Universität Trier

An Unconventional Approach to Evaluate the Bank of England's Asset Purchase Program

Matthias Neuenkirch

Research Papers in Economics
No. 11/16

An Unconventional Approach to Evaluate the Bank of England's Asset Purchase Program*

Matthias Neuenkirch†

University of Trier

First Draft: November 2, 2016
This Version: November 2, 2016

Abstract

Empirical papers analysing the transmission of (unconventional) monetary policy typically rely on a vector autoregressive framework. In this paper, I complement these studies and employ a matching approach to examine the impact of the Bank of England's asset purchase program on macroeconomic quantities in the UK. My sample covers the period March 2001–December 2015 and five small open inflation targeting economies. Using entropy balancing, I create a synthetic control group comprised of credible counterfactuals for the sample of observations subject to the asset purchase program. My key results are that a 100 bn GBP increase in asset purchases has a significant and positive effect on GDP growth with a peak effect of 0.66–0.69 percentage points (pp) after 30 months. The same increase leads to a reduction in the inflation gap with a peak effect between −0.77 and −0.94 pp after 30 months. An in-depth analysis reveals that the latter finding is not driven by the choice of the empirical methodology. In contrast, I find that the returns on asset purchases are decreasing (i) over time and (ii) with the level of asset purchases. This causes the impact of asset purchases on the inflation gap to eventually become negative.

JEL codes: E52, E58.

Keywords: Asset Purchases, Bank of England, Entropy Balancing, Matching, Quantitative Easing, Treatment Effects, Unconventional Monetary Policy.

*Thanks to Tobias Kranz, Florian Neumeier, and Peter Tillmann for their helpful comments on earlier versions of the paper.

†Corresponding author: University of Trier, Department of Economics, D-54286 Trier, Germany, Tel.: +49-651-2012629, Email: neuenkirch@uni-trier.de.

1 Introduction

Many of the world's central banks engage in unconventional monetary policy measures since the outbreak of the global financial crisis and the subsequent Great Recession. Arguably, the most important measure is commonly referred to as quantitative easing (QE) and aims, along with other measures, such as forward guidance, at overcoming the zero-lower bound of interest rates and providing additional monetary stimulus. The Bank of England (BoE) started QE in March 2009 when it announced the intention to use central bank money to purchase 75 bn GBP of public and private assets over the next three months. The purpose of this policy is to boost nominal spending and to help achieve the 2% inflation target (IT) (Joyce et al, 2011b). Since then, the volume of these purchases was increased six times reaching 375 bn GBP.[1]

The extant empirical literature on the macroeconomic effects of QE in the UK mostly supports the notion that expansionary monetary policy leads to an increase in prices and output. Kapetanios et al (2012) show that the asset purchase (AP) program had a peak effect on the level of real GDP of around 1.5 percentage points (pp), and on consumer price index (CPI) inflation of around 1.25 pp. Baumeister and Benati (2013) derive from their results that without QE annualised inflation in the UK would have fallen to −4 percent and output growth would have reached a trough of −12 percent in the first quarter of 2009. Weale and Wieladek (2016) document that an AP announcement of 1% of the GDP leads to a statistically significant rise of 0.25 pp and 0.32 pp in real GDP and the CPI, respectively. Gambacorta et al (2014) estimate a panel model for eight advanced economies and find that an exogenous increase in central bank balance sheets leads to a temporary rise in economic activity that is similar to the effects of conventional monetary policy. The effect on consumer prices, however, is weaker and less persistent. Numerous studies on the macroeconomic effects of QE for other central banks document similar findings and are summarised, inter alia, in the overview articles by Joyce et al (2012) and Martin and Milas (2012).[2]

[1] See also the top left panel of Figure A1 in the Appendix.

[2] The financial market impact of QE in the UK is documented, inter alia, by Breedon et al (2012), Christensen and Rudebusch (2012), Joyce et al (2011a), and Meaning and Zhu (2011).

All these papers rely on a vector autoregressive (VAR) framework to quantify the impact of unconventional monetary policy, an approach that goes back to Sims's (1980) seminal paper. In this paper, I complement these studies and employ a matching approach to examine the impact of the BoE's AP program on macroeconomic quantities in the UK. Matching approaches are commonly used in microeconometric applications. The macroeconomic literature employing such a technique is scant. For instance, Lin and Ye (2007) quantify the macroeconomic impact of IT in seven economies. Neuenkirch and Tillmann (2016) find that central bank governors influence macroeconomic outcomes and expectations above and beyond the institution they head. To the best of my knowledge, there is no literature assessing the impact of QE on macroeconomic quantities using a matching approach.

My sample covers the period March 2001–December 2015, and includes five small open economies with an explicit IT (Australia, Canada, New Zealand, Norway, the United Kingdom). I estimate treatment effects for the country "exposed" to quantitative easing (i.e., the UK from March 2009 onwards) compared to non-treated countries that are as similar as possible along observable dimensions (i.e., the other four economies and the UK before March 2009).[3] First, I create a weighted control group comprised of countries not exposed to QE that is similar to the treatment group with regard to macroeconomic characteristics that potentially affect the outcome variables of interest. Second, I take into account that selection into the treatment group might be endogenous and control for why QE might be implemented.

