

Dreuw, Peter

Article — Published Version

Structural time series models and synthetic controls— assessing the impact of the euro adoption

Empirical Economics

Provided in Cooperation with:

Springer Nature

Suggested Citation: Dreuw, Peter (2022) : Structural time series models and synthetic controls—
assessing the impact of the euro adoption, Empirical Economics, ISSN 1435-8921, Springer, Berlin,
Heidelberg, Vol. 64, Iss. 2, pp. 681-725,
<https://doi.org/10.1007/s00181-022-02257-x>

This Version is available at:

<https://hdl.handle.net/10419/307077>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



<https://creativecommons.org/licenses/by/4.0/>



Structural time series models and synthetic controls—assessing the impact of the euro adoption

Peter Dreuw¹ 

Received: 11 July 2021 / Accepted: 30 April 2022 / Published online: 25 June 2022
© The Author(s) 2022

Abstract

So far, empirical research on an ex-post benchmark of the euro adoption has relied on the synthetic control method by Abadie & Gardeazabal (Am Econ Rev 93:112–132, 2003) and Abadie et al. (J Am Stat Assoc 105:493–505, 2010, Am J Polit Sci, 59:495–510, 2015). However, the evidence obtained with this method is not overly consistent, leading to the conclusion that the method is not too robust to different settings of the adjustment screws. Using a new method developed by Harvey & Thiele (J Appl Econ 36:71–85, 2021) based on structural time series models, I find that France and Italy are clear losers from the euro while there are no real winners until 2019. Spain, Netherlands, Greece gain in the period before the financial crisis, but afterwards they lose. In fact, only the German economy is robust to the two crises but it loses until 2008. Relating the theories of optimum currency areas to the estimated gaps of the euro adopters from their synthetic controls, I find that openness, real convergence, and net migration are the main drivers of gains from a euro adoption, while the main drivers of losses are low levels of competitiveness, fiscal instability and labour market rigidity. Examining the crisis channels of the financial and euro crisis shows that fiscal instability, labour market rigidity and also business cycle synchronization cause large losses during the crises. Net migration helps to dampen these shocks. While openness is beneficial during the pre-crises period, it cannot help dampen the shock of the crises.

Keywords Economic growth · Synthetic control methods · Structural time series models · Optimum currency areas · Euro area

JEL Classification C22 · C23 · N14 · F45 · E65

✉ Peter Dreuw
peter.dreuw@uni-osnabrueck.de

¹ University of Osnabrueck, Rolandstr. 8, 49069 Osnabrueck, Germany

1 Introduction

Ever since the founding of the European Economic and Monetary Union (EMU) in 1990, the euro as a single common currency for the EMU members is controversially discussed by economists. The controversy arose because the literature on optimum currency areas (OCA) pioneered by Mundell (1961), McKinnon (1963) and Kenen (1969) does not clearly point out how high costs and benefits due to a membership in a common currency area are. Further, it is controversial whether the nineteen members of the euro area sufficiently meet the criteria and requirements stated in the OCA literature to minimize the costs that arise in common currency areas and what the consequences are of not fulfilling these criteria. By now, empirical research on costs and benefits from the euro is mainly about whether the EU and EMU member states satisfy the OCA criteria or not. Overall, the OCA literature lacks on a formal universal theory that clearly identifies the costs and benefits of a common currency adoption and the empirical literature rarely indicates whether a euro adoption is associated with costs or benefits for a particular country. Thus, a meaningful measure of an ex-post benchmark has the potential to shed new light on the euro debate.

More recently, several empirical studies by Fernandez & Garcia Perea (2015), Verstegen et al. (2017), Lin & Chen (2017), Gomis-Porqueras & Puzello (2018) and Gabriel & Pessoa (2020) have attempted to provide such an ex-post benchmark by applying the synthetic control method (SCM) by Abadie & Gardeazabal (2003) and Abadie et al. (2010, 2015) (AGDH hereinafter). However, the evidence by applying this method is not overly consistent, leading to the conclusion that the method is not too robust to different settings of the adjustment screws such as changes in the donor pool, selection of covariates and the use of pre-intervention outcomes as separate predictors. Although the empirical approach of all of the above studies is very reasonable, we do not know which of the approaches we can trust or how large the euro net effect really is for each EMU member, since the selection of the donor pool and covariates is based on more or less subjective measures and these choices change the outcomes.

Recently, a new SCM based on structural time series models (STMs) and cointegration has been developed by Harvey & Thiele (2021) (HT hereinafter). The approach of the HT-SCM is to identify control series that are moving to the target series prior the intervention, i.e., the euro adoption. When modelling target and control series in a multivariate STM and restrict them to share a common trend, the synthetic control is obtained by forecasting the common trend for the target series to the post-intervention period. The attraction of this method is, that no covariates are needed, time series dynamics such as structural breaks (e.g., the 2008/2009 financial crisis) can be easily incorporated into the model, and the selection of the control group can be based on cointegration - i.e., the long-term co-movement of two series - as an objective measure. The presence of cointegration can be tested using the test of Kwiatkowski et al. (1992), which is shown to be powerful for small time samples by HT. Plausibility considerations to avoid cointegration by chance in a short time sample can also be taken into account when selecting the control group. Thus, the method is able to solve several issues which occur in the AGDH-SCM studies so far.

This paper aims to contribute to the existing synthetic control studies on the euro net effect by providing new evidence for eight early euro adopters (Austria, Belgium,

France, Germany, Greece, Italy, the Netherlands and Spain) and for the EMU aggregate (EA19) by applying the HT-SCM. Once the synthetic controls are estimated, the OCA literature is related to the estimated gaps, and panel regressions are used to test the hypotheses about the effect of the derived main drivers of gains and losses. I find that the euro net effect on GDP per capita is positive for Greece and Spain over the period 1999–2007, the Netherlands gain a bit, Austria and Belgium are only hardly affected, while France and Germany lose over this period. Italy is a clear loser of the euro and loses from 1999 to 2019, as is France, but its losses are not as drastic as Italy's. The financial and euro crises have a massive impact on losses from the euro adoption, so that after the financial crisis all of the eight EMU members considered (CEA8 hereinafter), with the exception of Germany, lose relative to their synthetic controls. In particular Greece, Italy and Spain are big losers from the euro crisis. Relating the estimated gaps to the main drivers of gains and losses derived from the OCA literature, I find that main drivers of gains are openness, real convergence and labour mobility (which is proxied by net migration). Main drivers of losses are low levels of competitiveness and fiscal instability and labour market rigidity. Examination of the crisis channels of the financial and euro crises shows that fiscal instability, labour market rigidity but also business cycle synchronization cause the severe losses from the crises. Net migration helps to dampen these shocks. While openness is beneficial during the pre-crisis period, it could not help to dampen the shock of the crises.

The remainder of this paper is as follows. Section 2 discusses the related literature. Section 3 presents the empirical method of the HT-SCM, it explains the empirical strategy and describes the used data. Section 4 shows the empirical results of the euro net effect for the CEA8 and the EA19 aggregate and performs regression analysis to relate the OCA literature to the estimated gaps. Section 5 concludes.

2 Related literature

2.1 Common currencies and EMU: theory and empirical evidence

Sharing a common currency comes with the cost of losing an individual monetary policy. Instead of having its own central bank, a common central bank manages the monetary policy for the whole currency area. In particular, the inability to devalue or revalue one's currency in response to asymmetric shocks is seen as a major cost. But there are also substantial benefits of a common currency resulting from improved trading conditions with currency partners (economies of scale, no exchange-rate risk, lower transaction- and information costs, price transparency) and increasing trade to these improved conditions (see De Grauwe (2020) for a comprehensive review on costs and benefits of common currency). The early theories of OCAs consider several adjustment and flexibility requirements that members of a common currency should satisfy in order to be able to adjust for asymmetric shocks and minimize their economic losses from abandoning their individual monetary policy. These requirements include labour mobility (Mundell 1961), wage and price flexibility (ibid.), a high degree of openness (McKinnon 1963), product diversification (Kenen 1969), fiscal integration of

common currency members (*ibid.*) and risk sharing through capital market integration (Ingram 1969, 1973; Mundell 1973).¹

More modern theories focus on the conditions under which a common central bank policy is optimal, even if some of the OCA criteria of the early theories are not satisfied. These conditions are synchronous business cycles (Frankel & Rose 1997, 1998) and synchronous reaction to asymmetric shocks (*ibid.*, Alesina et al. 2003). Moreover, Frankel & Rose (1997, 1998) propose the view that common currency areas are endogenous. The authors argue that sharing a common currency will lead to an increase in economic relations between currency partners, which in turn results in a stronger correlation of their business cycles. As a result, their response to shocks will also become more synchronized. In the end, the common central bank policy may be optimal for all currency members as they converge economically. Thus, even if a monetary union is not optimal *ex ante*, it can be optimal *ex post*. Contrary to this view, Krugman (1993) argues that increasing economic relations lead to specialization, which in turn desynchronizes business cycles and responses to shocks. Thus, the common central bank policy may become more expensive for all currency members in the end.

Numerous empirical studies have been conducted on the criteria and conditions of early and modern OCA theories in the EU and EMU. While it seems to be common knowledge that labour mobility, price and wage flexibility and fiscal integration are (too) low and trade openness and product diversification are (sufficiently) high in EU and EMU member states (see Baldwin & Wyplosz 2019, Sect. 15.5), empirical evidence on the degree of risk sharing is mixed. Bekaert et al. (2017) and Hoffmann et al. (2018) find that integration into financial markets has not increased in the euro area. Ferrari & Picco (2016) find that risk sharing across EMU members is decreasing rather than increasing, while Cimadomo et al. (2017) find that 40% of pre-crisis and 65% of post-crisis shocks are dampened by financial integration among euro area members. Integrating financial markets to share risk appears to be an effective tool, but eurozone members are not integrating deeply enough. The empirical literature on the synchronization of business cycles and shocks (see Campos et al. (2017), Campos & Macchiarelli (2016), Wortmann & Stahl (2016), Fingleton et al. 2015, Pentecôte & Huchet-Bourdon (2012)) concludes that there is a core-periphery pattern in the EMU, first identified in Bayoumi & Eichengreen (1993). The core consists of Germany, Austria, Netherlands, Belgium, France, Luxembourg and Finland and is more synchronized, which makes the core countries (better) suited for the formation of a monetary union than the periphery (Greece, Ireland, Italy, Spain and Portugal). With regard to real convergence processes, the empirical literature finds some evidence of beta convergence in the sense of a catching-up process of the peripheral countries (which are also the relatively poorer countries in EMU) (see, for example, Dreyer & Schmid (2016) or Franks et al.). Economic relations between the currency partners are estimated on the basis of their multilateral trade relations. The empirical evidence on the trade effect in the euro area is also mixed, but differs only in the size of the trade increase due to the introduction of the euro. Rose (2016) examines a meta-analysis on

¹ For a comprehensive review of the OCA literature, see Mongelli (2002, 2008), Broz (2005), Cesarano (2006) Dellas & Tavlas (2009), or Kunroo (2015).

the euro effect on trade and finds that trade between euro partners increased by 8.9% to 12.3%.

There is a relatively small strand in the empirical literature that focuses on ex ante benchmarks of euro adoption in a single country. Bayoumi & Eichengreen (1997), Horvath & Komarek (2003), Horvath (2005), and Frydrych & Burian (2017) attempt to provide ex ante benchmarks by quantifying OCA indices that measure a country's eligibility for eurozone membership based on some of the aforementioned OCA criteria. However, these indices are very sensitive to input data, and minor methodological differences lead to different results. As a result, the indices are hardly considered reliable (*ibid.*, p. 196).

