Staying out of the Euro: Evaluating the Impact of the United Kingdom's Decision on its Economic Growth using the Synthetic Control Method

MPhil in Economic Research

The University of Cambridge

Candidate Number: 5320H

Deadline Date: 25th August 2016

Word limit: 10,000   Actual word count: 9991

I acknowledge that this is my own work, and I have read and understood Section 4 on plagiarism in the Course Handbook.

Abstract:
This paper evaluates the effect of the decision made by the United Kingdom to not join the euro zone in 1999. We use the synthetic control method first used by Abadie and Gardeazabal (2003), and conclude that GDP per capita would not have been significantly different if the United Kingdom had adopted the euro. The results suggest that GDP per capita would have been slightly lower between 1999 and 2008 if it had been adopted, but the difference is not extreme enough to establish confidence in the result. An evaluation of the euro effects for the time period from 2009-2015 is distorted by the financial crisis and the euro debt crisis. However, our analysis indicates that the UK's economic performance is no longer strongly influenced by its euro membership decision after 2008. We also raise concerns about limitations of the synthetic control method when applied to this particular question. Placebo tests indicate that no perfect control unit can be found for the United Kingdom and that there are idiosyncratic shocks other than the euro which distort our analysis.

I am grateful to Dr Solomos Solomou and Maarten de Ridder for their helpful comments.

1 Introduction

The euro is one of the biggest and most controversially debated projects of the European Union and marks a further step towards European integration. Its main benefits are a reduction in transaction costs and the elimination of exchange rate risks, which come at the cost of the loss of independent national monetary policy. Since a path to the creation of the euro had first been formally laid out in the Maastricht treaty in 1992, the question of whether the UK should join the euro zone has been controversially debated in both politics and academia. The economic profession is far from unified on the issue, with economists arguing both for a UK entry ((Layard et al. 2002), (Artis 2000), (Buiter 2008)), as well as against it ((Minford 2008), (Hallett 2002)).[1] This divide highlights the importance of explicitly quantifying the effect that joining the euro would have had on the UK. However, there have only been a few attempts to do so in terms of GDP per capita. This paper addresses this gap in the literature.

We employ the synthetic control method for our analysis and use it to create a counterfactual for the UK. This "artificial UK" is constructed by forming a weighted average of the countries that did join the euro. The results suggest that the UK's GDP per capita would have been 6.5% lower on average, during the time period from 1999-2008. While this effect seems economically significant initially, various robustness and placebo tests reduce our confidence in the estimate. Due to the financial crisis and the euro debt crisis, we exclude the time period from 2009 to 2015 from our analysis. The many shocks and country specific responses during this time period distort our analysis of the euro effects. However, we can still conclude that during this period of crises, the UK's performance no longer seems to be strongly influenced by its decision to not join the euro.

Most early research that tries to quantify the effect of currency unions in general, and the euro zone in particular, looks at the effects on trade, as an increase in trade due to a reduction in transaction costs was one of the key benefits expected from the creation of the euro zone. Rose (1999), using a gravity model of international trade, estimates that entering a currency union has very large positive effects. Two countries that have a common currency will trade up to three times as much as the same two countries if they didn't share a common currency, according to the paper by Rose. This result sparked further research using variations of the gravity model, which produced results that suggested more modest increases in trade (between 3% and 15%) as a result of adopting the euro (Baldwin et al. 2008, Bun & Klaassen 2007, Flam & Nordström 2003).[2]

The following three papers quantify the effect of a UK entry into the euro zone in terms of GDP per capita, but using different methods than this paper. Pesaran et al. (2007) use a global vector autoregressive (GVAR) model. The GVAR model uses data from 1979-2003 to estimate the effects of a UK entry on GDP and prices. They estimate probability distributions of GDP and prices, conditional on observations up to 1999, the year of the hypothetical UK entry. The probability distributions are calculated under two different scenarios. One scenario is the UK not adopting the euro, whilst the other scenario is of the UK adopting the euro. The probability distribution of the UK adopting the euro is calculated by fixing interest rates and exchange rates in relation

[1] The optimum currency area theory, as first introduced by Mundell (1961), strongly influences this discussion. It provides a blueprint, based on which the situation of the UK can be evaluated.

[2] For a more detailed review of the effects on trade, see Baldwin (2006).

to the rest of the euro area at the point of UK entry. They therefore obtain two distributions of what GDP and prices would have been in each case. To obtain estimates of the effect of entering into the euro zone, the difference in probability distributions is calculated. The GVAR model's main strength compared to the synthetic control method is that it allows the entry of the UK to have an effect on the outcomes of the other countries in the euro zone. The synthetic control method on the other hand, assumes that the treatment status of the treated unit is irrelevant for the units that are used as potential control units. While allowing for those types of complex cross country interdependencies, the GVAR model also requires restrictive assumptions. Assumptions have to be made regarding the lag structure, the type of trend in the data, exogeneity/endogeneity of each variable for each country, linearity of the model and the mechanisms with which shocks are transmitted. The GVAR analysis also never uses actual data on the UK after 1999 to evaluate the hypothetical scenario of the UK adopting the euro, or to test the predictive power of the model. A further assumption has to be made about the exchange rate between the pound and the euro zone at the point of the hypothetical entry of the UK. The results obtained by Pesaran et al. are sensitive to different levels at which the exchange rate is fixed. Their baseline assumption is that exchange rates are fixed at the level observed in the last quarter of 1998. They estimate that an entry of the UK into the euro zone would have been likely to lower its GDP per capita initially and raise it in the medium term (between 2000 and 2003). However, the estimated probabilities do not allow for clear conclusions. A further assessment of the effects of a hypothetical UK entry into the euro zone in 1999, has been undertaken by HM Treasury in 2003. The report by HM Treasury uses the National Institute Global Econometric model (NiGEM) by the National Institute of Economic and Social Research (NIESR) for the quantitative analysis. The NiGEM is an economic model based on a New Keynesian framework and a Dynamic Stochastic General Equilibrium model (DSGEM). It models each major country and the components of its economy separately (demand side, supply side, investment, unemployment,...). Details regarding the model are available on the NIESR homepage. The key assumption of the NiGEM, as identified in the report by HM Treasury, is that a change in the monetary policy regime cannot alter any outcomes in the long run. Due to this assumption, GDP per capita will be the same in the long run, whether the UK joins the euro zone or not. Despite this assumption, short run effects on GDP per capita can still be possible because prices and wages are assumed to be sticky. The report presents results on several economic indicators including output growth. It estimates that if the UK would have adopted the euro in 1999, output growth would have been lower on average between 2001 and 2007. However, the results are reported under the caveat that they are "likely to be surrounded by a wide margin of uncertainty" (HM Treasury 2003, p.43). Mazumder and Pahl (2013) provide a third approach to quantify the effect on the UK of staying out of the euro. Their evaluation period is from 1999 to 2010. Counterfactuals on what the outcomes would have been if the UK had joined the euro are obtained using variations of the Phillips curve. They obtain estimates of the parameters of the Phillips curve using data on the UK between 1999 and 2010. The counterfactuals are calculated using the estimated coefficients of the Phillips curve and a proxy on what inflation would have been in the scenario of a UK entry. Their results suggest that the UK would have had a lower output if it had adopted the euro. Again, this result is consistent with the result we obtain when using the synthetic control method. However, one needs to be aware that the estimation of counterfactuals using the Phillips curve rests on very strong assumptions and simplifications that reduce the confidence in the results. To obtain a proxy for what inflation would be in the UK if it had joined the euro zone, assumptions about purchasing power parity between the euro area and the UK have to be made. Furthermore, assumptions are made about the natural level of output, unemployment, and the level of productivity growth in the hypothetical scenario of the UK adopting the euro. The fact that expectations for inflation are not observable when estimating the New Keynesian Phillips curve introduces further room for errors, because an instrument has to be used. Additionally, the series of counterfactuals reported has been smoothed using the Hodrick-Prescott filter, which might artificially introduce features into the data. A much more general concern also arises when considering the Lucas critique. It might be problematic to use parameters estimated using real UK data to obtain counterfactuals for the hypothetical world, in which the UK did join the euro. Optimising agents might make different choices given the policy change, therefore also potentially changing the parameters of the Phillips curve.

