A. & Gardeazabal, J. (2003, March). The economic costs of conflict: A case study of the basque country. *American Economic Review*, *93*(1), 113-132. Retrieved from `https://www.aeaweb.org/articles?id=10.1257/000282803321455188` doi: 10.1257/000282803321455188

Arnorsson, A. & Zoega, G. (2018). On the causes of Brexit. *European Journal of Political Economy*, *55*, 301–323. doi: 10.1016/j.ejpoleco.2018.02.001

Dennison, J. & Carl, N. (2016). The ultimate causes of Brexit: history, culture, and geography. *British Politics and Policy at LSE*.

Dhingra, S. & Sampson, T. (2022). Expecting Brexit. *Annual Review of Economics*, *14*, 495–519. doi: 10.1146/annurev-economics-051420-104231

Di Stefano, R. & Mellace, G. (2020). *The inclusive synthetic control method* (Discussion Paper on Business and Economics No. 14). University of Southern Denmark. doi: 10.2139/ssrn.3737491

Divya, K. H. & Devi, V. R. (2014). A study on predictors of GDP: Early signals. *Procedia Economics and Finance*, *11*, 375-382. Retrieved from https://www.sciencedirect.com/science/article/pii/S2212567114002056 (Shaping the Future of Business and Society) doi: https://doi.org/10.1016/S2212-5671(14)00205-6

Farid, M. (2020). The effect of Brexit on United Kingdom productivity: Synthetic control analysis. *Journal of Applied Economic Sciences (JAES)*, *15*, 791–800.

Jens Hainmueller aut cre, Alexis Diamond aut & Alberto Abadie aut. (2011). Synth: An R package for synthetic control methods in comparative case studies. *Journal of Statistical Software*, *42*(13), 1–17. Retrieved from https://www.jstatsoft.org/v42/i13/

Saunders, R. (2018). *Yes to Europe! The 1975 referendum and seventies Britain.* Cambridge University Press.

Tournier-Sol, K. (2021). From UKIP to the Brexit party: the politicization of European integration and disruptive impact on national and European arenas. *Journal of Contemporary European Studies*, *29*(3), 380–390. doi: 10.1080/14782804.2020.1785849

Varga, T. (2023). What was David Cameron thinking? Thoughts of a British prime minister regarding Brexit. *Journal of Global Awareness*, *4*(1).

Whiteley, P., Goodwin, M. J. & Clarke, H. D. (2018). *The rise and fall of UKIP 2010-17.* Manchester University Press.