In contrast to a VAR approach, I do not identify the model via, for instance, recursive identification (Sims, 1980), sign restrictions (Uhlig, 2005), or a narrative approach (Romer and Romer, 2004; Cloyne and Hürtgen, 2016). My approach balances the treatment group and the control group based on observable pre-treatment characteristics. By doing so, I eliminate pre-treatment differences across treated and non-treated units and ensure that differences in the macroeconomic performance of the UK under the AP

[3]Another reason for the focus on the UK is the existence of a monthly dataset for asset purchases. In the case of the Federal Reserve and the European Central Bank, for instance, one has to resort to a hand full of announcements to identify QE measures and changes thereof.

program and that of countries without QE are due to the "QE treatment." In addition, my approach allows for an analysis of expected and unexpected changes in asset purchases, whereas VAR analyses typically focus on the consequences of monetary policy shocks. Due to the immense communication efforts by the BoE, however, there are very few unexpected changes in monetary policy over the recent years and, consequently, a VAR model that focuses only on shocks could paint a less than complete picture of the monetary transmission mechanism.

My key results are as follows. Confirming the previous empirical VAR literature, a 100 bn GBP increase in asset purchases has a significant and positive effect on GDP growth with a peak effect of 0.66–0.69 pp after 30 months. The same increase leads to a reduction in the inflation gap (i.e., the inflation rate minus its target value) with a peak effect between −0.77 and −0.94 pp after 30 months, a result in sharp contrast to the existing literature. An in-depth analysis reveals that the latter finding is not driven by the choice of the empirical methodology. In contrast, I find that the returns on asset purchases are decreasing (i) over time and (ii) with the level of asset purchases. This causes the impact of QE on the inflation gap to eventually become negative.

The remainder of the paper is organised as follows. Section 2 introduces the empirical methodology and Section 3 the dataset. Section 4 presents the empirical results. Section 5 concludes with some implications for researchers and monetary policy makers.

2 Empirical Methodology

The biggest challenge of the empirical work presented below is to establish a causal link between the implementation of the AP program and the subsequent reactions of macroeconomic variables in the UK. The reasons why a country might choose to implement QE are obviously related to the country's macroeconomic situation. To overcome this endogeneity problem with regard to the implementation of QE, I employ a matching approach.

My analysis is based on the idea that the implementation of QE represents a treatment. Observations with QE in place (i.e., the UK from March 2009 onwards) comprise the treatment group, observations without QE (i.e., AUS, CAN, NOR, NZ, and the UK before March 2009) make up the potential control group. My measure of interest is the average treatment effect on the treated (ATT), defined as follows:

$$\tau_{ATT} = E[Y(1)|T = 1] - E[Y(0)|T = 1] \tag{1}$$

$Y(\cdot)$ represents the outcome variable of interest, that is, GDP growth and the inflation gap at several horizons in the future (3, 6, 12, 18, 24, 30, and 36 months ahead). T indicates whether a unit is exposed to treatment ($T = 1$) or not ($T = 0$). Accordingly, $E[Y(1)|T = 1]$ is the expected outcome after treatment, and $E[Y(0)|T = 1]$ the counterfactual outcome, that is, the outcome a unit exposed to treatment would have achieved if it had not received treatment.

As the counterfactual outcome is not observable, I need a suitable proxy to identify the ATT. If the treatment is randomly assigned, then the average outcome of units not exposed to treatment, $E[Y(0)|T = 0]$, is a proper substitute. However, as discussed before, the implementation of QE and, thus, selection into treatment is endogenous. In general, the idea of matching estimators is to mimic randomization with regard to assignment of the treatment. The unobserved counterfactual outcome is imputed by matching the treated units with untreated units that are as similar as possible with regard to all pre-treatment characteristics that (i) are associated with selection into treatment, and (ii) influence the outcome of interest. The realizations of the macroeconomic variables in the future are then used as an empirical proxy for the unobserved counterfactual. Formally, the estimate of the ATT based on matching is defined as follows:

$$\widehat{\tau}_{ATT} = E[Y(1)|T = 1, X = x] - E[Y(0)|T = 0, X = x] \tag{2}$$

x is a vector of relevant pre-treatment characteristics, which are described in Section

3, $E[Y(1)|T=1, X=x]$ is the expected outcome for the units that received treatment, and $E[Y(0)|T=0, X=x]$ is the expected outcome for the treated units' best matches.

In this paper, I use entropy balancing to select matches for the units exposed to treatment and to estimate the ATT, a method proposed by Hainmueller (2012). Entropy balancing is implemented in two consecutive steps. First, weights are computed that are assigned to units not subject to treatment. These weights are chosen to satisfy pre-specified balanced constraints involving sample moments of pre-treatment characteristics by remaining, at the same time, as close as possible to uniform base weights. In my analysis, the balance constraints require equal covariate means across the treatment and the control group, which ensures that the control group contains, on average, units not subject to treatment that are as similar as possible to units that received treatment. Second, the weights obtained in the first step are used in a regression analysis with the treatment indicator as an explanatory variable. This yields an estimate for the ATT, that is, the conditional difference in means for the outcome variable between the treatment and control group.

In this paper's context, the intuition behind entropy balancing is to compare the macroeconomic performance of the UK under the AP program to that of countries without QE that are as similar as possible to the treatment group. The average difference in GDP growth and the inflation gap between the country under QE and the "closest" non-QE observations must then be due to treatment. In this sense, the empirical approach mimics a randomised experiment by balancing the treatment and the control group based on observable characteristics.