This paper adds to the existing literature that uses GDP per capita as the outcome variable of interest. The motivation for using GDP per capita and not trade is rooted in the perspective we adopt for our analysis. It is that of a UK policy maker who is interested in finding out whether the UK made the right decision when it chose to not adopt the euro in 1999. GDP per capita, a measure of economic performance, is closely linked to the welfare of a country, which should ultimately be of highest interest to a policy maker. It is also important to be aware of the limitations we face as a result of choosing GDP per capita as the outcome variable. By making this choice, we solely focus on the economic consequences of the euro, which is mainly a political project that has to be seen in the context of European integration (Wyplosz 2006). Nonetheless, evaluating the euro's economic consequences is a crucial task, since they lie at the heart of the question whether the euro will prevail as the key European currency. The political ideals of the euro might be noble, but if membership in the euro zone is not connected with economic benefits for its member countries, the euro project is unlikely to survive due to the loss of public support. Therefore, economic analysis such as done in this paper should guide any country's decision on whether to join the euro zone.

The results obtained by our analysis are similar in nature to the ones obtained using different methods. Since the synthetic control method is a fundamentally different approach than the three outlined above, it has very different limitations and assumptions. It is thus even more remarkable that all the methods produce fairly consistent results.

The rest of this paper is structured as follows. Section 2 explains the synthetic control method and section 3 applies it to our question. In section 4, we apply several robustness tests to our results. Section 5 analyses the effects of the euro during the financial crisis, and section 6 addresses limitations of the methodology. Section 7 concludes.

2 The synthetic control method

This section explains the intuition and the functioning of the synthetic control method and is based on the paper by Abadie and Gardeazabal (2003) that introduced the method. The synthetic control method was first used to estimate the effect of terrorism on GDP per capita in the Basque country (Abadie & Gardeazabal 2003) and has since been used to analyse a wide range of topics.[3] This paper closely follows the approach used by Abadie et al. (2015) in the way the synthetic control method is applied and it also uses the terminology established by Abadie et al. (2015). All analysis is done using the statistical software R and the R package "Synth" developed by Abadie et al. (2011). These papers can also be consulted for a more extensive description of the methodology.

The basic idea behind the synthetic control method is to make policy evaluation possible in the case where there is no obvious control unit. We have one treated unit $i = 1$ and several non-treated units $i = 2, 3, ..., n$. At time $t = T$, unit $i = 1$ is subjected to a treatment. We are interested in the effect of the treatment on our outcome variable Z. The synthetic control method proposes to create an artificial control unit which resembles the treated unit as closely as possible. This artificial control unit is created by forming a weighted average of the untreated units (the donor pool). Let us denote the vector of pre-treatment observation for the treated variable $i = 1$ by Z_1, and the matrix of pre-treatment observations for the units in the donor pool $i \neq 1$ by Z_0. They have the following dimensions:

$$Z_1 : \Big((T-1) \times 1\Big) \text{ and } Z_0 : \Big((T-1) \times (n-1)\Big)$$

We have observations for the following time periods:

$$t = 1, 2, 3..., T-1, T, T+1, ..., T+j$$

We calculate a vector of country weights

$$w^T = \begin{bmatrix} w_2 & w_3 & w_4 & ... & w_n \end{bmatrix},$$

such that we minimise the squared difference between Z_1 and $w \times Z_0$ over the pre-treatment period. More formally we choose w such that the following is satisfied:

$$\min_w \Big((Z_1 - Z_0 \cdot w)^T \times (Z_1 - Z_0 \cdot w)\Big). \tag{1}$$

We match the outcome variable of the treated unit as closely as possible with the artificial control unit during the pre-treatment period. This simple example, where we only match the outcome variable, captures the basic idea of the synthetic control method.

It is easily expanded to allow for more complex and thorough pre-treatment matching of the treated unit and the synthetic control unit. Instead of only trying to match the outcome itself, there might be further variables that are good predictors of the future development of Z. For example, if as in our case Z is GDP per capita, we would not only want the synthetic control unit and the treated unit to be very similar in the pre-treatment GDP per capita path, but also regarding certain other characteristics. Only if the treated unit and the synthetic control unit are comparable in terms

[3]For further applications of the synthetic control method see Abadie et al. (2010, 2015), Acemoglu et al. (2013), Campos et al. (2014) and Fernández & Perea (2015).

of investment, schooling and other growth predictors, can we be confident that the synthetic control unit has the same growth potential as the treated unit at time T. If they are comparable, we can be confident that they would follow the same growth path in the absence of the treatment. Ideally, we would want the treated unit and the synthetic control unit to be exactly identical except for the treatment. If this was the case, any difference in outcomes in the post-treatment periods must be due to the treatment. The synthetic control method is based on the assumption that if we are able to match the path of the outcome variable for a sufficiently long time period before the treatment, then the treated unit and the synthetic control unit "are alike in both observed and unobserved determinants of the outcome variable as well as in the effect of those determinants on the outcome variable" (Abadie et al. 2015, p.498). Hence, if the treated country and the synthetic control unit are very close matches, we can be confident that any divergence is due to the treatment.

We choose k predictors for the outcome variable Z in order to ensure that the treated unit and the synthetic control unit are comparable. The outcome variable can be part of the predictors. We denote the matrix of observations of pre-treatment predictors by X_1 and X_0 for the treated unit and the donor pool respectively. X_1 and X_0 thus have the following dimensions:

$$X_1 : \Big(k \times 1\Big) \text{ and } X_0 : \Big(k \times (n-1)\Big)$$

We also introduce predictor weights

$$v^T = \begin{bmatrix} v_1 & v_2 & v_3 & \ldots & v_k \end{bmatrix}.$$

They determine how important each individual predictor is. A higher predictor weight means that we care very much about the synthetic control unit and the treated unit being similar in this characteristic. We therefore choose w as follows:

$$\min_{w} \Big((X_1 - X_0 \cdot w) \times v^T \times (X_1 - X_0 \cdot w) \Big). \tag{2}$$

Both vector weights and predictor weights are chosen so that they sum to one. However, the result wouldn't change if they were all scaled by the same factor. The country weights (w) depend on the values chosen as predictor weights (v). The choice of predictor weights is thus crucial, as it determines the country weights and therefore the performance of the synthetic control unit in the post-treatment period.