# A Programming code

```
library(Synth)
library(SCtools)
library(tidyverse)
library(skimr)
library(readxl)
library(dplyr)
library(naniar)
library(visdat)
library(MSCMT)
library(xtable)

##Main results

#Read data, assign NA's to empty variables, and rename variables

df <- read_excel("Transformed dataframe.xlsx")

df[df == ".."] <- NA

df <- rename(df, "Country" = "Country Name",
             "Capital account balance" = "Capital account balance (% of GDP))",
             "Current account balance" = "Current account balance (% of GDP)",
             "GDP per capita" = "GDP per capita (current US$)",
```

```
            "Gross capital formation" = "Gross capital formation (% of GDP)",
            "Industry rate" = "Industry (including construction), value added (% of GDP)",
            "Inflation rate" = "Inflation, consumer prices (annual %)",
            "Schooling" = "School enrollment, secondary (% net)",
            "Trade openness" = "Trade (% of GDP)")

#Assign a unit number to each country and convert each variable to numeric

df$Unit <- match(df$Country, unique(df$Country))

df$Unit <- as.numeric(df$Unit)

df <- as.data.frame(df)

df$Year <- as.numeric(df$Year)
df$`Capital account balance` <- as.numeric(df$`Capital account balance`)
df$`Current account balance` <- as.numeric(df$`Current account balance`)
df$`GDP per capita` <- as.numeric(df$`GDP per capita`)
df$`Gross capital formation` <- as.numeric(df$`Gross capital formation`)
df$`Industry rate` <- as.numeric(df$`Industry rate`)
df$`Inflation rate` <- as.numeric(df$`Inflation rate`)
df$Schooling <- as.numeric(df$Schooling)
df$`Trade openness` <- as.numeric(df$`Trade openness`)

#Create some new variables

df["GDP growth rate"] = c(NaN,100*diff(log(df$`GDP per capita`))) #not used
df["GDP pc thousands"] = df["GDP per capita"]/1000
df["BoP"] = df$`Capital account balance`+df$`Current account balance` #not used

#Calculate the averages of the predictor variables for each country over the
#pre-intervention period

averages <- df %>%
  filter(Year >= 1995 & Year <= 2015) %>%
  group_by(Country) %>%
  summarize(
    GDPpc = as.character(round(mean(`GDP pc thousands`, na.rm = TRUE), 1)),
    trade = as.character(round(mean(`Trade openness`, na.rm = TRUE), 1)),
    capital_acc = as.character(format(signif(mean(`Capital account balance`, na.rm = TRUE),
                                  scientific = TRUE)),
    current_acc = as.character(round(mean(`Current account balance`, na.rm = TRUE), 1)),
```

```
    investment = as.character(round(mean(‘Gross capital formation‘, na.rm = TRUE), 1)),
    industry = as.character(round(mean(‘Industry rate‘, na.rm = TRUE), 1)),
    inflation = as.character(round(mean(‘Inflation rate‘, na.rm = TRUE), 1))
  )

print(xtable(averages),include.rownames = FALSE, include.colnames = FALSE)

#Create vectors for the  control units

all_available_countries <- c(1:5, 7, 9, 11:14, 16, 17, 19, 21:36, 38)
available_countries_excl_EU <- c(1, 4, 5, 7, 16,21,25,27,28,35,36,38)
EU_countries <- c(2,3,9,11,12,13,14,17,19,22,23,24,26,29,30,31,32,33,34)


#Perform Synth for 7 different intervention years using all of the
#available countries in the donor pool.

intervention_years <- 2012:2016

for(year in intervention_years){

  dataprep.out <- dataprep(
    foo = df,

    predictors = c("GDP per capita",
                   "Gross capital formation",
                   "Inflation rate",
                   "Trade openness"),

    special.predictors = list(
      list("Industry rate", 2007:(year-1), "mean"),
      list("Capital account balance", 2007:(year-1), "mean"),
      list("Current account balance", 2007:(year-1), "mean")
    ),


    predictors.op = "mean",

    time.predictors.prior = 2007:(year-1),

    dependent = "GDP pc thousands",
```

```
  unit.variable = "Unit",

  unit.names.variable = "Country",

  time.variable = "Year",

  treatment.identifier = 37,

  controls.identifier = all_available_countries,

  time.optimize.ssr = 1995:(year-1),

  time.plot = 1995:2022
)

print(paste("Time of intervention: ", year))

synth.out <- synth(data.prep.obj = dataprep.out)

print(paste("Average of absolute gap pre-treatment: ",
            mean(abs(dataprep.out$Y1plot[1:19] -
            (dataprep.out$Y0plot[1:19,] %*% synth.out$solution.w)))))

print(paste("Average gap pre-treatment: ",
            mean(dataprep.out$Y1plot[1:19] -
            (dataprep.out$Y0plot[1:19,] %*% synth.out$solution.w))))

print(paste("Average gap post-treatment: ",
            mean(dataprep.out$Y1plot[20:28] -
            (dataprep.out$Y0plot[20:28,] %*% synth.out$solution.w))))

path.plot(synth.res = synth.out,
           dataprep.res = dataprep.out,
           tr.intake = year,
           Ylab = "GDP per capita (in thousands)",
           Xlab = "",
           Legend = c("UK", "Synthetic UK"),
           Legend.position = "topleft",
 )