By combining matching and regression analysis, entropy balancing has some advantages over other treatment effect estimators. A particularly important advantage over "simple" regression-based approaches, as well as matching methods relying on propensity scores, is that entropy balancing is non-parametric in the sense that no empirical model for either the outcome variable or selection into treatment needs to be specified. Hence, potential types of misspecification, like those, for instance, regarding the functional form of the empirical model, which likely lead to biased estimates, are ruled out. Also, in contrast to regression-based analyses, treatment effects estimates based on entropy balancing do not suffer from multicollinearity, as the re-weighting scheme orthogonalises the covariates with respect to the treatment indicator.

Moreover, in contrast to other matching methods, entropy balancing ensures a high covariate balance between the treatment and control group even in small samples. With "conventional" matching methods such as, for instance, nearest neighbour matching or propensity score matching, each treated unit—in the simplest case—is matched with the one untreated unit that is closest in terms of a metric balancing score. Accordingly, the control group is comprised of only a subset of the units that are not subject to treatment (Diamond and Sekhon, 2013; Hainmueller, 2012). However, when the number of untreated units is limited and the number of pre-treatment characteristics is large, this procedure does not guarantee a sufficient balance of pre-treatment characteristics across the treatment and control groups. This is a serious problem, as a low covariate balance may lead to biased treatment effect estimates. In contrast, with entropy balancing, the vector of weights assigned to the units not exposed to treatment is allowed to contain any non-negative values. Thus, a synthetic control group is designed that represents a virtually perfect image of the treatment group. Entropy balancing thus can be interpreted as a generalization of conventional matching approaches. Finally, by combining a re-weighting scheme with a regression analysis, entropy balancing allows to properly address the panel structure of the data. In particular, I am able to control for unobserved heterogeneity by using country-fixed effects and for global shocks by time-fixed effects in the second step of the matching approach, that is, the regression analysis.

3 Data

I select a control group comprised of untreated units that is, on average, as similar as possible to the treatment group with regard to relevant pre-treatment characteristics. Since the BoE has an explicit IT mandate, my control group consists of four IT central banks in small open advanced economies that did not implement QE measures after the recent financial crisis and the Great Recession, namely, the Reserve Bank of Australia, the Bank of Canada, the Norges Bank, and the Reserve Bank of New Zealand. Consequently, my sample starts in March 2001 with the inception of the IT in Norway. With this step, I ensure that the treated and non-treated units are as similar as possible with respect to the institutional environment.

I generate two different sets of weights for the empirical analysis. The first one relies on the full sample period (i.e., March 2001–December 2015). The second one focuses on the period after the Lehman collapse (i.e., September 2008–December 2015) in an effort to further harmonise the economic environment. The matching variables capture factors that influence the likelihood of being selected into treatment and are related to the subsequent macroeconomic performance. Vector x, therefore, includes the yield curve slope (i.e., the 10-year government bond interest rate minus the 3-month money market interest rate), the GDP growth rate, the unemployment rate, the inflation gap (i.e., the inflation rate minus its target value), and the growth rate of credit to the non-financial sector.[4] Figure A1 provides a plot of these variables for the five economies.

Turning to the treatment variable, I rely on the AP dataset provided by the Bank of England on its website. In total, I have 82 observations when QE was in place. In contrast, the total number of observations without QE is 808. Therefore, the potential control group is roughly 10 times larger than the treatment group, which allows me to easily obtain a weighted control group that satisfies the condition of covariate balance. In the balancing stage of the matching approach, I employ a binary variable indicating whether or not the AP program was in place. In the regression stage, I use a continuous indicator that measures the AP volume in 100 bn GBP.

Table 1 shows the sample means of all matching covariates and the full sample period split into two groups. Column (1) contains observations with QE in place (the treatment group). Column (2) contains observations without QE in place (the potential

[4]Note that the inclusion of further macroeconomic variables is not feasible as this would violate the condition of perfect covariate balancing.

control group). Column (3) shows differences in means between both groups alongside the p-value of a corresponding t-test.

Table 1: Descriptive Statistics: Full Sample

	(1) AP	(2) No AP	(3)=(2)−(1) Difference		(4) Control	(5)=(4)−(1) Difference	
Yield Curve Slope	2.01	0.55	1.46	[0.00]	2.01	0.00	[1.00]
GDP Growth	1.27	2.27	−0.99	[0.00]	1.27	0.00	[1.00]
Unemployment Rate	7.20	5.32	1.88	[0.00]	7.20	0.00	[1.00]
Inflation Gap	0.39	−0.05	0.44	[0.00]	0.39	0.00	[1.00]
Credit Growth	4.36	7.25	−2.89	[0.00]	4.36	0.00	[1.00]
Observations	82	808			82		

Notes: Column (1) shows the average conditions for observations with the AP program in place (AP) and column (2) the average conditions for observations without the AP program in place (No AP). Column (3) shows differences in the average conditions between both groups alongside p-values of the corresponding t-tests. Column (4) shows the average conditions of the synthetic control group (Control), which is created by entropy balancing. Column (5) shows differences in the average conditions between the synthetic control group and the treatment group alongside p-values of the corresponding t-tests.