The vector of predictor weights can be chosen in different ways. They can be chosen manually, if one wants to have specific weights for each predictor. If they are chosen manually, a strong economic argument has to be made to justify the choice of predictor weights. Alternatively, one can simply choose the predictor weights to obtain the best possible fit of the outcome variable between the treated unit and the synthetic control unit in the pre-treatment period. This means we choose the predictor weights in order to satisfy

$$\min_{v} \Big((Z_1 - Z_0 \cdot w(v))^T \times (Z_1 - Z_0 \cdot w(v)) \Big). \tag{3}$$

We solve equation (3), given the country weights as a function of the predictor weights. The country weights as a function of the predictor weights are obtained from (2). This method is called the "eclectic method" (Abadie & Gardeazabal 2003). A further possibility used by Abadie et al. (2015), is to divide the pre-treatment period into two. We can then choose the predictor weights based on the first sub-period, in order to minimise losses when forecasting the outcome variable for the later sub-period. The latter two methods are strictly data driven, whereas the first method relies on an economic argument being made.

3 Data and main results

The question this paper aims to answer is what would the GDP per capita of the UK have been if it had adopted the euro in 1999? Since none of the countries that did adopt the euro is perfectly comparable to the UK, we do not have an obvious counterfactual for our analysis. The synthetic control method therefore is an attractive approach to address our question, because it provides a solution to the problem of not having a perfect control unit. The treatment in this particular scenario is "not adopting the euro". Hence, the donor pool consists of countries that did adopt the euro. In 1999, the following countries fixed the exchange rates of their national currencies to the euro: Austria, Belgium, Finland, France, Germany, Ireland, Italy, Luxembourg, Monaco, Netherlands, Portugal, San Marino, Spain and Vatican. Due to their very specific economies, Luxembourg, Monaco, San Marino and Vatican City are omitted. The countries included in the donor pool need to be comparable to the treated unit. Any countries that display idiosyncratic behaviour or are subjected to severe aggregate idiosyncratic shocks during the post-treatment period need to be removed from the donor pool. Otherwise, their idiosyncratic post-treatment characteristics would distort the analysis. We exclude Ireland from the sample since Ireland's economy grew extremely fast due to country specific circumstances until the early 2000's (Celtic Tiger) (Abadie et al. 2015). It would therefore distort the analysis in the post-treatment period if it was part of the synthetic control group. The remaining nine countries form the donor pool.

The dataset used for the analysis contains data on GDP per capita from 1970-2015. This means that there are 29 years of pre-treatment data to build the synthetic control unit and 17 years of post-treatment data to evaluate the performance of the synthetic control unit compared with the actual UK data. The financial crisis and the euro debt crisis occur during the evaluation period, which poses particular challenges for our evaluation of the euro effects. We need to be confident that the synthetic control unit and the UK would have displayed a similar GDP per capita path during the crises period in the absence of the euro. To be confident that this is the case, we include a wide set of predictor variables in our analysis. Besides the standard growth predictors (gross capital formation, population growth, trade openness, inflation, size of industrial sector, schooling), we include domestic credit to the private sector by banks (% of GDP) and the size of the service sector among our predictor variables.[4] The idea behind including a measure of credit and the size of the service sector among our predictor variables is that both should give us an indication of how severely a country will be affected by the crises. Therefore, if the synthetic control unit and the UK are very

[4]More information on the data is given in the appendix.

closely matched based on these two variables, we would expect that any deviation in GDP per capita between 2008 and 2015 is due to the euro and not any other special country characteristics. However, robustness tests show that even when including these predictors, the GDP per capita path of the synthetic control unit strongly relies on individual countries. This suggests that there are further determinants of a country's performance during the crises that aren't captured by the predictors. These could be idiosyncratic responses of the national governments which we cannot capture with the predictors. The financial crisis and the euro debt crisis therefore distort the analysis after 2008, and the synthetic control method can only be used reliably before this time period. This leaves 10 years of post-treatment data for evaluation. Since we exclude the crises period from our data, the additional predictors do not add any additional insight and we return to the standard growth predictors, similar to the ones used by Abadie et al. (2003).

Using the eclectic method, we calculate the country weights (w) and predictor weights (v), as shown in table 1 and 2. Italy, Finland and France obtain the biggest weights, which is not surprising since GDP per capita in those countries evolves similarly to the UK during the pre-treatment period. The predictor weights seem rather peculiar, since the industry share in GDP receives nearly the entire weight. However, we have chosen the eclectic method to choose the predictor weights. When using the eclectic method, the predictor weights are chosen with the sole aim of allowing for the closest possible fit of the outcome variable during the pre-treatment period. They are therefore mainly reported for completeness, as they help us understand how the synthetic control unit is constructed. The synthetic control unit and the treated unit (UK) are compared in table 3 in terms of their pre-treatment characteristics. If we are unsatisfied with how closely a certain predictor is matched, the table of predictor weights allows us to determine if this was due to the predictor receiving a low weight. This can then guide our decision on whether we want to manually select a higher predictor weight for a specific variable, thereby forcing the synthetic control method to closely match the selected predictor. In our case, schooling receives a very low weight (approximately zero), which explains why schooling is the predictor where the treated unit and the synthetic unit differ the most. The fact that schooling obtained a predictor weight of $v \approx 0$ implies that if we would like to achieve a better fit in schooling, this would compromise the fit in the GDP per capita path. Given this trade-off, we do not assign a higher weight to schooling manually, but prioritise the pre-treatment fit in the outcome variable.

Table 1: Country weights

Country	**Weight** (w)
Austria	0
Belgium	0.002
Finland	0.272
France	0.147
Germany	0.002
Italy	0.553
Netherlands	0.022
Portugal	0
Spain	0.001

Table 2: Predictor weights

Predictor	**Weight** (v)
GDP per capita	0.008
Capital formation	0.035
Population growth	0.028
Trade	0
Inflation	0
Industry	0.928
Schooling	0.001

Table 3 shows that the synthetic control unit is very similar to the UK (treated unit), and a better counterfactual than a simple sample average would be. We have chosen to match the predictors based on their mean value in the 10 years prior to treatment, where sufficient data was available. It is important to note that these predictors are merely additional predictor in the sense that the synthetic control method still primarily matches the GDP path of the UK for every individual pre-treatment year, according to equation 3. The average GDP per capita for the time period from 1989-1998 has been included as an additional predictor, to emphasise the importance of the treated unit and the synthetic control unit being comparable in terms of the outcome variable in the years directly prior to the treatment. For predictors where data isn't available for the entire 10 years before treatment, a shorter time period has been chosen. Observations older than 10 years before treatment are becoming less and less relevant and have not been included as additional predictors.

Table 3: Pre-treatment fit

	Treated	**Synthetic**	**Sample Mean**
GDP per capita	20132.193	20157.819	19546.732
Capital formation	20.014	21.285	23.313
Population growth	0.269	0.213	0.352
Trade	49.618	46.880	64.826
Inflation	2.098	2.468	2.251
Industry	29.258	29.241	29.411
Schooling	77.400	60.110	60.911

GDP per capita, capital formation, population growth and trade are averaged from 1989-1998, inflation is averaged from 1994-1998, industry from 1995-1998 and schooling is only matched in the year 1999 due to data availability.