Considered as a whole, the theoretical and empirical literature on OCAs does not provide clear evidence of a negative or positive euro net effect, and ex-ante benchmarks are only hardly reliable. Most studies focus on the OCA criteria and whether or not they are sufficiently met by EMU members. What has been missing from the OCA literature is a reliable empirical measure of how large the gains or losses from joining EMU are for a given country. Empirical research provides only a vague suggestion that peripheral countries may catch up due to convergence processes, while core countries may perform better in times of crises because their response to exogenous shocks is more synchronous and ECB policies to respond to shocks are therefore more optimal for core countries than for peripheral countries. Therefore, a meaningful and reliable measurement of an ex-post benchmark for the success of a euro adoption could open new perspectives on the relative importance of the OCA criteria, and assessing the impact of the euro effect for several countries could have great potential to provide clarity on the euro debate.

2.2 Synthetic control methods to estimate euro net effects

An empirical procedure for ex post evaluation of the euro adoption should give a reasonable answer to the question “what would be the level of prosperity in a country if it had not joined EMU?”. To answer this question, a control is needed for exactly the same country that has not joined the euro. In macroeconomics, however, these exact controls do not exist, so a meaningful synthetic control must be constructed for these types of questions. So far, the most popular and widely used method to construct such a synthetic control is the AGDH-SCM. This method uses a data-driven procedure to construct a synthetic counterfactual as a weighted combination of control series. The algorithm chooses the weights, which sum to one, such that the characteristics (certain covariates) of the synthetic control best match those of the target series during the period before the specific intervention. In other words, the algorithm minimizes the root mean square prediction error of the pre-intervention period with respect to the weights. The development of this methodology, as well as a reasonable length of the post-euro adoption period, prompted a growing number of researchers to conduct empirical studies of the net euro effect for several euro area members.

More recently, the studies by Fernandez & Garcia Perea (2015), Verstegen et al. (2017), Lin & Chen (2017), Puzzello & Gomis-Porqueras (2018) and Gabriel & Pessoa (2020) have used this method to estimate the euro net effect on GDP per capita

for some or all of the twelve early euro adopters. Table 1 summarizes their results. For some countries the estimates of the synthetic controls are quite similar (Italy, Portugal, France) but for others they are not very consistent (Germany, Austria, Belgium, Netherlands). Overall, there is some noticeable variation within the results and we do not know which model or which study to trust (more). As a result, there is no robust evidence about winners and losers from the euro adoption in these studies.

There are some reasons that could explain the different results among the studies:

- a) The use of the "donor pool" of potential control countries varies across studies and the selection of the donor pool is based on more or less subjective plausibility considerations. While Lin & Chen (2017) restrict the potential control countries to EU members or to geographically close countries, Fernandez & Garcia Perea (2015), Verstegen et al. (2017) and Gabriel & Pessoa (2020) restrict the potential control countries to OECD members that are economically close to the country of interest. Puzzello & Gomis-Porqueras (2018) restrict the donor pool to all countries in the world that are economically close to the country of interest.
- b) The covariates to explain GDP per capita differ among the studies.

To a): The data driven algorithm tries to match the country of interest as closely as possible to a weighted average of the countries in the donor pool in the pre-intervention period. Clearly, the control group chosen is different if the donor pool is different, and the synthetic control is more or less based on these subjective measures, even if the actual selection of the control group is data-driven. Hence, the selection of the donor pool should be based on an objective measure to overcome the problem of subjectivity. To b): The data-driven procedure for selecting weights also depends on the relative importance on the covariates in explaining GDP per capita. If the covariates are different, then the relative importance and hence the estimated weights may also be different.

If there are differences between studies in both (a) and (b), the differences in estimated synthetic controls in the post-intervention period may be even larger. This point is also made by Ferman et al. (2020) who focus on "specification searching opportunities" and discuss that different specifications of pre-intervention periods and covariates can lead to substantially different synthetic controls. However, because the algorithm always estimates the weights so that the synthetic control is as close as possible to the target series in the pre-intervention period, it is possible that the synthetic controls in the pre-intervention period are very similar.

Kaul et al. (2015) give another reason why the results might differ. They show that when all pre-intervention outcomes are used as separate predictors, all other covariates become irrelevant and only the pre-intervention fit of the synthetic control is optimized. They also show that the estimation of the synthetic control can change drastically if the number of pre-intervention outcomes as predictors is restricted. Therefore, the post-intervention synthetic control may differ even if the donor pool and the covariates are the same. While Lin & Chen (2017) do not use all pre-intervention outcomes as predictors, it is not clear from the description of the other studies whether any, some or all pre-intervention outcomes are used as separate predictors.

In addition to these issues related to the AGDH-SCM, there are some general issues that may arise when using synthetic control methods to assess the euro net effect

Table 1 Overview about euro net effect studies so far

	Fernandez & Garcia Perea (2015)	Verstegen et al. (2017)	Lin & Chen (2017)	Puzzello & Gomis-Porqueras (2018)	Gabriel & Pessoa (2020)
Estimation sample	1992–2013	1986–2014	1991–2013	1971–2007	1970–2007
Date of Intervention	1999	1999	1999, 2001 (Greece)	1995	1999, 2001 (Greece)
Synthetic controls for EMU members	Finland, Ireland , Belgium, Portugal, Italy, Germany, Netherlands, Austria, France, Spain, Greece	Austria, Belgium, Finland, Ireland , Netherlands, Italy, France, Germany, Greece, Portugal, Spain	Finland, France, Germany, Italy, Netherlands, Ireland , Luxembourg , Portugal, Austria, Belgium, Spain, Greece	Ireland, Belgium, France, Germany, Italy, Netherlands	Ireland, France, Germany, Italy, Portugal, Austria, Belgium, Finland, Greece, Netherlands, Spain
Donor pool	11 OECD members that are not EU members	14 OECD members (some of them are EU members)	24 EU members or geographically close to EU	Individual restricted donor pool for all euro members consisting of 5 to 9 countries of the world	14 OECD members (some of them are EU members)
Positive net effect	Finland	Austria, Belgium, Finland, Netherlands	Finland	Ireland	Ireland
Negative net effect	Belgium, Portugal, Italy	Italy	France, Germany, Italy, Netherlands Portugal	Belgium, France, Germany, Italy	France, Germany, Italy, Portugal,
No net effect	Germany, Netherlands, Austria, France	France	Austria, Belgium	Netherlands	Austria, Belgium, Finland, Greece, Netherlands, Spain

Table 1 (continued)

	Fernandez & Garcia Perea (2015)	Verstegen et al. (2017)	Lin & Chen (2017)	Puzzello & Gomis-Porqueras (2018)	Gabriel & Pessoa (2020)
Mixed net effects for different periods	Spain (+ up to 2011,-until 2012), Greece (+ up to 2011,-until 2012)	Germany (0 up to 2007, + until 2008), Greece (+ up to 2009,-until 2010), Portugal (+ up to 2009,- until 2010) Spain (+ up to 2011,-until 2012	Spain (+ up to 2004,-until 2005), Greece (+ up to 2009,-until 2010),		

Bold markers mean that the synthetic control is not adequate and there is little confidence in the results of these models. The results of these models are not analysed in the corresponding studies

(or other macroeconomic issues). For example, an implicit and inherent assumption of any synthetic control method is that the difference between the target series and its synthetic control series in the post-intervention period is entirely driven by the intervention. However, it is possible that countries will be affected by country-specific shocks after the introduction that are not due to this intervention (e.g., Germany and France were affected by the “dot-com” crisis, or early 2000s recession, much more than other EMU-members and control countries) and the synthetic control is not able to distinguish whether this shock is caused by the intervention or by another event. Conversely, a control country could be affected by an individual shock in the post-intervention period. Since the synthetic control is created with some weight on that country, it may also be biased. If a synthetic control method cannot capture these country-specific shocks in some of the control series during the post-intervention period, these series should be excluded from the donor pool so that the synthetic estimate is not biased due to these country-specific shocks. Another issue to consider is global shocks such as the financial crisis, oil price shocks, or the rise of China. Since these shocks apply to both the target series and the control group, synthetic control methods implicitly capture these effects as they affect the control group.

Taking all of the above problems into account, the AGDH-SCM provides reasonable measures of the performance of the various EMU members after euro adoption, but the rather mixed results from the various studies suggest that it is sensitive to the donor pool and the choice of covariates. A new method proposed by HT (2021) based on structural time series models (STMs) and cointegration is able to solve the major issues of the AGDH-SCM. In this approach, the synthetic control is implicitly estimated by a multivariate balanced growth STM in which all the series in the model are described with the same trend and only individual differences in height from it. Cointegration serves as an objective criterion for the selection of the control group and is the basis for the validity of the balanced growth STM. The rational of this approach is that

countries that are cointegrating share a common trend, and predicting this common trend in the post-adoption period is the synthetic control. Plausibility considerations, such as similar economic development, could also be taken into account when selecting the control group. Weights that sum to one are implicitly given by variance–covariance relations. No covariates are needed to estimate the common trend. An important feature of the HT-SCM approach is that time series dynamics are tracked in the model and common structural breaks like the financial crisis in 2008–2009 as well as country-specific structural breaks (e.g., the German reunification) can be easily accounted for both in the pre and post adoption period. But like the AGDH-SCM, the HT-SCM is unable to distinguish between the effect of an intervention and other country-specific shocks that occur at the same time in the post-intervention period. For this study, as for all other synthetic control studies so far, we must assume that differences of the synthetic controls and the target series are entirely due to the euro adoption.

The HT-SCM approach has two major weaknesses compared to the AGDH-SCM approach. First, estimated weights can be negative. Second, a large number of dummy variables need to be included in the model to capture outliers and to model the intervention of the euro adoption. While the first issue can be solved by removing the corresponding control series (this series does not contribute anything to the synthetic control), the second issue mainly concerns the length of the post-intervention period and the detection of outliers. The only way to deal with the number of parameters is to model the intervention with as few parameters as necessary, but as many as needed, and to perform a careful outlier analysis. Robustness analyses to address these issues (see Sect. 4.3) show that the HT-SCM estimates are robust to the number of parameters and modelling outliers. Thus, the HT-SCM can address several problems of the AGDH-SCM studies discussed above, while its weaknesses appear to be minor.

The aim of this paper is to contribute to the existing synthetic control literature on the euro effect by applying the new method of HT (2021). After estimating the synthetic controls, hypotheses about the main drivers of the gains and losses from EMU membership are tested with a panel regression analysis. New evidence on both the estimated synthetic controls and the main drivers of gains and losses could clarify the controversial debate over the adoption of the euro.

3 Methodology

3.1 Structural time series models and synthetic controls

The basic idea of STMs is to decompose a time series into several components that are not directly observable but can be directly interpreted, such as trend, seasonality, and cycle.² The basic univariate STM for the empirical research of this paper consists of

² For a comprehensive discussion of univariate and multivariate structural time series models, see Harvey (1989) and Harvey & Koopman (1997).

a decomposition of the GDP per capita series of any country into a local linear trend model, which is

$$y_t = \mu_t + \varepsilon_t, \quad \varepsilon_t \sim N(0, \sigma_\varepsilon^2), \quad t = 1, \dots, T, \quad (1)$$

$$\mu_{t+1} = \mu_t + \beta_t + \xi_t, \quad \beta_{t+1} = \beta_t + \zeta_t, \quad \xi_t \sim N(0, \sigma_\xi^2), \quad \zeta_t \sim N(0, \sigma_\zeta^2) \quad (2)$$

where y_t is annual GDP per capita in logs, μ_t is the level component, β_t is the slope component of the trend and ε_t , ξ_t and ζ_t are mutually independent disturbance terms with zero mean and variances σ_ε^2 , σ_ξ^2 and σ_ζ^2 respectively. If $\zeta_t = 0$, the trend component is called a random walk with drift, while if $\xi_t = 0$ the trend component is called an integrated random walk. Equation (1) is called observation equation. The equations for the level and slope in (2) together form the state equations.