The numbers reveal that, with respect to all relevant pre-treatment characteristics, times with QE in place differ, and notably so, from times when it is not. The yield curve slope is steeper indicating a more expansionary monetary policy when QE is in place. In addition, real and nominal macroeconomic conditions are worse as, (i) GDP growth is lower, (ii) the unemployment rate is higher, (iii) the inflation gap is higher, and (iv) credit growth is lower under QE. These descriptive findings illustrate why it is important to select an appropriate control group when using a matching approach before calculating treatment effects as, otherwise, the effect of QE on GDP growth and the inflation gap might be incorrectly estimated.

Column (4) shows the sample means of all matching covariates for the synthetic control group obtained via entropy balancing. Column (5) shows differences in means between the synthetic control group and the treatment group alongside p-values of the corresponding t-tests. Comparing the average realizations of the pre-treatment characteristics of the treatment group to those of the synthetic control group reveals the efficacy of entropy balancing. All covariates are virtually perfectly balanced and no statistically significant difference remains. Consequently, I am confident that the control group in the subsequent empirical analysis is comprised of credible counterfactuals for the sample of observations subject to the BoE's AP program.

Table 2 repeats this exercise for the financial crisis subsample. The picture is very similar to the full sample. First, monetary policy is more expansionary and the real and nominal macroeconomic situation is less favourable under QE.[5] Second, entropy balancing yields virtually perfectly balanced covariates.

Table 2: Descriptive Statistics: Financial Crisis

	(1) AP	(2) No AP	(3)=(2)−(1) Difference		(4) Control	(5)=(4)−(1) Difference	
Yield Curve Slope	2.01	0.96	1.05	[0.00]	2.01	0.00	[1.00]
GDP Growth	1.27	1.57	−0.30	[0.21]	1.27	0.00	[1.00]
Unemployment Rate	7.20	5.65	1.55	[0.00]	7.20	0.00	[1.00]
Inflation Gap	0.39	−0.29	0.68	[0.00]	0.39	0.00	[1.00]
Credit Growth	4.36	5.47	−1.11	[0.00]	4.36	0.00	[1.00]
Observations	82	358			82		

Notes: See Table 1.

Table 3 and Figure 1 provide an overview of the entropy balancing weights by country and over time.

Table 3: Entropy Balancing Weights by Country

	Full Sample	Financial Crisis
Australia	2.99	0.38
Canada	53.71	38.74
Norway	1.26	0.64
New Zealand	22.38	41.38
United Kingdom	1.66	0.85

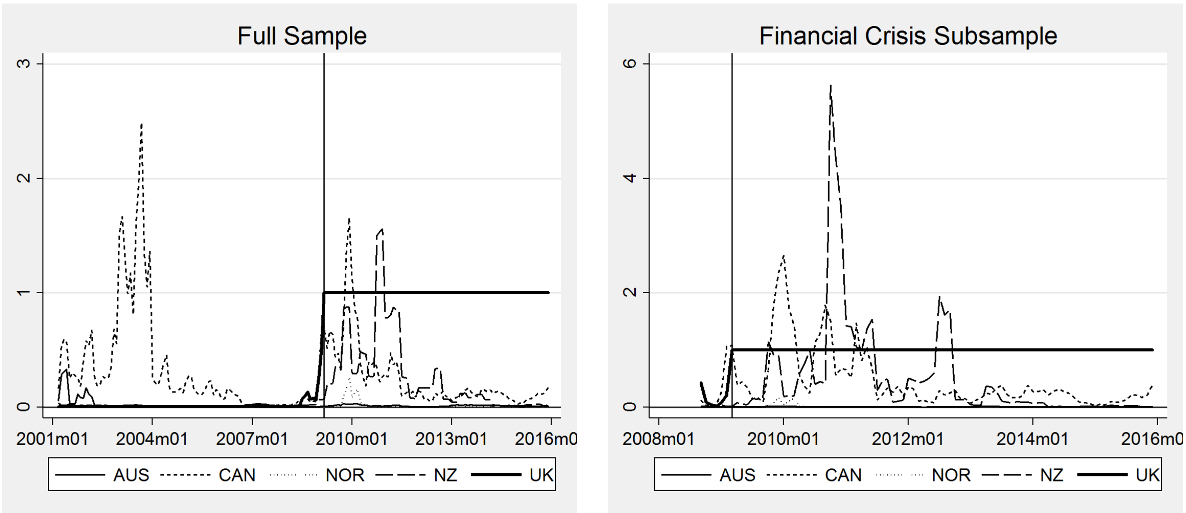
Notes: Table 3 shows the cumulative frequency of entropy balancing weights per country during the full sample period (left panel) and the financial crisis subsample (right panel).

The best match for the macroeconomic conditions in the UK under the AP program is a mix of the macroeconomic environment in Canada and New Zealand. The situation in Australia and Norway is apparently too different to be an important contributor in the matching algorithm. There are some differences across both weight sets as the role for Canada (New Zealand) is larger (smaller) when the sample is not restricted

[5]Note that the difference in GDP growth between treated and non-treated units is not statistically significant when focusing on the financial crisis subsample.

to the financial crisis period. Nevertheless, the correlation across both sets of weights is 0.88. In general, the largest weights are found during the first two years of the AP program (until June 2011) and, for the unrestricted sample, Canada during the year 2003.