Figure 1 captures the key results of our analysis. It graphs the GDP per capita path for the UK and the synthetic control unit. We can see that the synthetic control unit closely matches the UK before the treatment of not adopting the euro in 1999.[5] In 1999, the UK starts performing better than the synthetic control unit, which suggests that the UK would have had a lower GDP per capita if it had adopted the euro. On average, GDP per capita would have been lower by 6.5% between 1999 and 2008 if the UK had adopted the euro according to our analysis. This implies that for the UK, the costs of the euro would have outweighed its benefits during this time period. Having an independent monetary policy thus seems to be particularly important for the UK economy. From a purely economic perspective, the UK therefore made the right decision when it decided to not join the euro. Towards the end of the evaluation period, the GDP paths converge again, which illustrates that the conclusion that the UK would have been worse off with the euro is not a final verdict, but merely holds for our evaluation period. While the estimated effects are economically significant, placebo tests greatly reduce our confidence in the results.

[5] All figures in this report have been created using R and use the same formatting as the figures created in the R-Code made available by Abadie et al. (2015).

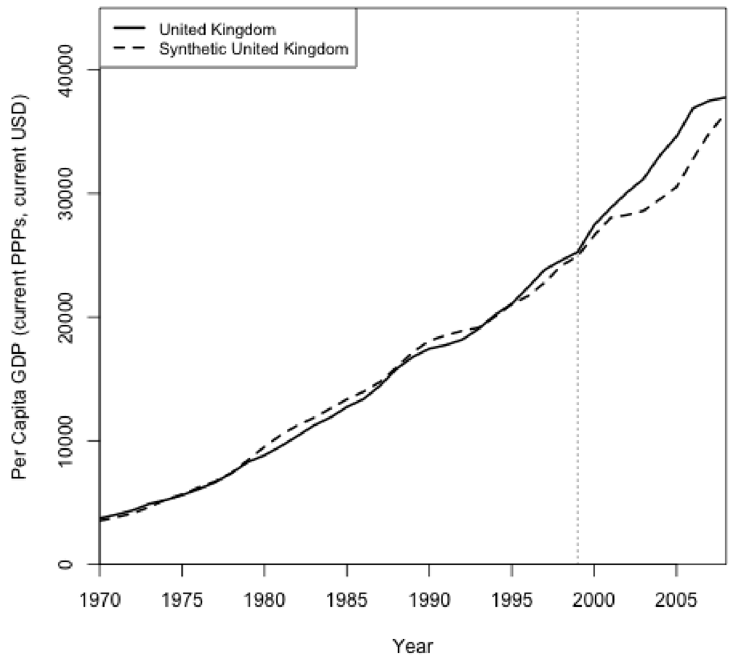

Figure 1: UK and Synthetic UK

4 Robustness tests

The synthetic control method provides the possibility to perform a variety of robustness tests to establish confidence in the results. We will test how sensitive the results are to changes in the predictor weights and how strongly they rely on specific countries. Furthermore, we can also perform placebo tests, which are similar in nature to the ones performed for difference in difference analysis.

4.1 Choosing predictor weights

The industry share in total GDP receives a predictor weighting very close to 1 when using the eclectic method. This seems problematic, but it should be remembered that the predictor weights are only selected in order to obtain the closest possible fit between the UK and the synthetic control method in the pre-treatment period. This remains our main objective, as we have argued that being able to closely match the outcome variable over an extended period of time is the crucial criteria we need to satisfy in order to have confidence in our results. Furthermore, the results do not change significantly if a different method is chosen. Running the synthetic control method with predictor weights selected manually, we calculate the following corresponding country weights.

Table 4: Country weights

Country	**Weight** (w)
Austria	0
Belgium	0.001
Finland	0.330
France	0.291
Germany	0
Italy	0.376
Netherlands	0.002
Portugal	0
Spain	0

Table 5: Predictor weights

Predictor	**Weight** (v)
GDP per capita (1970)	1/13
GDP per capita (1975)	1/13
GDP per capita (1980)	1/13
GDP per capita (1985)	1/13
GDP per capita (1990)	1/13
GDP per capita (1995)	1/13
GDP per capita (1998)	1/13
Capital formation	1/13
Population growth	1/13
Trade	1/13
Inflation	1/13
Industry	1/13
Schooling	1/13

We choose GDP per capita values in five year intervals over the pre-treatment period as additional predictors. It is necessary to include individual data points of the outcome variable, since we now choose country weights using only equation 2 from section 2. Equation 2 minimises the squared difference between the predictors (matrix X) of the UK and the synthetic control unit. Unless we specifically include the outcome variable as part of the predictors, it is not considered for the optimisation. For the eclectic method on the other hand, we also use equation 3, which optimises over the outcome variable (matrix Z). We have distributed the weights equally amongst the predictors, since there is no convincing reason why any particular one should receive a higher weight. There could be various arguments made in favour of an alternative distribution of the predictor weights, which illustrates the weakness of this approach: there is no clearly correct way to manually select the predictor weights. It is ultimately an arbitrary distribution that can be justified with strong economic intuition at best. The synthetic control method still chooses Italy, Finland and France as the closest matches. This illustrates that in this case, there is no significant difference between the eclectic approach and the manual approach above. As both approaches yield similar results, we prefer the eclectic method as it is strictly data driven.

A further method to determine the predictor weights is the cross validation method. It is the method that puts the largest requirements on the data. We need data on all predictors to be available for both sub-periods into which the pre-treatment period is divided. Several predictors had to be dropped due to data availability when using the cross validation approach. This might have caused the slightly bigger differences in the country weights. However, the two countries that receive the biggest weights are France and Italy, which are also part of the synthetic control unit when using the previous approaches. As a result, the GDP per capita path of the synthetic UK would have been extremely similar to the other two approaches and the answer to our main question wouldn't have changed.

4.2 Reliance on individual countries

Another important question is how much the result relies on any individual country. In our case, we would be particularly interested to see how the result changes if we exclude one of the countries that received the biggest country weights (France, Italy or Finland) from the donor pool. Figure 2 illustrates how the result would change if we left out any of the countries that received a positive weight.