 gaps.plot(synth.res = synth.out,
           dataprep.res = dataprep.out,
```

```
            tr.intake = year,
            Ylab = "Gap in GDP per capita",
            Xlab = "",
            Main = ""
   )
}

#Perform Synth for 7 different intervention years using all of the
#available countries excluding EU countries in the donor pool.


for(year in intervention_years){

  print(paste("Time of intervention: ", year))

  dataprep.out <- dataprep(
    foo = df,

    predictors = c("GDP per capita",
                   "Gross capital formation",
                   "Inflation rate",
                   "Trade openness"),

    special.predictors = list(
      list("Industry rate", 2007:(year-1), "mean"),
      list("Capital account balance", 2007:(year-1), "mean"),
      list("Current account balance", 2007:(year-1), "mean")
    ),


    predictors.op = "mean",

    time.predictors.prior = 2007:(year-1),

    dependent = "GDP pc thousands",

    unit.variable = "Unit",

    unit.names.variable = "Country",

    time.variable = "Year",
```

```
    treatment.identifier = 37,

    controls.identifier = available_countries_excl_EU,

    time.optimize.ssr = 1995:(year-1),

    time.plot = 1995:2022
  )

  synth.out <- synth(data.prep.obj = dataprep.out)

  print(paste("Average of absolute gap pre-treatment: ",
              mean(abs(dataprep.out$Y1plot[1:19] -
              (dataprep.out$Y0plot[1:19,] %*% synth.out$solution.w)))))

  print(paste("Average gap pre-treatment: ",
              mean(dataprep.out$Y1plot[1:19] -
              (dataprep.out$Y0plot[1:19,] %*% synth.out$solution.w))))

  print(paste("Average gap post-treatment: ",
              mean(dataprep.out$Y1plot[20:28] -
              (dataprep.out$Y0plot[20:28,] %*% synth.out$solution.w))))

   path.plot(synth.res = synth.out,
             dataprep.res = dataprep.out,
             tr.intake = year,
             Ylab = "GDP per capita (in thousands)",
             Xlab = "",
             Legend = c("UK", "Synthetic UK"),
             Legend.position = "topleft"
   )

   gaps.plot(synth.res = synth.out,
             dataprep.res = dataprep.out,
             tr.intake = year,
             Ylab = "Gap in GDP per capita",
             Xlab = "",
             Main = ""
   )
}

##Permutation methods
```

```
#Create variable for the year of intervention. Note that the year of
#intervention is 2014 but 2013 is stored into this variable as it is
#the last year of data that is used in the optimisation procedure of
#synth
year_of_intervention <- 2013

all_countries_incl_uk <- c(37, all_available_countries)

rmspe_ratios_2014 <- numeric(0)

gaps_permutation_2014 <- matrix(NA, length(1995:2022), 32)

col_counter <- 1

#Perform Synth for each unit, including the treated unit
#Note that the only unit to which the intervention is not
#applied is Ireland. Unfortunately, synth is not ablo estimate
#a synthetic control for this country. This unit is excluded from
#the loop.
for(unit in all_countries_incl_uk[-14]){
  dataprep.out <- dataprep(
    foo = df,

    predictors = c("GDP per capita",
                   "Gross capital formation",
                   "Inflation rate",
                   "Trade openness"),

    special.predictors = list(
      list("Industry rate", 2007:(year_of_intervention), "mean"),
      list("Capital account balance", 2007:(year_of_intervention), "mean"),
      list("Current account balance", 2007:(year_of_intervention), "mean")
    ),


    predictors.op = "mean",

    time.predictors.prior = 2007:(year_of_intervention),

    dependent = "GDP pc thousands",
```

```
    unit.variable = "Unit",

    unit.names.variable = "Country",

    time.variable = "Year",

    treatment.identifier = unit,

    controls.identifier = all_countries_incl_uk[! all_countries_incl_uk %in% c(unit)],

    time.optimize.ssr = 1995:(year_of_intervention),

    time.plot = 1995:2022
  )

  synth.out <- synth(data.prep.obj = dataprep.out)

  #Calculate pre-intervention RMSPE

  rmspe_pre <- sqrt(synth.out$loss.v)

  #Calculate post-intervention RMSPE

  rmspe_post <- sqrt((mean((dataprep.out$Y1plot[20:28] -
                            dataprep.out$Y0plot[20:28,]%*% synth.out$solution.w)^2)))

  #Calculate ratio of post-intervention RMSPE relative to pre-intervention RMSPE, see
  #Equation 12 in Abadie (2021)

  r_j <- rmspe_post / rmspe_pre

  rmspe_ratios_2014 <- append(rmspe_ratios_2014, r_j)

  gaps_permutation_2014[,col_counter] <- dataprep.out$Y1plot -
    (dataprep.out$Y0plot %*% synth.out$solution.w)

  col_counter <- col_counter + 1
}

#Calculate p-value for the inferential procedure based on the permutation distribution
#of r_j
```

```
p_value <- (1/32)*sum((rmspe_ratios_2014 - rmspe_ratios_2014[1])*
                        if_else(rmspe_ratios_2014 - rmspe_ratios_2014[1] > 0, 1, 0))


#Calculate pre-intervention MSPE's and select units that attain an MSPE of at most