If N univariate series are modelled in this way, the univariate STM can be generalised to the following multivariate STM

$$\begin{aligned} y_t &= \mu_t + \varepsilon_t, \\ \mu_{t+1} &= \mu_t + \beta_t + \xi_t, \quad t = 1, \dots, T \\ \beta_{t+1} &= \beta_t + \zeta_t, \\ \varepsilon_t &\sim N(0, \Sigma_\varepsilon), \quad \xi_t \sim N(0, \Sigma_\xi), \quad \zeta_t \sim N(0, \Sigma_\zeta) \end{aligned} \quad (3)$$

where y_t is an $N \times 1$ vector of GDP per capita series in logs, μ_t is an $N \times 1$ vector of individual trend components of each series, β_t is an $N \times 1$ vector of individual slope components of the individual trends and ε_t , ξ_t and ζ_t are $N \times 1$ vectors of disturbances with zero mean and $N \times N$ variance covariance matrices Σ_ε , Σ_ξ and Σ_ζ respectively. ε_t , ξ_t and ζ_t are assumed to be uncorrelated to each other. The advantage of this multivariate model is that if Σ_ε is non-diagonal the disturbances ε_t of the series are correlated, leading to efficiency gains in estimation and forecasting. The correlation between disturbances can be seen as a control for the general economic environment for all series. Balanced growth means that the N series have a stable relationship over time, implying a stationary difference and hence cointegration between each pair of the series. Thus, a balanced growth model is a special case of (3) where $\text{rank}(\Sigma_\xi) = 1$, so that it can be simplified to

$$\begin{aligned} y_t &= i\mu_t + \bar{\mu} + \varepsilon_t, \quad t = 1, \dots, T \\ \varepsilon_t &\sim N(0, \Sigma_\varepsilon) \end{aligned} \quad (4)$$

where μ_t is a univariate local linear trend as in (2), i is an $N \times 1$ vector of ones and $\bar{\mu}$ is an $N \times 1$ vector of unrestricted constants that captures the height differences of the series. The elements of ε_t are assumed to be correlated so that Σ_ε is non-diagonal.

HT (2021) show how the balanced growth model can be used to construct a synthetic control and assess the impact of an intervention such as the euro adoption. Assume, that the first series in (4), y_{1t} , is affected by the intervention of the euro adoption, while all other series are independent of this intervention. The euro adoption may affect this

series in order to change its level or its slope. The empirical analysis of this paper concentrates on potential level changes due to the intervention. Such an intervention can be modelled as a combination of pulse and step dummy variables in the target series equation. After modelling the intervention, the first series in (4) becomes

$$y_{1t} = \mu_t + \bar{\mu}_1 + \sum \lambda_p d_{pt} + \sum \lambda_s d_{st} + \varepsilon_{1t}, \quad t = 1, \dots, T. \quad (5)$$

where pulse dummies are denoted with subscript p and step dummies are denoted with subscript s . λ_p and λ_s are their parameters. Pulse and step dummies are defined as

$$d_{pt} = \begin{cases} 0, & t \neq \tau \\ 1, & t = \tau \end{cases}, \quad t = 1, \dots, T, \\ d_{st} = \begin{cases} 0, & t < \tau_l \\ 1, & \tau_l \leq t \leq \tau_u \\ 0, & t > \tau_u \end{cases}, \quad t = 1, \dots, T, \quad (6)$$

where τ is a single time point and τ_l and τ_u are the lower and upper bounds of a time range. Pulse dummies capture effects only at a single time point, so they can be used to model temporary intervention effects, whereas step variables capture effects at more than one time point, so they can be used to model intermediate intervention effects if $\tau_u < T$ or permanent intervention effects if $\tau_u = T$.

Step and pulse dummy variables can also be included in the model to account for temporary and intermediate country-specific deviations from the common trend or individual structural breaks. A structural break can be modelled with a step variable with $\tau_u = T$. To allow for country-specific deviations, a more general balanced growth model is formulated as

$$y_{it} = \mu_t + \bar{\mu}_i + \sum \psi_{ip} d_{ipt} + \sum \psi_{is} d_{ist} + (1i = 1) * \left(\sum \lambda_p d_{pt} + \sum \lambda_s d_{st} \right) + \varepsilon_{it}, \quad (7) \\ i = 1, \dots, N, \quad t = 1, \dots, T.$$

Again, the subscripts p and s denote pulse and step dummies in the form of (6) and λ_p and λ_s are the corresponding intervention parameters, while ψ_{ip} and ψ_{is} are the corresponding parameters to account for deviations from the common trend of country i . The first equation in (7) corresponds to an EMU member while the other $N - 1$ equations correspond to the control countries. The indicator variable ensures that the intervention modelling only appears in the first equation.

Also consider common structural breaks that can occur when all series in the balanced growth model are affected by the same shock, such as the 2008–09 financial crisis. These types of shocks affect the common trend. We can capture such shocks by adding pulse and step dummy variables as defined in (6) to the trend equation, which then becomes

$$\mu_{t+1} = \mu_t + \beta_t + \sum \varphi_{bp} d_{bpt} + \sum \varphi_{bs} d_{bst} + \xi_t \quad (8)$$

φ_{bp} and φ_{bs} are the parameters corresponding to the pulse and step dummies. Subscript b means that the variable occurs in the trend equation, while the other subscripts are analogous to those in Eq. (6).³

The basic model for the empirical analysis of this paper is the balanced growth model given in (7), where the common trend is modelled as in (2) or (8), depending on whether there are any structural breaks. To estimate the model, it must be put in state space form. The regression components, i.e., the dummy variables are included in the state vector. Then the Kalman Filter and Smoother are applied to estimate the unobserved components. The hyperparameters corresponding to the variances and covariances of the error components are estimated using maximum likelihood. The likelihood function in this regard can be computed directly by the Kalman Filter equations (see Durbin & Koopman 2012, Sect. 2.10).

Weights that sum to one are implicitly given by variance–covariance relations of the series. If $\mathbf{w}_1$ denotes the $(N - 1) \times 1$ vector of weights to construct the synthetic control y_{1t}^c , they are given by

$$\begin{aligned} \mathbf{w}'_1 &= \boldsymbol{\beta}'_1 + s \mathbf{i}' \boldsymbol{\Sigma}_{\varepsilon(-1)}^{-1}, \\ \boldsymbol{\beta}'_1 &= \boldsymbol{\sigma}'_{\varepsilon} \boldsymbol{\Sigma}_{\varepsilon(-1)}^{-1}, \\ \boldsymbol{\sigma}_{\varepsilon} &= \text{cov}(\varepsilon_{1t}, \boldsymbol{\varepsilon}_{(-1)t}), \\ s &= (1 - \boldsymbol{\beta}'_1 \mathbf{i}) / \mathbf{i}' \boldsymbol{\Sigma}_{\varepsilon(-1)}^{-1} \mathbf{i} \end{aligned} \quad (9)$$

where $\boldsymbol{\beta}_1$ and $\boldsymbol{\sigma}_{\varepsilon}$ are $(N - 1) \times 1$ vectors and $\mathbf{i}$ is an $(N - 1) \times 1$ vectors consisting of ones. $\boldsymbol{\Sigma}_{\varepsilon(-1)}$ is an $(N - 1) \times (N - 1)$ matrix and the subscript (-1) means that the components referring to the first row are omitted in the respective vector or matrix. A shortcoming of the HT-SCM is that “there is nothing to prevent some of the weights being negative” (HT 2021, p. 4). In relation to the AGDH-SCM, the synthetic control could be constructed directly as a weighted combination of the control series. However, constructing the synthetic control this way ignores common and country-specific structural breaks and country-specific deviations from the common trend. A more accurate measure of the synthetic control can be constructed from the output of the Kalman Smoother,⁴ which is given by

$$\begin{aligned} \hat{y}_{1t}^c &= \hat{\mu}_t + \hat{\bar{\mu}}_1 + \sum \hat{\psi}_{1p} d_{1pt} + \sum \hat{\psi}_{1s} d_{1st} + \hat{\varepsilon}_{1t}, \\ \hat{\varepsilon}_{1t} &= \hat{\boldsymbol{\beta}}' \left(\hat{\mathbf{y}}_{(-1)t}^- - \hat{\bar{\mu}}_{(-1)} - \mathbf{i} \hat{\mu}_t \right) \\ \hat{\mathbf{y}}_t^- &= \hat{\mathbf{y}}_{it}^- = y_{it} - \sum \hat{\psi}_{ip} d_{ipt} - \sum \hat{\psi}_{is} d_{ist}. \end{aligned} \quad (10)$$

³ Structural breaks could also be modelled in the observation equations. To avoid possible multicollinearities between euro intervention effects and level breaks occurring in the same period, I decide to model structural breaks in the trend equation. Furthermore, note that modelling structural breaks in the observation equation is done using step dummies, while modelling structural breaks in the trend equation is done using pulse dummies shifted one period into the past (see Commandeur & Koopman 2007, p. 78ff.). The term $\sum \varphi_{bs} d_{bst}$ is only needed if structural breaks suppress or push the trend for more than one period.

⁴ Each component estimated with the Kalman Smoother is marked with a hat.

Note that if there are country-specific outliers and structural breaks, the construction of the synthetic control is not based on the original series of the control group, $y_{(-1)t}$, but on the model prediction adjusted for these deviations, $\widehat{y}_{(-1)t}$. When there are no structural breaks or country-specific deviations, as in the model (4) + (5), the construction of the synthetic control using (10) simplifies to

$$\begin{aligned}\widehat{y}_{1t}^c &= \widehat{\mu}_t + \widehat{\mu}_1 + \widehat{\varepsilon}_{1t}, \\ \widehat{\varepsilon}_{1t} &= \widehat{\beta}' (y_{(-1)t} - \widehat{\mu}_{(-1)} - i\widehat{\mu}_t)\end{aligned}\quad (11)$$

The summations of the intervention parameters, $\sum \widehat{\lambda}_p d_{pt} + \sum \widehat{\lambda}_s d_{st}$, in (5) and (7) is the estimate of the gap between the original target series, y_{1t} , and its synthetic control $\widehat{y}_{1t}^c$ at time $t \geq \tau$. The intervention modelling is chosen so that $\sum \widehat{\lambda}_p d_{pt} + \sum \widehat{\lambda}_s d_{st} \approx y_{1t} - \widehat{y}_{1t}^c$ with as many parameters as necessary but as few as possible.

3.2 Diagnostics checking and empirical remarks

To ensure that a model is well specified, several diagnostic tests must be performed on standardized innovations (i.e., the one-step-ahead-prediction errors of the Kalman Filter) as well as an outlier analysis using standardized smoothed state residuals and standardized smoothed observation residuals (irregular hereinafter). According to Harvey & Koopman (1992) and Harvey et al. (1998) outliers in standardized irregular residuals indicate series-specific temporary effects that can be captured with a pulse dummy in the respective observation equation. In the case of the balanced growth model, an outlier in standardized irregulars indicates a strong deviation of the respective time series from the common trend. Outliers in standardized smoothed state residuals indicate a common structural break representing a level change in the common trend.

In the balanced growth model, there is no disturbance term that directly indicates structural breaks in a particular series,⁵ since the standardized smoothed state residuals indicate only common structural breaks, and a structural break in a series accounts for only a weighted portion of the total residual. If such a weight is small, the structural break in the residual cannot be traced; if such a weight is large, the residual may indicate a common structural break even though the outlier is generated by only a single series. In addition, a series-specific structural break will increase the variance of the irregulars of that series, so the standardized irregulars may not indicate outliers. However, the standardized irregulars will have a structure: Suppose there is a series-specific negative structural break at time τ while there are no other series-specific outliers. Then, standardized irregulars will fluctuate around a positive mean in $t = 1, \dots, \tau$ and then around a negative mean in $t = \tau + 1, \dots, T$. The weighted sum of the positive mean in $t = 1, \dots, \tau - 1$, and the negative mean in $t = \tau, \dots, T$, will be close to zero as the standardized irregular are $\sim N(0, 1)$.

⁵ In a balanced growth model, a series-specific structural break is a structural deviation from the common trend.