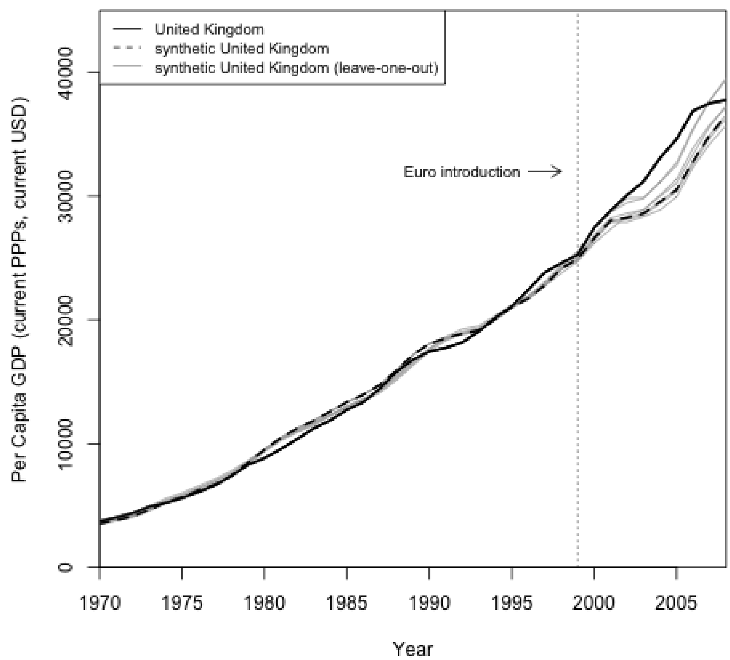

Figure 2: Excluding countries from donor pool

The top line of the "leave-one-out lines" is in fact two lines lying above of each other. They belong to the synthetic control unit when leaving out Italy or France respectively. Leaving out any of the other countries does not change the result noticeably. However, leaving out either France or Italy would lead us to reduce our estimate of the losses the UK would have experienced if it had joined the euro. We can therefore interpret our result as an upper bound of the losses. Even though the result changes slightly when leaving out France or Italy, there are no particularly severe idiosyncratic shocks to the two countries in the post-treatment period that would suggest that they should be excluded from the donor pool.

4.3 In-time placebos

In-time placebos are used to test whether there is an effect at a time period where there was no treatment. If this was the case, we could no longer be confident that our method is indeed picking up the effect of the euro. In-time placebo tests are particularly important for the question of what the effect of the euro would have been on the UK, since there are several dates that could potentially serve as the treatment period.

Intuitively, one might be inclined to use 2002 as the treatment date, as this is the date the euro was physically introduced, and the year the wider public generally associates with the beginning of the euro era. However, as can be seen in figure 1, the GDP paths of the synthetic control unit and the UK already start diverging before 2002. Therefore, 2002 is not the appropriate treatment period, since the euro being introduced as a virtual currency already had an effect on the euro zone countries. Further possible dates after which there might have already been anticipation effects are 1992 and 1998. The UK achieved the inclusion of an opt-out clause from the euro zone in the Maastricht Treaty in February 1992, which made it less likely for the UK to adopt the euro along with the other EU states in 1999. In October 1997, Tony Blair exercised the opt-out clause which ruled out that the UK would join the euro zone in 1999. Therefore, there might be anticipation effects as early as 1992. Figures 3 and 4 show the results of the in-time placebo tests for 1992 and 1998 as the treatment period.

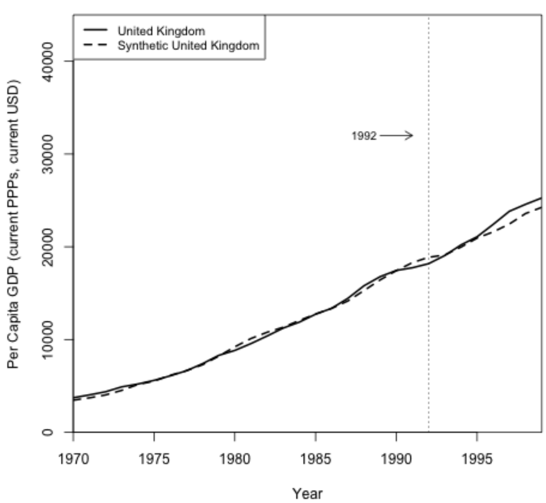

Figure 3: Maastricht Treaty (1992)

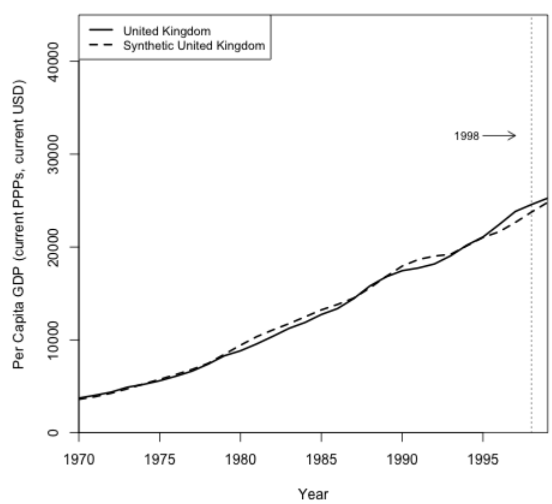

Figure 4: Use of the opt-out clause (1998)

We are interested in the fit between the synthetic control unit and the UK in the time period between the selected treatment date and the year 1999, which was the year selected as the treatment date originally. Since the synthetic control unit matches the GDP per capita path of the UK very closely in both cases, we conclude that there were no anticipation effects after 1992 or 1998.

In-time placebo tests for further dates (20, 15 and 10 years before 1999) also closely match the GDP path of the UK in the pre-treatment period.[6] The outcome variable of the treated unit thus

[6]We excluded Finland from the donor pool for all in time-placebos that had the treatment period before 1994, since Finland was subject to idiosyncratic shocks (banking crisis) in the early 1990s (Abadie et al. 2015, Drees & Pazarbasioglu 1998).

does not show any changes at time periods where there was no treatment, which makes us confident that the appropriate treatment period is 1999, the year the national currencies were fixed to the euro.

4.4 In-space placebos

In-space placebos are another type of placebo test that can be performed to find out how robust the results are. The in-space placebo tests turn out to be the crucial placebo tests for the analysis of this paper, as they challenge our confidence in the significance of the results obtained. The idea behind the in-space placebo test is to see if there are any changes in the outcome variable in the treatment period for a unit that did not receive treatment. We would like to see no deviation of the GDP path after the introduction of the euro between the synthetic control unit and the actual country for the placebo analysis. However, we cannot simply evaluate the in-space placebo based on the post-treatment period alone. A big deviation in the GDP path after the introduction of the euro can only be seen as problematic if the synthetic control unit was able to closely match the actual GDP path before the treatment period. Post-treatment deviations are not surprising if there were already big deviations in the pre-treatment period. We thus use the concept of the root mean square prediction error (RMSPE) to evaluate in-space placebo tests (Abadie et al. 2015).[7]

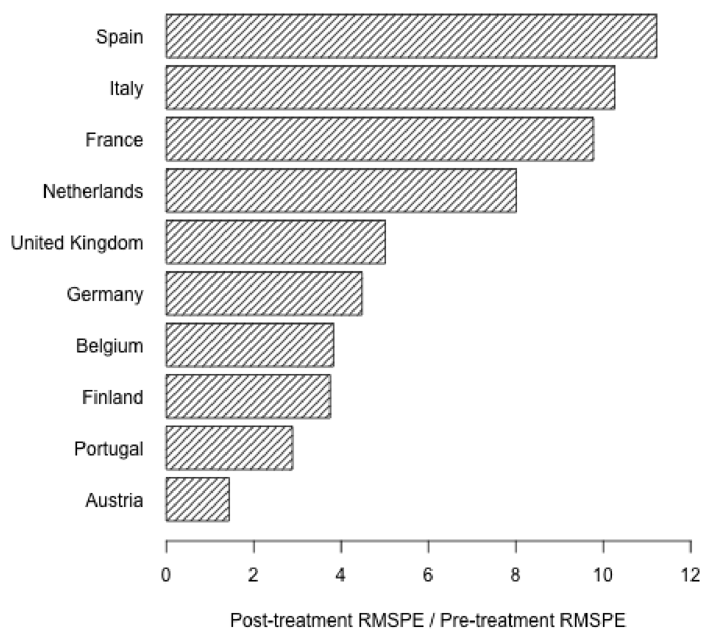

Figure 5: RMSPE Plot

[7]The pre-treatment RMSPE is calculated as $RMSP_{pre} = \left(\frac{1}{T-1} \sum_{t=1}^{t=T-1} (Z_1 - Z_0 \cdot w)^2 \right)^{\frac{1}{2}}$. The post-treatment RMSPE is calculated equivalently (Abadie et al. 2015).