#5 times the MSPE of the treated unit to include in the plot
mspe_2014 <- apply(gaps_permutation_2014[1:19,]^2,2,mean)
uk.mspe_2014 <- as.numeric(mspe_2014[1])
gaps_permutation_controls_2014 <- gaps_permutation_2014[,mspe_2014<5*uk.mspe_2014]
Cex.set <- .75

#Make plot for the permutation analysis
plot(1995:2022,gaps_permutation_2014[,1],
     ylim=c(-18,18),xlab="",
     xlim=c(1995,2022),ylab="Gap in GDP per capita",
     type="l",lwd=2,col="black",
     xaxs="i",yaxs="i")

#Add lines for control units
for (i in 1:ncol(gaps_permutation_controls_2014)) {
  lines(1995:2022,gaps_permutation_controls_2014[,i],col="gray") }

#Add line for the treated unit, i.e., the  UK
lines(1995:2022,gaps_permutation_2014[,1],lwd=2,col="black")

#Add grid
abline(v=2014,lty="dotted",lwd=2)
abline(h=0,lty="dashed",lwd=2)
legend("topleft",legend=c("UK","Control units"),
       lty=c(1,1),col=c("black","gray"),lwd=c(2,1),cex=.8)
arrows(2012.5,-6,2013.5,-6,col="black",length=.1)
text(2010.5,-6,"Time of intervention",cex=Cex.set)



##Random donor pool selection

data_available <- c(1,2,3,4,5,7,9,11,12,13,14,16,17,19,21,22,23,24,25,26,27,28,29,30,31,
                    32,33,34,35,36,38)

#Create dataframe in which all the results of the random donor pool selection will be
#stored. Note that it is not actually a bootstrap but I chose this name as it the idea
```

```
#is similar to that of bootstrap

bootstrap_result <- data.frame(
  Iteration = integer(500),
  Average_Gap_pre_treatment = numeric(500),
  Average_Gap_post_treatment = numeric(500),
  Average_Gap_whole_period = numeric(500),
  MSPE_Loss_V = numeric(500),
  Donor_Pool = character(500)
)

set.seed(123)

year_of_intervention <- 2014

gaps_bootstrap <- matrix(NA, length(1995:2022), 500)

col_counter <- 1

#Perform Synth 500 times for randomly selected donor pools

for(i in 1:500)
{
  donor_pool <- sample(all_available_countries, 25, replace = FALSE)

  dataprep.out <- dataprep(
    foo = df,
    predictors = c("Gross capital formation", "Inflation rate", "Trade openness"),
    special.predictors = list(
      list("Industry rate", 2007:(year_of_intervention-1), "mean"),
      list("Capital account balance", 2007:(year_of_intervention-1), "mean"),
      list("Current account balance", 2007:(year_of_intervention-1), "mean")
    ),
    predictors.op = "mean",
    time.predictors.prior = 1995:(year_of_intervention-1),
    dependent = "GDP pc thousands",
    unit.variable = "Unit",
    unit.names.variable = "Country",
    time.variable = "Year",
    treatment.identifier = 37,
    controls.identifier = donor_pool,
    time.optimize.ssr = 1995:(year_of_intervention-1),
```

```
    time.plot = 1995:2022
  )

  synth.out <- synth(dataprep.out)

  gaps_bootstrap[,col_counter] <- dataprep.out$Y1plot -
    (dataprep.out$Y0plot %*% synth.out$solution.w)

  col_counter <- col_counter + 1

  bootstrap_result[i, "Iteration"] <- i
  bootstrap_result[i, "Donor_Pool"] <- paste(donor_pool, collapse = ",")
  bootstrap_result[i, "Average_Gap_pre_treatment"] <- mean(dataprep.out$Y1plot[1:19] -
                                                             (dataprep.out$Y0plot[1:19,]
                                                              %*% synth.out$solution.w))
  bootstrap_result[i, "Average_Gap_post_treatment"] <- mean(dataprep.out$Y1plot[20:28] -
                                                              (dataprep.out$Y0plot[20:28,]
                                                               %*% synth.out$solution.w))
  bootstrap_result[i, "Average_Gap_whole_period"] <- mean(dataprep.out$Y1plot -
                                                            (dataprep.out$Y0plot
                                                             %*% synth.out$solution.w))
  bootstrap_result[i, "MSPE_Loss_V"] <- synth.out$loss.v
}

#Create a dataframe to store only the gaps of the synthetic control estimators that attain
#an MSPE that is at most 3 times the minimum MSPE

gaps_bootstrap_subset <- matrix(nrow = 28)

for(i in 1:500)
{
  if(bootstrap_result$MSPE_Loss_V[i] < 3*min(bootstrap_result$MSPE_Loss_V))
  {
    gaps_bootstrap_subset <- cbind(gaps_bootstrap_subset, gaps_bootstrap[,i])
  }
}

gaps_bootstrap_subset <- gaps_bootstrap_subset[,-1]

#Plot the results of the random donor pool selection

plot(1995:2022,gaps_bootstrap_subset[,1],
```

```
     ylim=c(-18,18),xlab="",
     xlim=c(1995,2022),ylab="Gap in GDP per capita",
     type="l",lwd=2,col="gray",
     xaxs="i",yaxs="i")

# Add lines for control units
for (i in 2:ncol(gaps_bootstrap_subset)) { lines(1995:2022,gaps_bootstrap_subset[,i],
                                                 col="gray") }

# Add grid
abline(v=2014,lty="dotted",lwd=2)
abline(h=0,lty="dashed",lwd=2)
arrows(2012.5,-6,2013.5,-6,col="black",length=.1)
text(2010.5,-6,"Time of intervention",cex=Cex.set)
```