Figure 1: Entropy Balancing Weights over Time



4 Empirical Results

Table 4 shows the ATTs for GDP growth. The panel "Non-Weighted" shows the results of unweighted estimations to illustrate a potential endogeneity bias. The panels "Weighted Full" and "Weighted Crisis" show the results of estimations with the weighting scheme obtained via entropy balancing for the full sample period and the financial crisis subsample, respectively. All estimations include country-fixed effects and year-fixed effects.

In all three panels, I observe a positive effect of QE on GDP growth in the UK. When taking into account the endogeneity associated with the selection into treatment (bottom two panels), the effects of asset purchases on GDP growth are notably larger compared to the unweighted estimations. The peak effect after 30 months is 0.66–0.69 pp for a 100 bn GBP increase in the AP program. These positive effects are in line with the earlier literature using VAR models. The only difference is that the peak effect occurs after a longer outside lag.

Table 4: Treatment Effects: GDP Growth

Non-Weighted	$t+3$	$t+6$	$t+12$	$t+18$	$t+24$	$t+30$	$t+36$
Treatment Effect	0.177	0.230	0.262	0.232	0.226	0.306	0.415
Standard Error	(0.058)	(0.060)	(0.063)	(0.067)	(0.072)	(0.076)	(0.084)
p-value	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]
Observations	875	860	830	800	770	740	710

Weighted Full	$t+3$	$t+6$	$t+12$	$t+18$	$t+24$	$t+30$	$t+36$
Treatment Effect	0.343	0.266	0.313	0.372	0.487	0.685	0.616
Standard Error	(0.052)	(0.053)	(0.056)	(0.055)	(0.049)	(0.048)	(0.048)
p-value	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]
Observations	875	860	830	800	770	740	710

Weighted Crisis	$t+3$	$t+6$	$t+12$	$t+18$	$t+24$	$t+30$	$t+36$
Treatment Effect	0.272	0.309	0.453	0.565	0.606	0.660	0.524
Standard Error	(0.077)	(0.075)	(0.082)	(0.076)	(0.064)	(0.067)	(0.081)
p-value	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]
Observations	425	410	380	350	320	290	260

Notes: Table 4 shows average treatment effects on the treated obtained by (weighted) least squares regressions. Standard errors are in parentheses and p-values in brackets. Panel "Non-Weighted" shows results of unweighted estimations and panels "Weighted Full" and "Weighted Crisis" show results of weighted estimations. Estimations include country-fixed effects and year-fixed effects.

Table 5: Treatment Effects: Inflation Gap

Non-Weighted	$t+3$	$t+6$	$t+12$	$t+18$	$t+24$	$t+30$	$t+36$
Treatment Effect	0.194	0.160	0.096	0.023	−0.060	−0.147	−0.249
Standard Error	(0.043)	(0.044)	(0.047)	(0.050)	(0.052)	(0.057)	(0.063)
p-value	[0.00]	[0.00]	[0.04]	[0.65]	[0.25]	[0.01]	[0.00]
Observations	875	860	830	800	770	740	710

Weighted Full	$t+3$	$t+6$	$t+12$	$t+18$	$t+24$	$t+30$	$t+36$
Treatment Effect	−0.210	−0.250	−0.455	−0.520	−0.634	−0.770	−0.669
Standard Error	(0.045)	(0.048)	(0.050)	(0.049)	(0.043)	(0.038)	(0.047)
p-value	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]
Observations	875	860	830	800	770	740	710

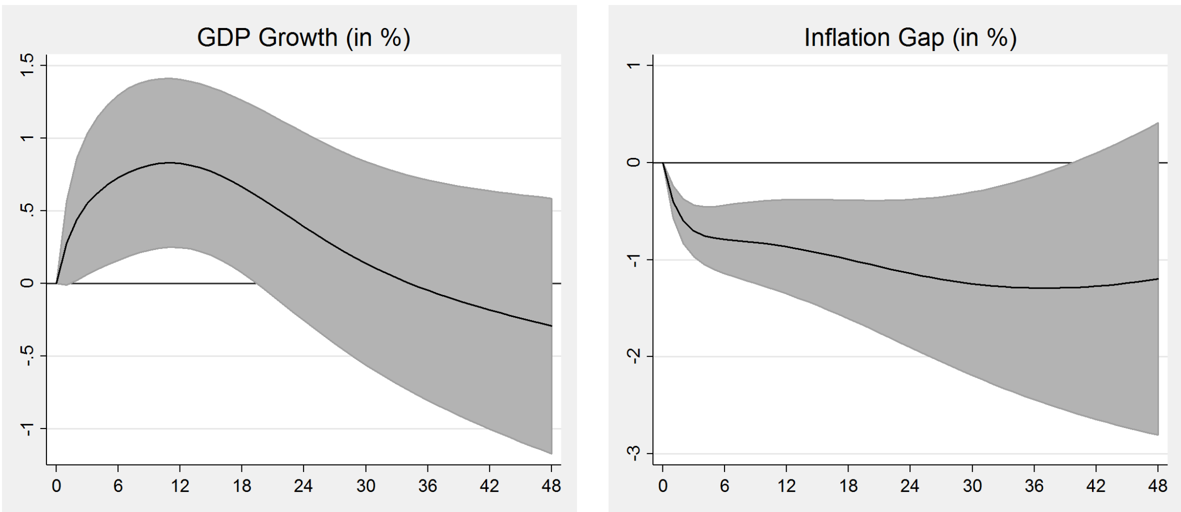
Weighted Crisis	$t+3$	$t+6$	$t+12$	$t+18$	$t+24$	$t+30$	$t+36$
Treatment Effect	−0.220	−0.249	−0.551	−0.738	−0.872	−0.941	−0.828
Standard Error	(0.071)	(0.077)	(0.077)	(0.068)	(0.054)	(0.056)	(0.079)
p-value	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]
Observations	425	410	380	350	320	290	260

Notes: See Table 4.