As mentioned by Commandeur & Koopman (2007, p. 96), outlier analysis is not about mindlessly adding every single outlier discovered. Especially adding structural breaks should be confirmed by a theory or a historical event concerning the possible cause of the structural break. In this regard, common structural breaks are justified in times of global economic crises like the financial crisis in 2008–2009, the oil crises in 1973 and 1979 or the early 1980s and early 1990s recession. Individual structural breaks or temporary outliers should be confirmed by country-specific recessions or booms. Moreover, outlier analysis should be a tool to confirm the results of visual analysis, and visual analysis should be a tool to confirm the results of outlier analysis. Therefore, in the empirical analysis of this paper, the outlier analysis is always performed together with the visual analysis.

Note that adding a pulse dummy to the observation equation actually pushes the respective series to the common trend for the respective time point, while the observation for which the pulse dummy is added has no weight in the construction of the common trend for that time point (This follows from the fact that adding a pulse variable is the same as replacing the respective observation with a missing value). Adding a step variable to an observation equation pushes the series to the level of the common trend for the length of the step. This is an important feature, especially in the post-adoption period, as it allows the common trend to be constructed without country-specific strong deviations (This is not possible with an AGDH-SC). The drawback of this feature is that adjustments in the post-adoption period can only be added if there are at least three controls. With only two controls, the common trend in the post-adoption period is made of only two series. If, a pulse or step variable is added to an observation equation of one of the series, the common trend in that time period consists of only one series and thus the synthetic control for those time points consists of only one series.

Model diagnostics such as normality, serial correlation and heteroscedasticity are checked with standardized innovations, as recommended by Harvey et al. (1998, p. 111f.). Appropriate standardized innovations for diagnostics testing in a model with regression components are the generalised least squares (GLS) residuals (Harvey (1989, p. 386f.)). The GLS residuals are the standardized innovations that are produced when the Kalman Filter is applied to y_t minus the estimated regression part of the model.⁶ Therefore, it is required that the Kalman Smoother estimates of the structural breaks modelled in the trend equation in (8) are reformulated as variables affecting the observations y_{it} . The reformulation is the cumulative sum of the original dummy variable (see Commandeur & Koopman (2007, p 78ff.)):

$$\begin{aligned} d_{bpt} &\rightarrow d_{bpt}^y = \sum_{t=1}^T d_{bpt}, \quad t = 1, \dots, T, \\ d_{bst} &\rightarrow d_{bst}^y = \sum_{t=1}^T d_{bst}, \\ \sum_{t=1}^T d_{bpt}, \quad \sum_{t=1}^T d_{bst} &= 0, \quad \text{when } t = 1 \end{aligned} \quad (12)$$

⁶ In other words, the Kalman Filter is applied to the stochastic part of the model. Instead of subtracting the regression part of y_t and then applying a Kalman Filter to a model for $\hat{y}_t^*$ as given in (12) and (13), one could apply a second Kalman Filter for y_t on the original model (7) + (8) and treat the estimated regression coefficients as fixed and known. Actually, this is the procedure I use.

To produce the GLS residuals, a second Kalman Filter is applied to (4) but with $\hat{y}_t^*$ instead of y_t . $\hat{y}_t^*$ is defined as

$$\begin{aligned} \hat{y}_t^* &= \hat{y}_{it}^*, \quad i = 1, \dots, N, \quad t = 1, \dots, T, \\ \hat{y}_{it}^* &= y_{it} - \sum \hat{\varphi}_{bp} d_{bpt}^y - \sum \hat{\varphi}_{bs} d_{bst}^y - \sum \hat{\psi}_{ip} d_{ipt} - \sum \hat{\psi}_{is} d_{ist} \\ &\quad - (1|i = 1) * \left(\sum \hat{\lambda}_p d_{pt} + \sum \hat{\lambda}_s d_{st} \right) \end{aligned} \quad (13)$$

and the parameters $\hat{\varphi}_{bp}$, $\hat{\varphi}_{bs}$, $\hat{\psi}_{ip}$, $\hat{\psi}_{is}$, $\hat{\lambda}_p$ and $\hat{\lambda}_s$ are estimates of the Kalman Smoother of the model in (7) with the trend component as in (8).

Finally, it should be noted that model diagnostics based on the GLS residuals can only be performed in the non-diffusion phase of the Kalman filter. The Kalman filter is initialised in the same way for all models. The trend component μ_t is initialized by setting $\mu_0 = 0$ so that $\bar{\mu}$ is subject to a constraint and contains only $N - 1$ free parameters as noted by Carvalho & Harvey (2005). Both the slope component β_0 as well as $\bar{\mu}$ are initialized with a diffuse prior. Thus, the diffuse phase of the Kalman Filter lasts for the first two periods and the model diagnostics can be checked with the GLS residuals in $t = 3, \dots, T$.

3.3 Data and empirical strategy

To analyse the euro net effect on euro area economies, annual data on real GDP per capita at constant 2010 prices from 1970 to 2019 are used and transformed into logarithms. The data stem from the World Bank's world development indicators database and are available for 122 countries for the mentioned time series length. Among these 122 countries are 14 countries that have adopted the euro.

At first, the donor pool of potential control countries consists of the 108 non-euro area countries, but in a second stage it is restricted to fulfil three control group criteria: cointegration, small variance in the difference between the target and control countries in the pre-intervention period and economic similarity. Therefore, only a small fraction of the 108 potential control countries remains as suitable control countries for the various models. The net effect of the euro on annual real GDP per capita is analysed for eight of the twelve countries that adopted the euro early (Austria, Belgium, France, Germany, Greece, Italy, the Netherlands and Spain) as well as for the EA19 aggregate.

Finland is excluded from the analysis because of its major crisis around 1990, which has made it impossible to construct a suitable model for evaluating the euro effect. For Portugal, the criteria for the control group are too restrictive, so that no suitable control group can be found. The same applies to Ireland and Luxembourg, both of which are highly individual in terms of their economic development. Ireland is considered a "Celtic tiger" due to its rapid growth since the mid-1990s, while Luxembourg is both one of the richest and one of the smallest countries in the world. The fact that no suitable control group could be found for Ireland and Luxembourg in the studies mentioned in Table 1 and in the present study could reflect their individual growth paths. The euro adoption date for all CEA8 is 1999, except for Greece, which officially joined in 2001. Despite this small temporal difference, as in Versteegen et al. (2017) and Fernandez

& Garcia Perea (2015), 1999 is chosen as the intervention date of euro adoption for Greece. This is justified by the fact that Greece had to fulfil the criteria for joining EMU for at least two years before officially joining and therefore had to adapt to the economic realities of the EMU system in 1999. Nonetheless, a model in which Greece adopts the euro in 2001 is estimated to test robustness. Due to economic turbulence and some economic differences between the control and target countries in the 1970s, the estimation sample for none of the models starts in 1970, yet the largest possible time series length is chosen for all models.

The empirical strategy for assessing the euro net effect is closely related to that proposed by HT (2021). First, cointegration between the target series and potential control countries is tested in the whole sample. To test for cointegration the test for level stationarity by Kwiatkowski et al. (1992) – KPSS hereinafter – is used by applying it to the difference between the target and control countries for the pre-intervention period. The difference is also examined graphically to ensure that the series are cointegrating and to look for distortions between the series that should be incorporated into the model by adding appropriate dummy variables. Another statistical criterion for selecting the control group is that the variance of the differences between the target and control groups should be relatively small because "controls with relatively large variances tend to be downweighted" (see HT 2021, p. 4). Consequently, including a control country with a relatively large variance in the model does not contribute much to the estimation of the synthetic control, while requiring more parameters to be estimated. Therefore, controls with relatively large variances are removed from the donor pool. Note that the difference between a perfect control and the target is zero in the pre-intervention period, implying that the control country is exactly equal to the target, with the only difference being that it is not affected by the intervention. Thus, the smaller the variance of the difference between control and target, the more suitable it is as a control, given that it is cointegrating. Plausibility considerations are also applied to the selection of the control group. In this context, only controls with a similar economic development in terms of GDP per capita are selected. To get an idea of the economic similarity, the mean difference in GDP per capita between the target country and the control group in the pre-euro period and the mean value of GDP per capita in the pre-euro period of the control country are examined. If the mean difference in GDP per capita is close to the mean of the control country, a control is considered economically similar. Once an appropriate control group is selected, KPSS tests for level and trend stationarity are performed to determine whether the series are best described by an integrated random walk, a random walk with drift, or with stochastic components for trend and slope.

The question is how to model the intervention. HT (2021) consider an approach based entirely on estimation under the balanced growth model. In this context, a model is estimated in which the values of the target series in the post-intervention period are replaced by missing values (this is the same as modelling impulse dummies for each post-intervention period), and the difference between the original target series and the synthetic control estimate of this model provides guidance on how to model the intervention. A balanced growth model in which target series values are replaced by missing values after the intervention is called a BGMTV (balanced growth with missing target values) model, and a synthetic control estimated with such a model is called

a BGMTV synthetic control. Based on the difference between the original target series and the BGMTV synthetic control in the post-intervention period, the intervention is modelled with as few parameters as possible and with as many parameters as necessary, so that whenever the difference indicates an intermediate or permanent step, these steps are estimated with a single step variable instead of a bundle of pulse dummies.

Adding outliers and common and country-specific structural breaks affects the estimation of the common trend as well as the weights and $\hat{\beta}$. Hence, to get an idea of how to model the intervention, the underlying BGMTV model should include all relevant step and pulse variables to account for these outliers. Therefore, a preliminary outlier analysis is performed with a BGMTV model to check whether the model is well specified. Only after ensuring that the BGMTV-model captures all the relevant structural breaks and outliers, the difference between the target and the estimated BGMTV synthetic control is explored to model the intervention. After estimating a full model with original target values in the pre- and post-adoption period that includes the relevant outliers and structural breaks, as well as the intervention modelling, a final check of outliers ensures that the model is well specified.

Finally, the model diagnostics are tested based on the GLS residuals for the target series in the full model. The Jarque–Bera (JB) test is used to test normality. Autocorrelation is tested with the Ljung-Box-Q(p,l)-test, which tests the Q-statistic for the first p autocorrelations against a $X^2(l)$ -distribution, where $l = p - w + 1$ and w denotes the number of estimated variances belonging to the tested series. For a univariate random walk with drift, $w = 2$ since σ_ε^2 and σ_ξ^2 must be estimated. Heteroscedasticity is tested with the H(h)-test (see Harvey 1989, p. 259f.) which compares the residual variance of the first h residuals with that of the last h , where h is the nearest integer to $(T-d)/3$ and d being the number of diffuse initial elements ($d = 2$ for this study), and the null hypothesis is that of homoscedasticity. The null is rejected, if $H(h) \geq F_{1-\alpha/2}(h, h)$ or $H(h) \leq F_{\alpha/2}(h, h)$. Alternatively, if $H(h) < 1$, the null is rejected if $1/H(h) \geq F_{1-\alpha/2}(h, h)$. Any estimated model passes the diagnostics if all tests indicate that the null cannot be rejected.

The empirical analysis is performed in R with help of the KFAS-package by Helske (2017).⁷ The Kalman-Filter and Smoother recursions in this package are based on the univariate treatment algorithms given in Durbin & Koopman (2000, 2003, 2012). While the univariate treatment procedure has some computational advantages, it produces cross sectionally uncorrelated innovations. Hence, it is appropriate to apply the univariate tests mentioned above to each univariate standardized innovation series (ibid. 2012, p. 188), rather than applying their multivariate analogues to the multivariate vector of standardized innovations.

⁷ The code used to estimate the models is able to replicate the results in Harvey and Thiele (2021) who apply a different computer package (STAMP by Koopman et al. (2008)).

4 Assessing the impact of the euro adoption

4.1 An instruction for the empirical analysis with the example of Austria

Step 1: Choosing the control group

The first step in building a model is to select an appropriate control group. The selection procedure includes visual analysis and KPSS tests to check for cointegration with the respective EMU member across the whole sample. The cointegration tests are evaluated together with the criteria of small relative variance and economic similarity in terms of a close GDP per capita level in the pre-1999 period. At the outset of the analysis, the full sample period is used, but it may be appropriate to shorten it due to convergence processes and economic turbulence in the mid-1970s or early 1980s.