The RMSPE ratio is calculated as ${}^{RMSPE_{post}}/_{RMSPE_{pre}}$, where post and pre stand for post-treatment period and pre-treatment period respectively. The smaller the RMSPE ratio of the placebo tests is in comparison to the one when using the UK as the treated unit, the more confident we are that our estimated treatment effect is significant. Figure 5 illustrates the RMSPE ratios for all of the countries. It allows us to determine the probability of obtaining a result at least as extreme as the one for the UK when choosing a country at random.

The in-space placebo analysis severely reduces the confidence in our estimate. Four countries achieve higher RMSPE ratios than we obtain when using the UK as the treated unit. This suggests that our estimate that GDP per capita would have been 6.5% lower is not very significant. There is a 50% chance of obtaining an RMSPE ratio at least has high as the one for the UK when picking a country at random from our sample. We would have liked the UK to have the highest RMSPE ratio to establish confidence in our estimate.

A high RMSPE ratio can be the result of two factors: an extremely good fit in the pre-treatment period and/or a very large deviation in the post-treatment period. Equivalently, a low RMSPE ratio is due to a bad pre-treatment fit and/or a small post-treatment deviation. We can look at the two countries with the highest RMSPE ratios to get a better understanding of why the UK's RMSPE ratio is merely average. Figures 6 and 7 show the results of the in-space placebo tests for Spain and Italy, the two countries with the highest RMSPE ratios.

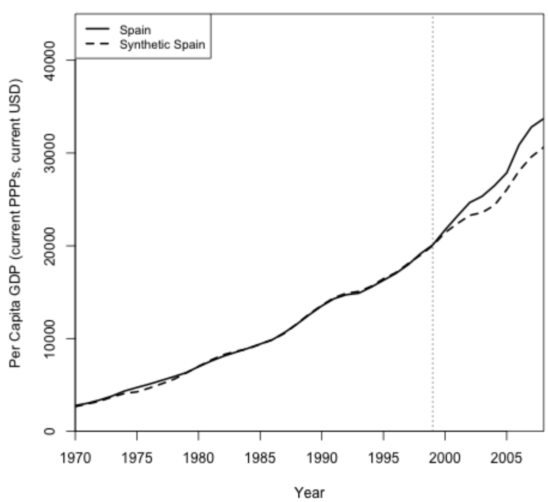

Figure 6: In-space placebo: Spain

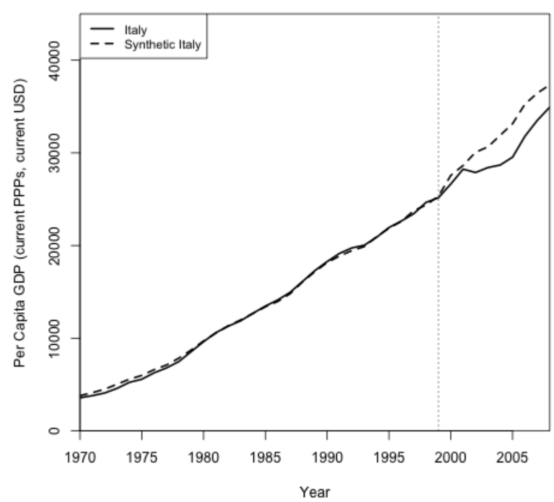

Figure 7: In-space placebo: Italy

Looking at the GDP per capita paths for Spain and Italy, as well as that of their synthetic controls, it becomes clear that their high RMSPE ratios are due to two reason: a better pre-treatment fit than the fit that was achieved for the UK, and deviations in the outcome variable comparable in size to those observed for the UK. It therefore follows that there are two reasons why the UK's RMSPE ratio is not the largest one.

Firstly, it is due to the fit between the synthetic control unit and the treated unit achieved during the pre-treatment period. While the fit achieved in the pre-treatment period appears to be satisfactory when looking at the UK in isolation, we realise that a better fit can be achieved for the other countries in the donor pool when performing the placebo tests. The reason for this might be that the UK's economy is somewhat different from all continental European economies. Therefore, a perfect counterfactual cannot be constructed using the synthetic control method. Some special characteristics of the UK economy have been controlled for, like the importance of the service sector. This has been controlled for implicitly, by making sure the synthetic control unit is comparable in terms of the industry share in GDP. Since a country with a large service sector is likely to have a relatively small industry sector, it is sufficient to control for only one of them (adding the service sector as an additional predictor did not change the results). However, there might be other characteristics that cannot be controlled for, such as the UK's special relationship to the US as well as the Commonwealth, and the resulting exposure to shocks from these parts of the world. These are merely guesses and there could be multiple reasons why we cannot find as good a match for the UK as we can for the continental European countries.

Secondly, the relatively small RMSPE ratio is due to the deviation not being very pronounced between the synthetic control unit and the treated unit in the post-treatment period. The deviation observed for the UK is not much bigger than the one observed for Spain and Italy, both countries where we would expect to see no deviation at all! This lets us draw two conclusions.

The first conclusion can be drawn from the observation that there is a change in the outcome variable after the treatment period for both Spain and Italy, which goes in opposite directions for the two countries. This could suggest that there is no homogeneous euro effect across all countries in the sample, even when controlling for the standard economic growth predictors. Figure 6 suggests that Spain benefitted more than the average country in the sample. Italy on the other hand, benefitted less from the euro than the average country. The other possible conclusion is that not adopting the euro simply wasn't a decision with big consequences for GDP per capita in the period from 1999-2008. The observed deviation is not big enough in order for us to say with confidence that there was a significant effect at all. This interpretation of the in-space placebo tests thus suggests that GDP per capita in the UK would not have been significantly different if the UK had joined the euro. The decision not to join therefore didn't have a significant impact on its economic growth. This is an important insight for policy makers in the UK and for other countries that are considering joining the euro zone. It suggests that the economic benefits from the euro due to trade effects and a reduction in transaction costs are not significant for the UK and might even be outweighed by the costs, with the main cost being the loss of independent monetary policy. Since the economic effects of the euro are not clear, the case for the euro has to be made using political arguments. It is worth noting that the result is only useful for policy makers from other countries in the sense that the effect on GDP per capita from joining the euro is not unequivocally positive. The result does not suggest that the outcome for other countries would be the same as for the UK, but rather highlights the importance of evaluating the effects for each country individually.