Table 5 repeats this exercise and shows the ATTs for the inflation gap. Here, I find a positive treatment effect only during the first 12 months and for the non-weighted estimations that ignore the endogenous selection into treatment. Again, the results in the bottom two panels indicate a bias, as the effects on inflation is negative over all horizons and notably larger in absolute terms. Similar to the estimations for GDP growth, the peak effect between −0.77 and −0.94 pp is found after 30 months.

These results are in sharp contrast to the previous literature. To confirm that these differences are not driven by the choice of the empirical methodology I estimate a standard VAR model for the period August 2008–December 2015. As favoured by the information criteria, I employ one lag of the six variables also used in the matching approach, that is, the unemployment rate, the GDP growth rate, the inflation gap, credit growth, the level of asset purchases, and the yield curve slope. The model is identified using Cholesky decomposition and in the order outlined in the previous sentence, that is, the unemployment rate is ordered first. Figure 2 shows the responses of GDP growth and the inflation gap to a 100 bn GBP shock in asset purchases.

Figure 2: Impulse Responses



Notes: The figure shows selected impulse responses to a 100 bn GBP shock in asset purchases. Cholesky decomposition is based on the ordering (i) unemployment rate, (ii) GDP growth rate, (iii) inflation gap, (iv) credit growth, (v) level of asset purchases, and (vi) yield curve slope. Shaded areas indicate 95% confidence bands. Full set of impulse responses is available on request.

A 100 bn GBP shock in asset purchases leads to an increase in GDP growth that is significant for a horizon of 2–19 months and exerts a peak effect of 0.83 pp after 11 months. The corresponding impulse response function for the inflation gap is significant for 39 months with a peak effect of −1.29 pp after 37 months.[6] These results confirm that in the case of the inflation gap, the results that contrast with the previous literature are not driven by the choice of the empirical methodology.

One possible explanation is that the estimations in my paper are based on a longer sample period than the estimations of Kapetanios et al (2012), Gambacorta et al (2014), Baumeister and Benati (2013), and Weale and Wieladek (2016). For the papers just cited, the estimations end in September 2010, June 2011, the fourth quarter of 2011, and May 2014, respectively, whereas my sample ends in December 2015. Indeed, Goodhart and Ashworth (2012) conclude that the returns on QE are diminishing and could potentially turn negative.

Consequently, I modify the regression model in two different ways. First, I include an additional interaction term of the volume of asset purchases with a variable that measures the number of months that have gone by since the inception of QE in the UK (i.e., a trend variable that takes the value 1 in March 2009, 2 in April 2009, ..., and 82 in December 2015). Second, I include an additional regressor "asset purchases squared" to test for potential non-linearities in the effect of QE on the inflation gap. Table 6 sets out the results.[7]

The results in the top panel show an initial positive effect of asset purchases on the inflation gap over all horizons. However, the interaction term of asset purchases with the number of months elapsed since the inception of the program indicates that the effect becomes weaker over time and eventually turns negative. The row "Negative Effect" shows when the positive effect turns negative. This figure varies considerably across models (16–59 months) but it is safe to say that the returns on QE are diminishing and indeed turn negative at some point in time.

[6] Note that this impulse response function approaches zero when considering horizons longer than 48 months.

[7] Note that applying the same exercises in case of the treatment effects on GDP growth does neither reveal a clear pattern over time nor a distinct non-linearity. The results are available on request.

Table 6: Additional Treatment Effects: Inflation Gap

Over Time	$t+3$	$t+6$	$t+12$	$t+18$	$t+24$	$t+30$	$t+36$
AP	0.787	0.628	0.389	0.458	0.530	0.310	0.732
	(0.095)	(0.107)	(0.116)	(0.112)	(0.090)	(0.072)	(0.074)
	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]
AP × Time	−0.013	−0.012	−0.012	−0.015	−0.019	−0.020	−0.029
	(0.001)	(0.001)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)
	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]
Negative Effect	59.0	51.7	32.1	31.0	27.8	15.6	25.1
Observations	875	860	830	800	770	740	710

Non-Linear	$t+3$	$t+6$	$t+12$	$t+18$	$t+24$	$t+30$	$t+36$
AP	0.163	0.279	0.737	1.276	1.223	0.561	0.359
	(0.173)	(0.182)	(0.182)	(0.162)	(0.126)	(0.105)	(0.128)
	[0.35]	[0.13]	[0.00]	[0.00]	[0.00]	[0.00]	[0.01]
AP^2	−0.080	−0.114	−0.259	−0.394	−0.417	−0.312	−0.257
	(0.036)	(0.038)	(0.038)	(0.034)	(0.027)	(0.023)	(0.030)
	[0.03]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]
Maximum Impact	1.02	1.22	1.42	1.62	1.47	0.90	0.70
Negative Effect	2.04	2.45	2.85	3.24	2.93	1.80	1.40
Observations	875	860	830	800	770	740	710

Notes: Table 6 shows average treatment effects on the treated obtained by weighted least squares regressions. Standard errors are in parentheses and *p*-values in brackets. Panel "Over Time" ("Non-Linear") shows treatment effects over time (non-linear treatment effects) with weights obtained using the full sample period. Estimations include country-fixed effects and year-fixed effects. The rows "Negative Effect" show the number of months after which (upper panel) and the level of asset purchases beyond which (lower panel) the effect on the inflation gap turns negative. The row "Maximum Impact" shows the level of asset purchases with the maximum positive effect on the inflation gap.