KPSS tests for the period 1970 to 1998 indicate that the control group would consist of Morocco, the UK and Tunisia. But Morocco and Tunisia are not at all comparable to Austria, since GDP per capita in both countries is about 20 times smaller than in Austria. Table 2 shows the results of the KPSS tests for the period 1976 to 1998 along with the other measures of the control group criteria. For this period, the KPSS tests show that the USA, Denmark, Israel, and the UK should form the control group. According to the cointegration, low relative variance, and economic similarity criteria, this is a much more plausible control group.

Visual analysis of the mean-adjusted series of Austria and its potential control group for the period 1970 to 2017 confirms these results. Figure 1 shows the GDP per capita series of Austria and its potential control countries and their mean-adjusted differences from Austria for this period. It can be clearly seen that Austria converges economically to the potential control countries until 1976, but then cointegrates for

Table 2 Ordered KPSS tests for cointegration for Austria

Name	KPSS-Level	Variance	Mean GDP per capita difference pre 1999	Mean GDP per capita pre 1999
USA	0.0672	0.0006	– 1,939	33,441
Paraguay	0.1049	0.0034	28,201	3,300
Morocco	0.1129	0.0012	29,924	1,578
Denmark	0.1352	0.0008	– 11,065	42,567
Israel	0.1628	0.0008	10,688	20,813
Lesotho	0.1741	0.0031	30,880	621
Nepal	0.1749	0.0012	31,162	340
Turkey	0.2018	0.0029	25,226	6,275
Oman	0.2354	0.0082	17,876	13,626
Bahamas	0.3138	0.0135	5,185	26,317
Uruguay	0.3167	0.0092	24,339	7,162
UK	0.3324	0.0010	5,327	26,175
Puerto Rico	0.3349	0.0015	14,667	16,834

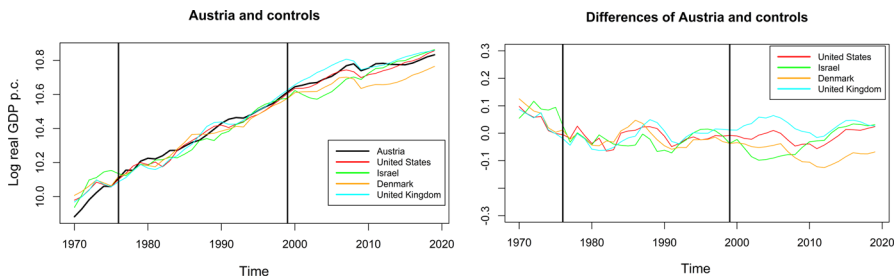


Fig. 1 Austria and control series (left) and differences of Austria and its controls (right). The series are adjusted for the mean prior to 1999

the entire period before the intervention. Taking the results of the KPSS tests and the visual analysis together, the USA, Denmark, Israel and the UK provide an adequate control group.

KPSS tests for trend and level stationarity are useful to get an idea of how to model the trend. Applied to the first difference of a series, the KPSS test for level stationarity indicates whether the trend is deterministic or stochastic. Applied to the original series, the KPSS test for trend stationarity indicates whether the slope is deterministic or stochastic. Table 3 shows the results for Austria and the selected control countries for the period 1976 to 1998. Only for Israel do the tests indicate that the slope is stochastic, but from the visual analysis (see Fig. 1) there is only a vague case for a stochastic slope. Finally, the picture that emerges from the tests and the visual analysis is that the trend is best modelled as a random walk with drift.

Step 2: Building a naïve model

Once we have selected suitable control countries, we can set up the model for the period 1976 to 2019. As a starting point, we create a "naïve" model as in (4), ignoring possible common and country-specific structural breaks as well as country-specific deviations from the common trend. The goal of this naïve model is to get an overview of the model prediction and check if there are any structural breaks or influential

Table 3 KPSS tests for level and trend stationarity for Austria and its controls

KPSS-Test:	Level stationarity	Trend stationarity
H0	Stochastic level	Deterministic trend
Crit. Value (10%)	0.347	0.119
Austria	0.0792	0.0580
USA	0.0583	0.0651
Israel	0.2625	0.1968
Denmark	0.0661	0.0798
UK	0.0611	0.0699

The sample range for the tests is 1976–1998. All tests are carried out with a truncation lag parameter of 2. Bold numbers indicate significance at (at least) the 10% level

observations. Since we do not know how to model the intervention, we estimate a BGMTV model. The resulting visual output of this model is shown in Fig. 2a. Now, the graph of the difference between Austria and its BGMTV synthetic control (shown in Fig. 2c) together with the table of differences (shown in Fig. 2d) gives an indication of how to model the intervention. According to the graph and table of differences, the periods 2001–2005 and 2015–2019 can be modelled with a step variable while the remaining time points need to be modelled with a pulse dummy. Modelling the intervention in this way and estimating the model with the full series of Austria yields the visual output shown in Fig. 2b. The synthetic control in Fig. 2a and b is constructed using (11). The model diagnostics can be found in Table 6 in Sect. 4.3 in the “Austria naïve” row.

Step 3: Checking the model

Checking the model includes testing for heteroskedasticity, normality and autocorrelation with standardized GLS residuals as well as performing an outlier analysis and a visual analysis of the mean-adjusted series for the target and control countries. The obtained estimates of the naïve model are close to the observed differences of Austria to its estimated synthetic control and the heteroskedasticity-test, normality-test and Q(5,4)- and Q(10,9)-test are satisfactory. However, the estimation of hypothetical Austria with the naïve model has some weaknesses. First, the estimated trend more or less ignores the large downshift of the 2008–2009 financial crisis and second, there are some individual trend deviations of Denmark in 1984–1988 and 2009–2019, of the

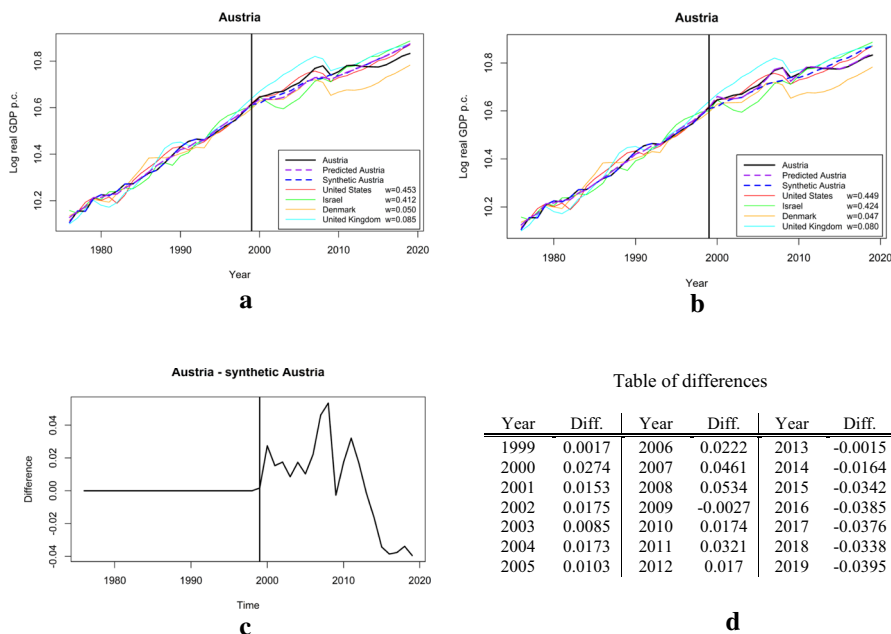


Fig. 2 Visual output of a BGMTV-model **a**, visual output of the full sample estimation **b**, difference of Austria and its BGMTV-model estimate **c** and the table of differences between Austria and its estimated BGMTV-model estimate **d**. Outliers or structural breaks are not captured in the model

USA in 1982 and of Israel in 2001–2007. In particular, the crisis in Israel around 2001, which recovers by 2007, and the much sharper downturn of Denmark in 2009 could affect the estimate of the synthetic Austria. Therefore, based on the visual analysis, a step variable 2009–2019 is added to the observation equation of Denmark and two step variables, 2001–2007 and 2002–2006 are added to the observation equation of Israel.

As mentioned in Sect. 3.3, adding outliers and common and country-specific structural breaks affects the estimation of the common trend as well as the weights and $\hat{\beta}$. Thus, before getting an idea of how to model the intervention based on the difference of the target series to its BGTMV-synthetic control, a preliminary outlier analysis is needed to ensure that all relevant outliers are captured in the model and that the model is well specified. Since we do not know how to model the intervention yet, the preceding outlier analysis is performed with BGTMV-models. In this process, the outlier analysis is performed repeatedly: If one or more plausible outliers are detected, corresponding step or pulse dummy variables are added to the respective observation or trend equation; then the model is estimated; then the model is checked again for plausible outliers; then plausible outliers are added to the model; then the model is estimated again, etc.

Applying this procedure after adding the aforementioned step variables for Israel and Denmark, the standardized smoothed state residuals reveal a consecutive structural break in 2008–2009, which is captured in the model by adding two pulse dummy variables to the trend equation. The standardized irregulars of the several series (step-wise) show outliers for Denmark in 1984, 1985, 1986 and 1987; for Israel in 1976 and 1989; for the USA in 1982 and 1983, for UK in 1980, 1981, 1982, 1983, 1988, 1989 and 1990; and for Austria in 1990, 1991 and 1992, that can be confirmed by the visual analysis. Capturing these outliers by adding appropriate pulse dummies to the respective observation equations and estimating the model again reveals that the parameters for 1991 and 1992 are almost equal and can therefore be combined into a single step variable, 1991–1992. After adding these outliers and structural breaks there are still some weak outliers in standardized irregulars of the USA in the mid-1980s, but adding these to the model can only hardly be justified by the visual analysis, so they are not added.

Step 4: Modelling the intervention and estimating the full model

Finally, all relevant outliers are captured in the model, and we can use the difference of Austria to its BGTMV-synthetic control to get an idea of how to model the euro adoption intervention. The resulting visual output of this model is shown in Fig. 3a. Figure 3c shows the graphical representation of Austria's difference from its synthetic control, and Fig. 3d shows the table of this difference. Figure 3c and d show that the intervention can be modelled with pulse dummies for 2005, 2006, 2007, 2008, 2011, 2012, 2013, 2014 and step dummies for 1999–2000, 2001–2002, 2003–2004, 2009–2010, 2015–2019. Finally, Fig. 3b shows the visual output of the full model containing all relevant dummies to capture common structural breaks, country-specific structural breaks and deviations from the common trend, as well as the intervention modelling. The variables used to construct this full model are listed in Table 4. The model diagnostics for this full model for Austria are reported in Table 5 in Sect. 4.2,

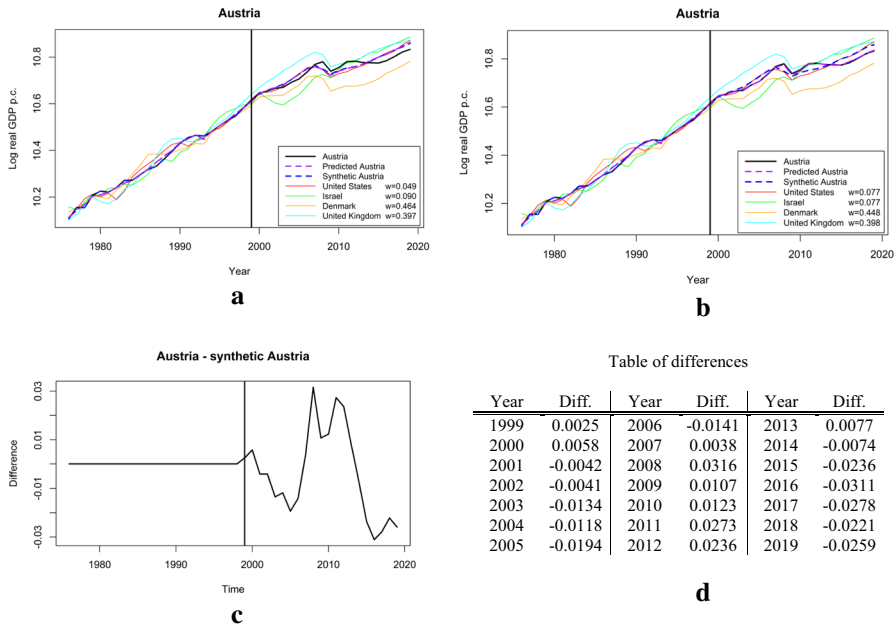


Fig. 3 Visual output of a BGMTV-model **a**, visual output of the full sample estimation **b**, difference of Austria and its BGMTV-model estimate **c** and the table of differences between Austria and its estimated BGMTV-model estimate **d**. Outliers and structural breaks are captured in the model