5 Evaluation post 2008

This section looks at the period from 2009-2015, which was excluded from the previous analysis. When looking at the time period until 2008, figure 2 showed that the results do not change significantly if any single country is excluded from the donor pool. This changes when the evaluation period is extended until 2015, which is illustrated in figure 8. The spread of the GDP per capita paths of the synthetic control units with one country left out increases significantly after 2008. This suggests that the countries in the euro zone were not affected uniformly by the financial crisis and the euro debt crisis, and that country specific measures to fight the crisis had varying success. It indicates that countries still had a certain flexibility to react to the crises despite not having independent monetary policy. Otherwise, a more uniform picture should have emerged during the period from 2008-2015. Therefore, the response of GDP per capita to the crisis and the recovery profile of a country does not just depend on whether a country is in the euro zone or not, but on other country specific measures like a country's policy response. These idiosyncratic shocks distort our analysis of the effects related to the euro. They are the reason why we cannot be confident in finding a synthetic control unit for the UK from 2008-2015 that only captures the effect of the euro. We have therefore excluded this time period from our main analysis.

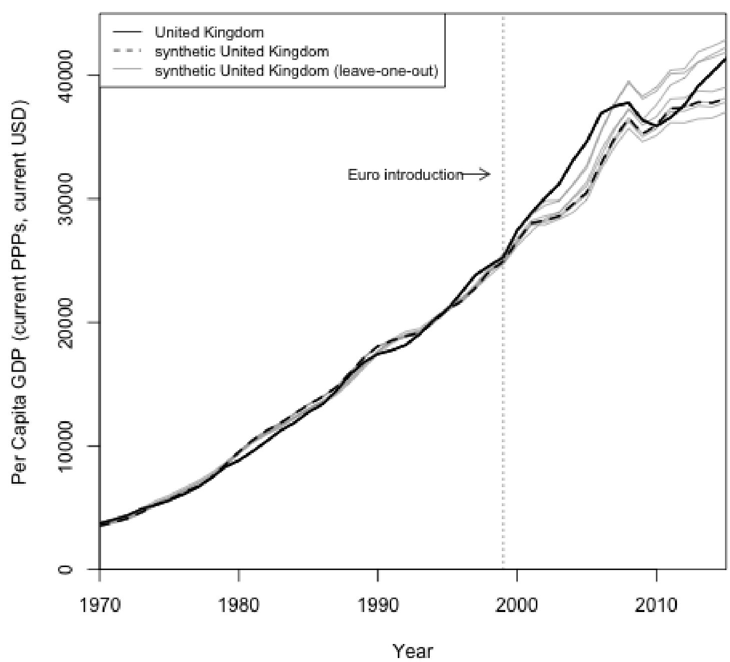

Figure 8: Excluding countries from donor pool (same as figure 2 with evaluation period extended to 2015)

6 Limitation of the synthetic control method

The synthetic control method is a useful tool to perform policy analysis when there isn't an obvious counterfactual. The idea behind the synthetic control method is easy to understand intuitively. However, it is important to be aware of its limitations as revealed by the analysis undertaken for this paper.

A key weakness of the synthetic control method becomes apparent when performing the in-space placebo tests for Spain and Italy. We observe a change in the outcome variable after the treatment period, even though we wouldn't expect to see one. This draws our attention to the limits of the statement that if the synthetic control unit perfectly matches the treated unit during the pre-treatment period, they "are alike in both observed and unobserved determinants of the outcome variable as well as in the effect of those determinants on the outcome variable" (Abadie et al. 2015, p.498). The pre-treatment fit is extremely close over 29 years for both Spain and Italy, and there are no particularly severe idiosyncratic shocks to any of the countries in the donor pool in the post-treatment period that could provide a convincing reason for the deviation after 1999. If the statement would hold true, we should not see a deviation after 1999 for either Spain or Italy. But the statement only holds true if the shocks and determinants of the outcome variable are similar in the post-treatment period and the pre-treatment period. The synthetic control method cannot create a control unit that is similar to the treated unit in its response to shocks for which there was no precedent in the pre-treatment period. This problem of not finding a convincing synthetic control unit is particularly likely to occur if the treatment is defined as "not being treated", as is the case for our analysis. In this case, all countries in the donor pool are subjected to some treatment. If there was no similar treatment in the period based on which the match is found that affected all countries, it is difficult to argue that the synthetic control unit behaves similarly to the unit of interest if it would receive the treatment. This problem occurs in our scenario. During our pre-treatment period (1970-1998), there have not been any events comparable to the creation of the euro. We therefore cannot be sure that the combination of countries chosen as the synthetic control unit reacts to the introduction of the euro in the same way the UK would react. Similarly, there have been some unprecedented shocks in the evaluation period, like the dot-com bubble. Since this shock is strongly tied to technologies which have not been used in the pre-treatment period, it is difficult to argue that the synthetic control unit's response to it is the same as that of the UK. With every additional year included in the evaluation period, and with every additional country added to the donor pool, the likelihood of shocks that are different in nature to the ones in the pre-treatment period increases. This puts a limit on the evaluation horizon. It suggests that the situations in which the synthetic control method can be used with confidence are actually rather limited. The condition of similar shocks has to be satisfied in addition to the condition of no idiosyncratic shocks/characteristics for countries in the donor pool during the evaluation period. We discussed this condition when preparing the dataset, and it was the reason we excluded certain countries from the donor pool.

A further condition to obtain valid estimates using the synthetic control method is that the treatment status of the treated unit has no effect on the countries in the donor pool. This means that we assume that the performance of the countries in the euro zone would not have been different if the UK had joined. It is not possible to test this assumption, but due to the overall size of the euro zone, with all major European economies except the UK being among the first euro members, a strong case can be made that this assumption holds. However, arguments for a different case could be made as well. If one argues that the countries in the donor pool would have benefitted from the UK adopting the euro, the estimate of the losses the UK would have made when entering the euro are an upper bound. Pesaran et al. (2007) address the same question as this paper using a GVAR model that allows for the type of interdependencies described above. They obtain similar results in the sense that they do not obtain results that allow for unambiguous conclusions. This suggests that the assumption made for the synthetic control method does not distort results significantly.

The synthetic control method also constraints country weights to be non-negative. This creates problems when the treated unit has the most extreme values either for the outcome variable or the predictors. The outcome values and predictor values of the synthetic control unit are constrained by the range of values given by countries in the donor pool. Thus, a match for a country cannot be found if it has the most extreme characteristics out of all the countries in the sample. This is also one of the key differences to the standard regression analysis, where extrapolation beyond the values given in the sample is possible. A more detailed discussion of this limitation and a comparison of the synthetic control method to the regression analysis is provided by Abadie et al. (2015). It is worth noting that this constraint imposed by the synthetic control method is also one of its strengths, as it forces researchers to carefully pick the donor pool to ensure that countries are comparable.

Another limitation arises due to the usage of the eclectic method to select the predictor weights (v). If a country from the donor pool matches the treated unit well in terms of the outcome variable, but is different in terms of one of the predictor variables, the eclectic method will simply assign a vector weight of $v = 0$ to that specific predictor. This is not problematic if the synthetic control unit still provides a good match based on the additional predictors (as is the case for the analysis done for this paper), but it can potentially lead to problems. If key predictors are matched very poorly by the eclectic method, one might want to consider assigning the predictor weights manually. It is also important to be aware that the more predictor weights of $v = 0$ there are, the closer we get to the very simple case, where we merely match based on the outcome variable.