The results in the bottom panel provide a similar picture. For lower AP levels, the effect is positive. However, the significant negative quadratic term indicates that there is a non-linearity. The row "Maximum Impact" indicates that the AP level with the maximum positive impact varies between 70 bn GBP and 162 bn GBP, depending on the specification. After these peaks, the effect of QE on the inflation gap becomes smaller and eventually negative at levels between 140 bn GBP and 324 bn GBP as indicated in the row "Negative Effect."

All these findings are in line with Goodhart and Ashworth (2012). The BoE's AP program indeed has diminishing returns for the inflation gap and the effect eventually turns negative. This finding, which also holds in a VAR context, is new to the literature, as previous papers focused on shorter sample periods.

5 Conclusions

Empirical papers analysing the transmission of (unconventional) monetary policy typically rely on a vector autoregressive framework. In this paper, I complement these studies and employ a matching approach to examine the impact of the Bank of England's asset purchase program on macroeconomic quantities in the UK. My sample covers the period March 2001–December 2015, and five small open economies with an explicit IT (Australia, Canada, New Zealand, Norway, the United Kingdom). Using entropy balancing, I create a synthetic control group comprised of credible counterfactuals for the sample of observations subject to QE based on observable pre-treatment characteristics.

My key results are as follows. Confirming the previous empirical VAR literature, a 100 bn GBP increase in asset purchases has a significant and positive effect on GDP growth with a peak effect of 0.66–0.69 pp after 30 months. The same increase leads to a reduction in the inflation gap with a peak effect between −0.77 and −0.94 pp after 30 months, a finding that is in sharp contrast to the existing VAR literature. An in-depth analysis reveals that the latter finding is not driven by the choice of the empirical methodology. In contrast, I find that the returns on asset purchases are decreasing (i) over time and (ii) with the level of asset purchases. This causes the impact of QE on the inflation gap to eventually become negative.

This paper has some implications for researchers and policymakers. For researchers, I provide some evidence that evaluating the transmission of (unconventional) monetary policy in a matching approach offers an interesting alternative. Monetary policymakers should be aware of the fact that the returns on QE might be non-linear and that QE eventually could have detrimental effects.

References

Baumeister, C. and Benati, L. (2013), "Unconventional Monetary Policy and the Great Recession: Estimating the Macroeconomic Effects of a Spread Compression at the Zero Lower Bound," *International Journal of Central Banking* 9 (2), 165–212.

Breedon, F., Chadha, J. S., and Waters, A. (2012), "The Financial Market Impact of UK Quantitative Easing," *Oxford Review of Economic Policy* 28 (4), 702–728.

Christensen, J. H. E. and Rudebusch, G. D. (2012), "The Response of Interest Rates to US and UK Quantitative Easing," *Economic Journal* 122 (November), F385–F414.

Cloyne, J. and Hürtgen, P. (2016), "The Macroeconomic Effects of Monetary Policy: A New Measure for the United Kingdom," *American Economic Journal: Macroeconomics*, 8 (4), 75–102.

Diamond, A. and Sekhon, J.S. (2013), Genetic Matching for Estimating Causal Effects: A General Multivariate Matching Method for Achieving Balance in Observational Studies," *Review of Economics and Statistics* 95, 932–945.

Gambacorta, L., Hofmann, B., and Peersman, G. (2014), "The Effectiveness of Unconventional Monetary Policy at the Zero Lower Bound: A Cross-Country Analysis," *Journal of Money, Credit and Banking* 46 (4), 615–642.

Goodhart, C. A. E. and Ashworth, J. P. (2012), "QE: A Successful Start May be Running into Diminishing Returns," *Oxford Review of Economic Policy* 28 (4), 640–670.

Hainmueller, J. (2012), "Entropy Balancing for Causal Effects: A Multivariate Reweighting Method to Produce Balanced Samples in Observational Studies," *Political Analysis* 20, 25–46.

Joyce, M. A. S., Lasaosa, A., Stevens, I., and Tong, M. (2011), "The Financial Market Impact of Quantitative Easing in the United Kingdom," *International Journal of Central Banking* 7 (3), 113–161.

Joyce, M., Tong, M., Woods, R. (2011), "The United Kingdom's Quantitative Easing Policy: Design, Operation, and Impact," *Bank of England Quarterly Bulletin* 2011Q3, 200–212.

Joyce, M., Miles, D., Scott, A., and Vayanos, D. (2012), "Quantitative Easing and Unconventional Monetary Policy – An Introduction," *Economic Journal* 122 (November), F271–F288.

Kapetanios, G., Mumtaz, H., Stevens, I., and Theodoridis, K. (2012), "Assessing the Economy-Wide Effects of Quantitative Easing," *Economic Journal* 122 (November), F316–F347.