Table 4 Pulse and step dummies in the full model for Austria

Equation	Pulse dummy	Step dummy
Austria (pre 1998)	1990	1991–1992
Austria (post 1998)	2005,2006, 2007,2008, 2011,2012,2013,2014	1999–2000, 2001–2002,2003–2004, 2009–2010,2015–2019
USA	1982,1983	
Denmark	1985, 1986, 1987	2009–2019
Israel	1976,1989	2001–2007, 2002–2006
UK	1980, 1981, 1982, 1983, 1988, 1989, 1990	
Level	2008, 2009	

along with the model diagnostics for the models of the other CEA8. Detailed information on the estimation procedure as given in Tables 2 and 3 and Fig. 3 is provided in an Online Appendix for all the remaining CEA8 countries as well as for the Euro Area aggregate (EA19) and for a control model for Greece in which its date of joining EMU is set to 2001. Details on the estimated parameters and their significances are also provided in the Online Appendix.

Table 5 Diagnostics tests and estimated weights for the full models of the CEA8, the EA19 aggregate and the Greek (2001) model

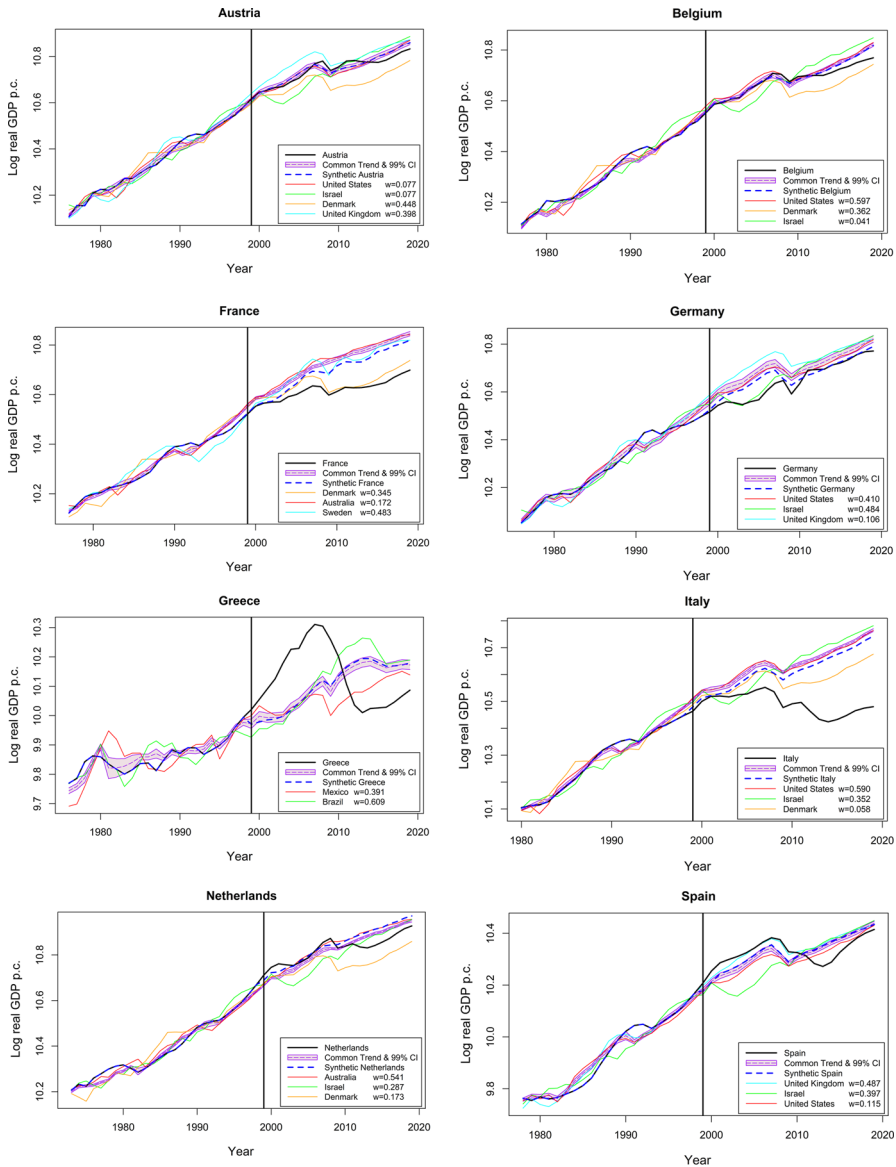
Model	H(h)	Q(5,4)	Q(10,9)	JB	Weights	
Austria (1976 – 2019)	H(14) 2.3170 (0.1278)	1.6299 (0.8034)	10.8072 (0.2892)	0.6736 (0.7140)	USA: Israel: Denmark: UK:	0.077 0.077 0.448 0.398
Belgium (1977 – 2019)	H(14) 1.1231 (0.8311)	0.6627 (0.9559)	7.503 (0.5849)	1.9594 (0.3754)	USA: Denmark: Israel:	0.597 0.362 0.041
France (1977 – 2019)	H(14) 1.8125 (0.2779)	11.3017 (0.0234)	15.6425 (0.0747)	0.0151 (0.9925)	Denmark: Australia: Sweden:	0.345 0.172 0.483
Germany (1976 – 2019)	H(14) 4.7372 (0.0063)	10.661 (0.0306)	18.4488 (0.0303)	4.164 (0.1247)	USA: Israel: UK:	0.392 0.474 0.134
Greece (1976 – 2019)	H(14) 2.6549 (0.0782)	15.8128 (0.0033)	16.9039 (0.0502)	0.9679 (0.6164)	Mexico: Brazil:	0.410 0.590
Italy (1980 – 2019)	H(13) 2.7857 (0.0759)	6.8798 (0.1424)	13.3641 (0.1468)	0.6396 (0.7263)	USA: Israel: Denmark:	0.574 0.358 0.067
Netherlands (1973 – 2019)	H(15) 1.8141 (0.2601)	11.3351 (0.0230)	15.1143 (0.0878)	0.7753 (0.6786)	Australia: Israel: Denmark:	0.532 0.276 0.192
Spain (1978 – 2019)	H(13) 1.9578 (0.2390)	7.1492 (0.1282)	10.8254 (0.2879)	1.4438 (0.4858)	UK: Israel: USA:	0.487 0.397 0.115
EA19 (1977 – 2019)	H(14) 2.6604 (0.0776)	5.4108 (0.2477)	16.582 (0.0557)	9.4272 (0.0090)	Denmark: USA: Israel:	0.294 0.446 0.260
Greece (2001) (1976 – 2019)	H(14) 2.0086 (0.2043)	16.1839 (0.0028)	18.5659 (0.0291)	1.447 (0.4851)	Mexico: Brazil:	0.360 0.640

P – values are given in parentheses. The estimation periods are given in parentheses in the first column.

4.2 Evaluating the euro net effect

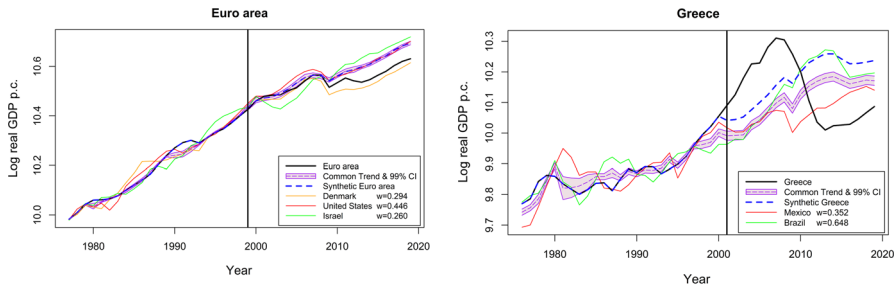
The procedure explained in the previous subsection is applied to each model for the CEA8 as well as for the EA19 aggregate and a control for Greece adopting the euro in 2001. The resulting visual output for each of the full models for CEA8 are shown in Fig. 4a. Figure 4b shows the visual results for the EA19 aggregate and Greece (2001).⁸ The outputs show the original target series, their synthetic control according to (10), the common trend with respect to the target series, i.e., $\hat{\mu}_t + \hat{\mu}_1$, and the mean-adjusted control series. Additionally, the estimated weights computed by (9) are given in the legend of each visual output. In each of the models, the common trend is modelled as

⁸ The EA19 aggregate is provided by the World Bank's world developments indicator database, too.



a Visual outputs of the full models for the CEA8 countries.

Fig. 4 a Visual outputs of the full models for the CEA8 countries. **b** Visual output of the full model for the EA19 aggregate and the Greek (2001) model.



b Visual output of the full model for the EA19 aggregate and the Greek (2001) model.

Fig. 4 continued

a random walk with drift.⁹ Before analysing these visual outputs, some notes on their generation and analysis are provided.

The first note concerns the choice of control group and negative weights in the estimation of full models. If the weight of a control country in a full model is negative, it is excluded from the control group. This applies to Denmark in the model for Germany and New Zealand in the model for Greece. In Sect. 4.3, robustness checks of the models are performed by estimating so-called "naïve" and "structural" models. In naïve models, only the euro adoption intervention is modelled, so that no structural breaks or outliers are captured in the model. In structural models, all common and country-specific structural breaks are added but no outliers in standardized irregulars. If the weight of a control country is negative in the full model, it is also dropped from the control group of the naïve and structural models, so that the robustness checks are performed with the same control group as in the full models. In some naïve and some structural models, the weights of the control countries are negative but become positive in the full model (compare Tables 5 and 7).

Second, country-specific structural breaks for Denmark and Israel are added to their observation equations in each model in which they are chosen as a control country. For Israel, two step variables are added: 2001–2007 and 2002–2007, and for Denmark one step variable is added: 2009–2019. This has the effect of pushing Israel to the common trend in 2001–2007 and Denmark in 2009–2019. Thus, Israel and Denmark serve only hardly as controls for these periods, but provide good control in the remaining periods. The addition of these level variables is closely linked to the assumption that the level changes in Israel and Denmark in the respective periods are the result of country-specific shocks, which should not affect the synthetic control estimate.