7 Conclusion

This paper provides quantitative evidence to the discussion on what would have happened to the UK economy if it had joined the euro zone. Using the synthetic control method, we conclude that it is unlikely that joining the euro zone would have had a significant effect on economic growth in the UK. There is some evidence, which suggests that GDP per capita would have been lower by 6.5% in the period from 1998-2008. However, the effect is not pronounced enough for the result to be clearly significant. After 2008, the GDP per capita path no longer seems to be heavily influenced by the UK's decision to not join the euro, but by its response to the financial crisis. The idiosyncratic shocks experienced by countries during this period and the country specific responses distort our analysis with regards to the effects of the euro, preventing a long term evaluation.

This result is consistent with the estimates obtained by the existing literature that uses GDP per capita as the outcome variable of interest (HM Treasury 2003, Mazumder & Pahl 2013, Pesaran et al. 2007). When considered in the context of the literature on trade, which concludes that the euro has had positive trade effects (Baldwin et al. 2008, Bun & Klaassen 2007), the result implies that there are counteracting effects for the UK that eliminate the potential gains from trade.

This paper also highlights several factors that indicate an analysis using the synthetic control method might be problematic in this specific case. There are two main reasons for concern: Firstly, the pre-treatment fit is not as good as hoped, which suggests that the UK economy has certain characteristics that cannot be perfectly matched by continental European countries. Secondly, placebo analysis creates doubt whether the analysis is correctly picking up the effects of the creation of the euro, or whether there are idiosyncratic shocks that disturb the results.

Finally, it is important to be aware that we have only looked at the economic effects of the euro, a project that had many political motivations. Since the economic effects of the euro appear to be ambiguous in the case of the UK, future research should focus on the political consequences of the euro, when assessing whether the UK should have joined. An analysis of how much the UK has benefitted/suffered as a result of being less closely involved in the decisions regarding the monetary union would be of particular interest.

A Data appendix

The GDP per capita series has been obtained from the OECD.Stat national accounts database. We use real GDP per capita measured in US Dollar. The term 'real GDP' refers to purchasing power parity (PPP) adjusted GDP in current prices, following the definition provided by the World Bank. Using real GDP allows for relative price levels between countries to change over time. This is important when considering series over an extended amount of time, as has been done for this paper.

All variables that are used as predictors have been obtained from the World Development Indicators database of the World Bank. For clarity of presentation, some of the predictor names have been shortened in table 1. Their full names and definitions are given below.

Capital formation refers to gross capital formation (% of GDP) and was formerly called gross domestic investment. Population growth is measured in annual %. Trade is measured as the sum of imports and exports as % of GDP. Inflation is measured as the annual % change in the consumer price index. Industry refers to the value added by the industry sector as % of GDP. Schooling measures the % of total labour force with secondary education or higher. It is obtained from the two separate series called "Labour force with secondary education (% of total)" and "Labour force with tertiary education (% of total)". Domestic credit to private sector by banks is measured as % of GDP, and the size of the service sector is measured by the value added by services as % of GDP.

References

Abadie, A., Diamond, A. & Hainmueller, J. (2011), 'Synth: An r package for snthetic control methods in comparative case studies', *Journal of Statistical Software* **42**(13).

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

Abadie, A., Diamond, A. & Heinmueller, J. (2015), 'Comparative politics and the synthetic control method', *American Journal of Political Science* **59**(2), 495–510.

Abadie, A. & Gardeazabal, J. (2003), 'The economic cost of conflict: A case study of the Basque country', *The American Economic Review* **93**(1), 113–132.

Acemoglu, D., Johnson, S., Kermani, A., Kwak, J. & Mitton, T. (2013), 'The value of connections in turbulent times: Evidence from the United States', *National Bureau of Economic Research Working Paper Series* .

Artis, M. (2000), 'Should the UK join EMU?', *National Institute Economic Review* **171**, 70–81.

Baldwin, R. (2006), 'The euro's trade effects', European Central Bank Working Paper Series No 594.

Baldwin, R., DiNino, V., Fontagné, L., Santis, R. A. D. & Taglioni, D. (2008), 'Study of the impact of the euro on trade and foreign direct investment', European Commission: Directorate-General for Economic and Financial Affairs.

Buiter, W. H. (2008), 'Why the United Kingdom should join the eurozone', *International Finance* **11**(3), 269–282.

Bun, M. J. & Klaassen, F. J. (2007), 'The euro effect on trade is not as large as commonly thought', *Oxford Bulletin of Economics and Statistics* **69**(4), 473–496.

Campos, N. F., Coricelli, F. & Moretti, L. (2014), 'Economic growth and political integration: Estimating the benefits from membership in the european union using the synthetic counterfactuals method', Institute for the Study of Labor, Discussion Paper.

Drees, B. & Pazarbasioglu, C. (1998), *The Nordic Banking Crisis: Pitfalls in Financial Liberalization*, International Monetary Fund.

European Commission (2006), 'One currency for one Europe - the road to the euro'.

European Community (1992), 'United Kingdom: EMU opt-out clause', Maastricht Treaty.

Fernández, C. & Perea, P. G. (2015), 'The impact of the euro on euro area GDP per capita', Banco de España Documentos de Trabajo.

Flam, H. & Nordström, H. (2003), 'Trade volume effects of the euro: Aggregate and sector estimates', Institute for international economic studies, Stockholm University.

Hallett, A. H. (2002), 'Britain and the euro', *Atlantic Economic Journal* **30**(3), 335–348.

HM Treasury (2003), 'Modelling the tranistion to EMU - EMU study by Dr Peter Westaway'.

Layard, R., Buiter, W., Huhne, C., Hutton, W., Kenen, P. & Turner, A. (2002), 'Why Britain should join the euro', *Britain in Europe Campaign Ltd* .

Lucas, Jr., R. E. (1976), Econometric policy evaluation: A critique, *in* K. Brunner & A. Meltzer, eds, 'The Phillips Curve and Labor Markets', North-Holland Publishing Co.

Mazumder, S. & Pahl, R. M. (2013), 'What if the UK had joined the euro in 1999?', *Open Econ Review* **24**, 447–470.

Minford, P. (2008), 'Why the United Kingdom should not join the eurozone', *International Finance* **11**(3), 283–295.

Mundell, R. A. (1961), 'A theory of optimum currency areas', *The American Economic Review* **51**(4).

National Institute of Economic and Social Research (NIESR) (2016), *National Institute's Global Econometric Model* .
URL: *https://nimodel.niesr.ac.uk (accessed 17/07/2016)*

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

Rose, A. K. (1999), 'One money, one market: Estimating the effect of common currencies on trade', *National Bureau of Economic Research Working Paper Series* .

The World Bank (2016), *What is the difference between current and constant data?* .
URL: *https://datahelpdesk.worldbank.org/knowledgebase/articles/114942-what-is-the-difference-between-current-and-constan (accessed 01/07/2016)*

Wyplosz, C. (2006), 'European monetary union - the dark sides of a major success', *Economic Policy* **21**(46), 207–261.