Lin, S. and Ye, H. (2007), "Does Inflation Targeting Really Make a Difference? Evaluating the Treatment Effect of Inflation Targeting in Seven Industrial Countries," *Journal of Monetary Economics* 54, 2521–2533.

Meaning, J. and Zhu, F. (2011), "The Impact of Recent Central Bank Asset Purchase Programmes," *BIS Quarterly Review* December 2011, 73–83.

Martin, C. and Milas, C. (2012), "Quantitative Easing: A Sceptical Survey," *Oxford Review of Economic Policy* 28 (4), 750–764.

Neuenkirch, M. and Tillmann, P. (2016), "Does a Good Central Banker Make a Difference?" *Economic Inquiry* 54 (3), 1541–1560.

Romer, C. D., and D. H. Romer (2004), "A New Measure of Monetary Shocks: Derivation and Implications," *American Economic Review* 94 (4), 1055–1084.

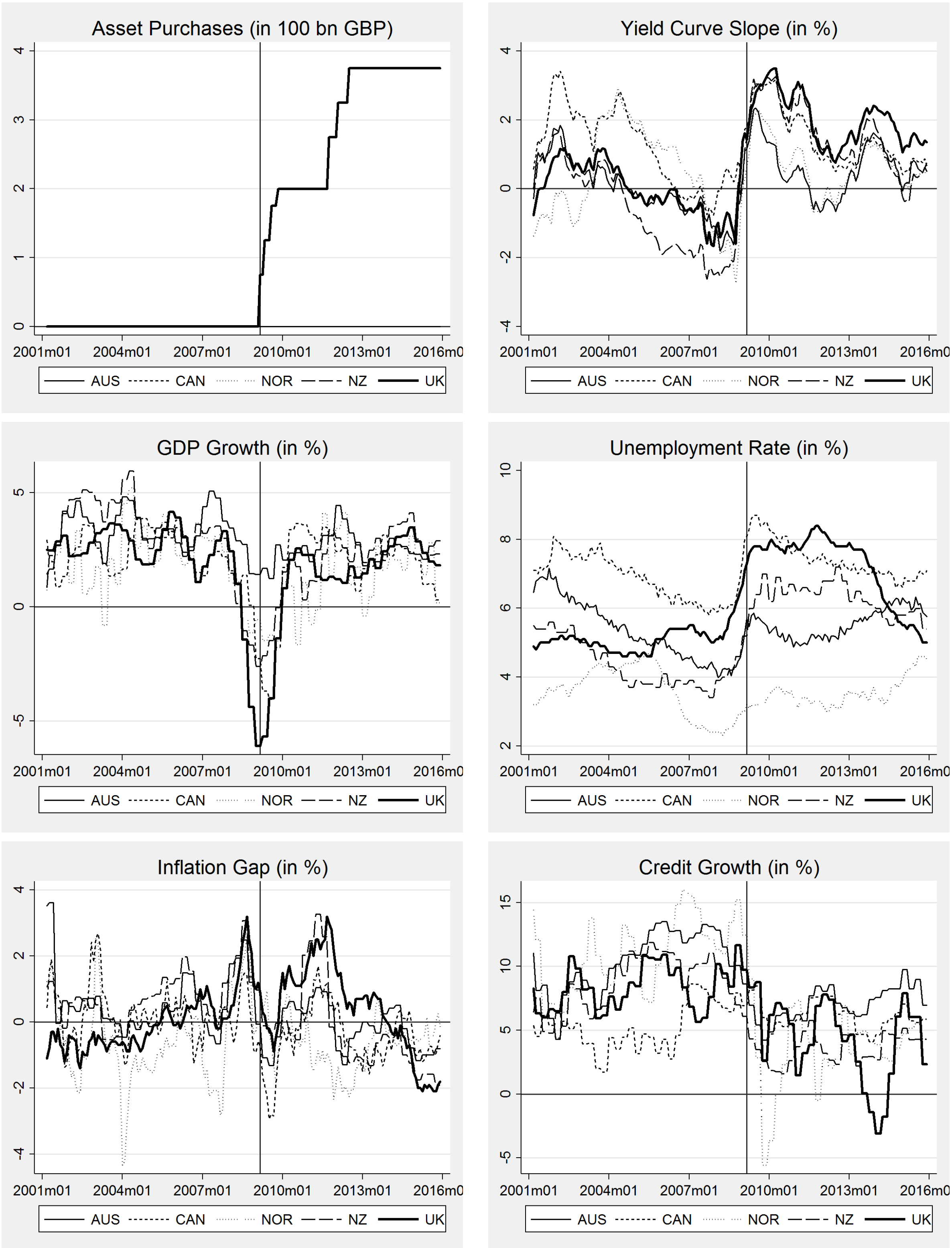
Sims, C. A. (1980), "Macroeconomics and Reality," *Econometrica* 48 (1), 1–48.

Uhlig, H. (2005), "What are the Effects of Monetary Policy on Output? Results from an Agnostic Identification Procedure," *Journal of Monetary Economics* 52, 381–419.

Weale, M. and Wieladek, T. (2016), "What are the Macroeconomic Effects of Asset Purchases?" *Journal of Monetary Economics* 79, 81–93.

Appendix

Figure A1: Macroeconomic Data



Source: Bank of England (asset purchases), OECD (interest rates, GDP growth, unemployment rate, and inflation rate), and BIS (credit growth).