Third, country-specific structural breaks are added to the observation equations for France: 1996–2019, Germany: 1997–2019, Italy: 1996–2019, and the Netherlands: 1997–2019. These structural breaks are added because these countries leave the common growth path at these points in time. If these step variables were not added, the

⁹ The KPSS tests for level stationarity clearly show that the trend is stochastic for all models (see Online Appendix B). For some models, the KPSS tests for trend stationarity are less clear and suggest that the trend for some target and control series should also be modelled with a stochastic slope. However, visual inspection of the series shows no clear case for adding a stochastic slope to a trend, so all trends are modelled as random walks with drift.

euro effect would be extended to include the level change, which would lead to an overestimation or underestimation of the euro effect by the parameter value of the step variables. The effects of these country-specific structural breaks can be seen very clearly in the different levels of the common trend and the estimated synthetic controls after 1996/1997. The fact that these EMU members leave the common trend just before euro adoption may be consistent with the euro anticipation effects examined by some of the AGDH-SCM studies discussed in Sect. 2.2. The argument is that EMU countries (citizens, politicians, banks, etc.) might change their economic actions in line with their expectations of the euro effect even before the actual introduction of the euro in 1999. However, it is at least equally plausible that these structural changes can be explained by a lagged effect of the fall of the Soviet Union, especially for Germany, but also for France and Italy, where the structural breaks are negative. Since we do not know whether these structural breaks are already caused by the euro introduction, step variables are added to separate these effects from the euro adoption. The robustness model for Greece (2001) shows that Greece is leaving the common trend in 1998, so a step dummy is added to this model as well. Apart from this difference, its synthetic estimate is very close to that of the model in which Greece is joining in 1999. On the one hand, the fact that Greece leaves the common trend in 1998 speaks for its “real” entry in 1999, so that the dating of the intervention is justified by meeting of the criteria for joining EMU. On the other hand, we do not know whether this is caused by the euro adoption, as in the case of the other countries discussed above. In this context, Greek development after 1999/2001 was much more dynamic than that of the other CEA8 countries, and the only suitable control countries found are Mexico and Brazil, which are much less developed (Both have less than half of Greek GDP per capita). Thus, both models for Greece are subject to some uncertainty regarding a suitable control group and an individual Greek growth path.

Fourth, the common structural breaks modelled in (8) are worth mentioning in the different models. Dummy variables to account for the structural break caused by the financial crisis in 2008 and/or 2009¹⁰ are added to all models except the model for France. Moreover, there are some differences between the target and control series around 1990 in all models except the models for Greece and the Netherlands, and appropriate step and step dummies are added to the observation equation for the target to account for these differences. In the models of Belgium, France, Germany, Italy and Spain, outliers also indicate the addition of a common structural break in 1990, which is also captured with a pulse dummy in (8). Further structural breaks are added around 1980 to the models of Belgium, Germany, Greece, Netherlands and Spain. In general, most control and target series contain only a few outliers. Exceptions are the United Kingdom in the model for Austria, Sweden in the model for France, and Denmark in the model for the Netherlands, where more than five outliers are included in the corresponding observation equations. The interest in the parameter estimates is actually only secondary. The main focus is on visually examining the difference between EMU

¹⁰ In some models (Greece, Netherlands), adding a pulse in 2009 captures the impact of the crisis, eliminating outliers in the smoothed standardized state residuals, while in some models (Austria, Belgium, Germany, Italy, Spain), two pulses must be added in 2008 and 2009 to eliminate outliers and adequately account for the financial crisis.

countries and their estimated synthetic controls, which provide information on gains and losses resulting from the euro adoption.¹¹

The fifth and final note relates to model diagnostics. Table 5 provides information on the $Q(5,4)$, the $Q(10,9)$, the JB, the $H(h)$ statistics and their p-values for the target observation equation (since this is the series of interest), as well as the estimated weights of the control countries, which are also given in the legends of the various graphs in Fig. 4. In general, the model diagnostics are acceptable, but there is some evidence of autocorrelation in the models of France, Germany, Greece, and the Netherlands, as the $Q(5,4)$ test rejects the null value at the 5% significance level. However, the $Q(10,9)$ test accepts the null in the model of the Netherlands at the 10% level, and it accepts the null at the 5% level in the model of France and Greece. One could account for autocorrelation by modelling ε_t as a stationary multivariate ARMA-process, but then many more parameters must be estimated. Since the trend modelling splits non-stationary and stationary components, the effect of autocorrelation is likely to be small, and it is preferable to tolerate some degree of autocorrelation (HT 2021, p. 7). This is confirmed in the robustness section (see Sect. 4.3), where not accounting for outliers in many models leads to a significant increase in the $Q(p,l)$ statistic, while the estimate does not change much. All in all, the model diagnostics provide confidence in the estimates.

Keeping the five notes in mind, the analysis of the visual results in Fig. 4 shows two clear losers from the euro: France and Italy. While they are only marginally affected until 2001, they gradually lose out until 2019. Italy and France show a similar negative trend until 2010, but then Italy is hit much harder by the euro crisis and significantly increases the gap with its synthetic estimate, while France's gap increases only slightly. The synthetic estimates for Austria and Belgium are close to the original series. Due to the euro crisis, they lose slightly compared with their synthetic controls. The Netherlands records some small gains but is close to its synthetic estimate until 2009 but has been losing steadily since 2010. Germany loses until 2010, but is no longer affected thereafter. Spain and Greece gain until 2010 but lose afterwards. Spain finally returns to the common trend after the euro crisis. Interestingly, the visual output for the EA19 aggregate gives a similar picture as when considering all the individual results in their entirety.

The results of the visual analysis are quantified in Table 6, which provides information on the absolute and percentage differences between the original EMU series and their synthetic controls at the end of the pre-financial crisis period (2007), at the end of the financial crisis (2009), at the end of the euro crisis (2013), and at the end of the sample in 2019. Italy's remarkable loss in 2019 is \$9341, a difference of 26.23% from its synthetic control that would not have adopted the euro. France's shortfall is also remarkable, as its loss in 2019 is 12.07%, equivalent to about \$5350. Greece's large gains in 2007, as well as its losses in 2013, are also notable. For Austria, Belgium, and the Netherlands, neither the losses nor the gains are very large in any period.

¹¹ More detailed information (estimates, significance) on common structural breaks for all full models shown in Fig. 4 is included in Online Appendix D, along with all other pulse and step variables added to the model in the corresponding observational equations, and the smoothed estimates of the state variables $\hat{\mu}_t$, $\hat{\mu}_t$, $\hat{\beta}$ in $t = 2019$. Online Appendix D therefore also includes information on the actual outlier modelling in all models.

Table 6 Absolute and percentage gaps of CEAs and EA19 aggregate and the Greek (2001) model at the end of selected time periods

	2007		2009		2013		2019	
	%	abs	%	abs	%	abs	%	abs
Austria	0.36	172	1.15	531	0.79	380	- 2.62	- 1326
Belgium	1.00	448	0.72	312	- 1.36	- 601	- 4.83	- 2298
France	- 5.90	- 2455	- 8.71	- 3491	- 10.36	- 4275	- 12.07	- 5350
Germany	- 5.50	- 2290	- 3.70	- 1472	- 0.29	- 126	- 1.86	- 884
Greece	20.10	6040	15.87	4525	- 18.07	- 4021	- 9.15	- 2198
Italy	- 6.93	- 2650	- 10.29	- 3652	- 21.07	- 7158	- 26.23	- 9341
Netherlands	1.85	960	- 1.52	- 768	- 6.92	- 3501	- 4.51	- 2509
Spain	2.73	882	3.91	1196	- 7.98	- 2308	- 1.92	- 639
EA19	0.10	41	- 2.60	- 959	- 6.25	- 2351	- 6.25	- 2585
Greece (2001)	14.04	4219	9.88	2816	- 24.73	- 5503	- 14.89	- 3576

Overall, the development of the gaps between the original series and their synthetic estimates in the CEA8 is individual, but there are some structural movements in the response to the 2008–2009 financial crisis and to the 2010–2013 euro crisis. During the financial crisis, Austria, Belgium, and Spain benefit from the euro by losing less than their synthetic controls, and Germany benefits by catching up with its synthetic control. During the euro crisis, all CEA8 start to lose or continue to lose, leading to a negative gap (Austria, Belgium, Netherlands, Spain and Greece) or a widening of the negative gap (France, Italy). The only country that succeeds relative to its synthetic control during the euro crisis is Germany, which catches up and overtakes in 2011 and 2012, but then loses a little and approaches the level of its synthetic control by 2018. In the period after the euro crisis, the Netherlands, Spain and Greece start to catch up with their synthetic controls, while the size of the gap remains more or less unchanged for the other CEA8. In summary, with the exception of France and Italy, the CEA8 seem to perform slightly better during the financial crisis, while sliding into a structural crisis in the period thereafter.

Finally, it is worth noting that the synthetic controls obtained by applying HT-SCM are close to some of the AGDH-SCM estimates from the studies presented in Table 1. The synthetic control estimates for Austria, France, Germany, and the Netherlands are close to those of Puzzello & Gomis-Porqueras (2018); the estimates of Lin & Chen (2017) are broadly consistent with those for Austria, France, Germany, and Italy; and the results of Fernandez & Garcia Perea (2015) are close to those for Austria, Spain, and Greece.

4.3 Robustness section

In the following section, I will show that the estimated synthetic controls presented in Fig. 4 are robust to changes in the modelling approach. Therefore, the "full" models presented in Sect. 4.2 are compared to the models referred to below as "naïve" and "structural." A naïve model, as in step 2 of Sect. 4.1, refers to a model in which all temporary country-specific outliers as well as common and country-specific structural breaks are excluded from the model and only the euro adoption is modelled. A structural model refers to a model that includes all country-specific and common structural breaks as well as modelling the euro adoption, but excludes temporary country-specific effects.

Figure 5 shows the estimated differences between the target series and their synthetic controls in the full model, naïve model, and structural model together in one graph for each of the CEA8, the EA19 aggregate, and the Greek control model. Table 7 provides information on model diagnostics based on the JB statistics, the $Q(5,4)$ and $Q(10,9)$ statistics, and the $H(h)$ statistics, along with the estimated weights of the naïve and structural models.

Considering at first only Fig. 5, I find that the estimated gaps of the structural model are close to the estimated gaps of the full model. Exceptions are the estimates for Belgium and the EA19 aggregate, for which the structural model predicts higher losses. In many cases, the estimates of the naïve models differ somewhat more from the estimates of the full models. According to the naïve model, Austria, Belgium

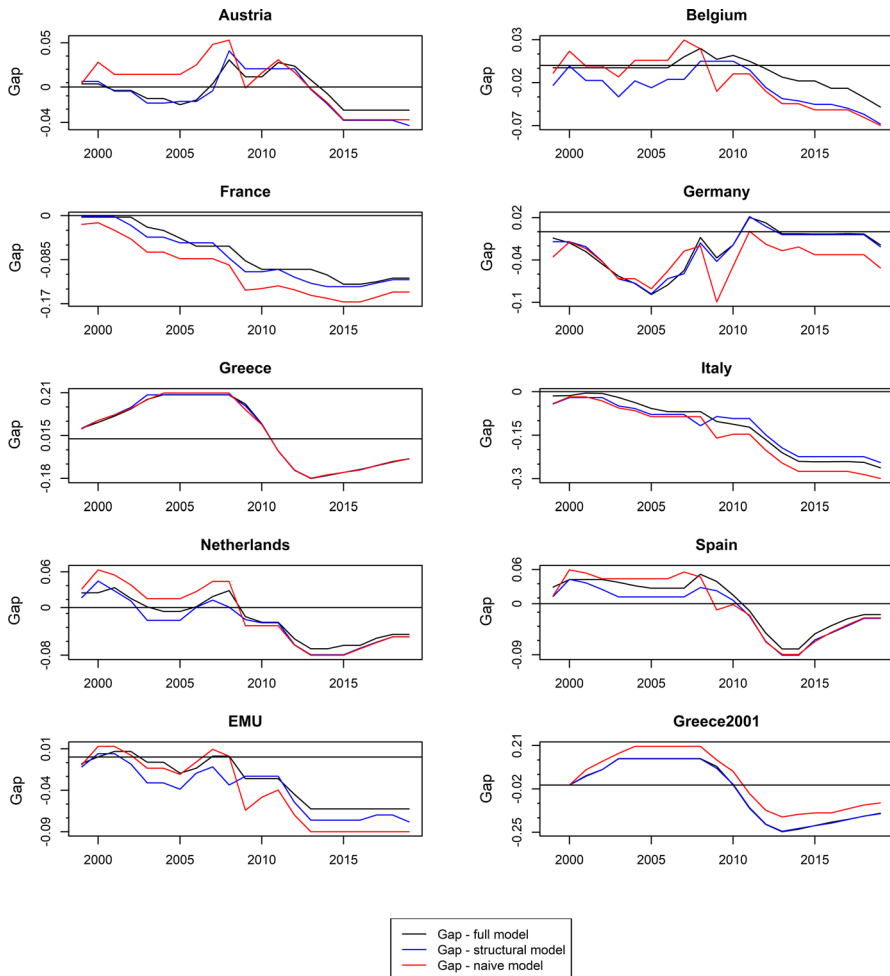


Fig. 5 Gaps of the several CEA8 countries, the EA19 aggregate and Greece (2001) estimated by the full model (black lines) the structural models (blue lines) and the naïve model (red lines). (Color figure online)

and the Netherlands would benefit by up to 3 to 5 percentage points more in the period before the financial crisis, while Germany and Belgium would lose up to 3 to 7 percentage points in the period during and after the financial crisis, compared with the gap estimated by the full model. For France and Italy, the gap estimated with the naïve model is larger by 2 to 5 percentage points for the entire period after the introduction of the euro. Despite these discrepancies, the naïve, structural and full models create synthetic controls that move in parallel, so they can be said to follow the same growth path. Note that the difference between the naïve and structural/complete models for France reflects the step variable that is added because France leaves the common trend even before the introduction of the euro. If this variable were not added, the euro net effect would be overestimated by the corresponding parameter, which is -0.0268 and

highly significant. The same applies to the model for Italy (where the parameter is -0.0236 and highly significant), to the model for the Netherlands in the period before the financial crisis (where the parameter is 0.0202 and highly significant) and to the model for Germany (where the parameter is -0.0331 and highly significant) in the period after 2008.

Table 7 shows that several of the estimated naïve and structural models fail some or all of the model diagnostics at the 5% significance level and that the estimated weights are negative in a number of models (especially for Denmark). Moreover, the estimated weights differ substantially between structural, naïve, and full models (compare the weights in Table 5 with those in Table 7). However, despite the failed diagnostic tests and the different estimated weights, which may even be negative, the three models yield similar synthetic controls. This leads to the conclusion that including outliers and structural breaks in the full model primarily eliminates misspecification, but does not have as large an impact on the results. In addition, adding country-specific outliers and structural breaks eliminates extrapolation in the estimation of weights to some extent, which is reasonable because influential observations bias the variance and the estimation of weights depends on Σ_ε .

From Fig. 5 and Table 7, I conclude that the full model estimates are mostly robust to (a) modelling outliers and thus the number of regression parameters included in the model, (b) changes in estimated weights, and (c) some degree of misspecification indicated by the diagnostic tests.

4.4 Explaining the gap

The estimation of synthetic controls for the CEA8 opens the door for further research on the economic impact of a euro adoption. We are now able to link the theoretical and empirical state of the literature, briefly described in Sect. 2.1, to the gaps between the CEA8 and their synthetic controls. Moreover, we are able to examine the channels of the financial and euro crises in more detail. In this subsection, I derive hypotheses about the effects of the main drivers of gains and losses due to EMU membership and test them using fixed-effects regressions. The aim of this section is to relate the estimated synthetic controls to some of the main drivers of gains and losses described in the literature. Since much of the literature on common monetary unions is concerned with the loss of individual monetary policy in response to (asymmetric) shocks, the crisis channel is of particular interest. Therefore, the regression analysis is divided into two parts. First, benchmark results are constructed for the entire sample. Second, the sample is split into four periods: the pre-crisis period, the financial crisis period, the euro crisis period, and the post-crisis period. While this drastically limits the degrees of freedom and may not be appropriate for drawing conclusions from these results, it allows for a more detailed examination of the crisis channel. In addition to testing the hypotheses stated in the OCA/EMU literature, the regression results also provide validation of the accuracy of the estimated synthetic controls.

Table 7 Diagnostics tests and estimated weights of the controls of the naïve and structural models of the CEA8, the EA19 aggregate and the Greek (2001) model

Model	H(h)	Q(5,4)	Q(10,9)	JB	Weights	
Austria naïve	H(14)	2.2303	7.3009	5.127	USA:	0.449
	3.5644	(0.6935)	(0.6058)	(0.0770)	Israel:	0.424
	(0.0235)				Denmark:	0.047
					UK:	0.080
Austria structural	H(14)	6.7179	9.6294	6.6641	USA:	0.358
	6.8079	(0.1516)	(0.3813)	(0.0357)	Israel:	0.403
	(0.001)				Denmark:	0.135
					UK:	0.105
Belgium naïve	H(14)	4.4201	15.5002	1.0022	USA:	0.668
	1.8897	(0.3521)	(0.0781)	(0.6059)	Denmark:	− 0.051
	(0.246)				Israel:	0.383
Belgium structural	H(14)	11.1734	22.8895	0.2977	USA:	0.635
	5.6135	(0.0247)	(0.0064)	(0.8617)	Denmark:	0.013
	(0.0027)				Israel:	0.352
France naïve	H(14)	62.4185	70.76	2.2363	Denmark:	− 0.020
	2.7736	(0.0000)	(0.0000)	(0.3269)	Australia:	0.821
	(0.0662)				Sweden:	0.199
France structural	H(14)	8.7995	9.0997	0.8896	Denmark:	0.122
	3.654	(0.0663)	(0.4281)	(0.6410)	Australia:	0.844
	(0.0211)				Sweden:	0.034
Germany naïve	H(14)	17.1052	19.3065	1.346	USA:	0.723
	1.3626	(0.0018)	(0.0227)	(0.5102)	Israel:	0.312
	(0.5704)				UK:	− 0.035
Germany structural	H(14)	14.8913	21.2613	35.5771	USA:	0.345
	6.2801	(0.0049)	(0.0115)	(0.0000)	Israel:	0.413
	(0.0015)				UK:	0.242
Greece naïve	H(14)	17.216	19.5629	16.3329	Mexico:	0.373
	6.4025	(0.0018)	(0.0208)	(0.0003)	Brazil:	0.627
Greece structural	H(14)	16.9108	19.6148	7.5326	Mexico:	0.375
	5.0564	(0.002)	(0.0204)	(0.0231)	Brazil:	0.625
Italy naïve	H(13)	14.8561	46.886	0.9814	USA:	0.969
	2.5771	(0.0050)	(0.0000)	(0.6122)	Israel:	0.191
Italy structural	(0.1000)				Denmark:	− 0.160
	H(13)	13.0855	31.1499	1.7619	USA:	0.903
	1.5132	(0.0109)	(0.0003)	(0.4144)	Israel:	0.188
Netherlands naïve	(0.4654)				Denmark:	− 0.090
	H(15)	19.7216	27.3074	13.4003	Australia:	0.768
	2.2301	(0.0006)	(0.0012)	(0.0012)	Israel:	0.388
Netherlands structural	(0.1315)				Denmark:	− 0.156
	H(15)	14.9231	19.1357	1.3999	Australia:	0.777
	10.6928	(0.0049)	(0.0241)	(0.4966)	Israel:	0.360
	(0.0000)				Denmark:	− 0.138

Table 7 (continued)

Model	H(h)	Q(5,4)	Q(10,9)	JB	Weights	
Spain naïve	H(13)	37.0932	58.4232	4.0587	UK:	0.744
	16.9553	(0.0000)	(0.0000)	(0.1314)	Israel:	0.329
	(0.0000)				USA:	− 0.074
Spain structural	H(13)	37.3795	47.5358	7.5588	UK:	0.757
	10.4885	(0.0000)	(0.0000)	(0.0228)	Israel:	0.318
	(0.0001)				USA:	− 0.075
EA19 naïve	1/H(14):	16.8520	36.4110	0.1168	Denmark:	− 0.190
	5.7956	(0.0021)	(0.0003)	(0.9433)	USA:	0.799
	(0.0022)				Israel:	0.392
EA19 structural	3.8420	7.6138	11.7700	0.4073	Denmark:	− 0.125
	(0.0168)	(0.1068)	(0.2266)	(0.8157)	USA:	0.759
					Israel:	0.366
Greece (2001) naïve	H(14)	15.9299	18.7041	10.0225	Mexico:	0.457
	8.9689	(0.0031)	(0.0278)	(0.0067)	Brazil:	0.543
	(0.0001)					
Greece (2001) structural	H(14)	14.0349	17.1936	6.4958	Mexico:	0.357
	4.6133	(0.0072)	(0.0458)	(0.0389)	Brazil:	0.643
	(0.0071)					

The estimation periods are identical to those given in Table 5. P – values are given in parentheses

4.4.1 Deriving hypotheses from the literature on common currency areas

The loss of an individual monetary policy and, in particular, the inability to respond to shocks with currency devaluations or revaluations has been seen as major costs of joining a common currency since the seminal work of Mundell (1961), McKinnon (1963), and Kenen (1969). As a result, much of the research on joining currency unions is concerned with finding alternative regulatory mechanisms. In this context, Mundell (1961) focuses on labour mobility and price and wage flexibility. In short, the rationale behind these criteria is as follows: An asymmetric shock can lead to demand and supply shocks, so that some countries face excess demand and others face excess supply. If workers who become unemployed in one country now simply move across the border where their labour is demanded, this will effectively dampen the shock. Moreover, currency devaluations and revaluations are indirect changes in the domestic price level. If wages and prices could be adjusted directly, for example, by structurally hiring or firing workers or by imposing new wages or prices on products, currency devaluations and revaluations become irrelevant. Unfortunately, the actual degree of labour mobility is difficult to observe, but it can be approximated by net migration. The more open an economy is to immigration, the higher the degree of labour mobility will be and the lower the cost of joining EMU. The data on migration come from Eurostat. Net migration is defined as the sum of immigrants minus emigrants as a percentage of the population. One indicator of wage flexibility is the employment protection index provided by the OECD database. The higher an index value, the lower the wage flexibility, so the expected sign for the employment protection index is negative.

McKinnon (1963) emphasizes that a higher degree of openness lowers the cost of joining a common monetary union because the domestic price level depends to a greater extent on the international prices of tradable goods. This makes currency devaluations and revaluations less effective and the loss of the ability to change the value of the currency less costly. In addition, trade with currency partners becomes cheaper because many trade costs are eliminated, so more open economies benefit more from joining a currency union if they trade more with currency partners. Consequently, small and open economies such as Austria, Belgium, and the Netherlands should benefit more from joining EMU, but openness should be beneficial for all EMU members. Openness is measured via trade openness as the sum of imports and exports as a percentage of GDP. Data on imports and exports stem from the World Bank's World Development Indicators database.

In the newer stream of the literature on EMU, Frankel & Rose (1997, 1998) argue for the endogeneity of a monetary union and conclude that the synchronization of business cycles is an important criterion for the success of a monetary union. Convergence effects are closely related to the endogeneity of a monetary union. According to Frankel & Rose (ibid.), convergence of business cycles means that they become more synchronous over time. Dellas & Tavlas (2009, p. 1128) elaborate that real convergence processes are also likely for economies joining a monetary union whose members are more economically developed. According to the authors, the relatively underdeveloped economy would benefit by adopting relatively high expected returns on investment and low real interest rates. These, in turn, lead to more optimistic income expectations and excess domestic demand. To test the real convergence hypothesis, 1998 GDP per capita is included in the regression. The convergence hypothesis states that the lower GDP per capita was before euro adoption, the higher the gains from joining EMU. Therefore, the expected sign of the parameter for GDP per capita in 1998 is negative.

The synchronization of business cycles is often measured in the recent empirical literature using the approach of Cerqueira & Martins (2009) and Cerqueira (2013). Their approach allows for the computation of the correlation of cycles at a given point in time rather than only in a given time range. Thus, it is possible to include temporal variability in the measure of co-movement, which is limited by the conventional correlation measure. The calculation for the degree of co-movement is as follows

$$\rho_{ijt} = 1 - \frac{1}{2} \left(\frac{(d_{jt} - \bar{d}_j)}{\sqrt{\frac{1}{T} \sum_{t=1}^T (d_{jt} - \bar{d}_j)^2}} - \frac{(d_{it} - \bar{d}_i)}{\sqrt{\frac{1}{T} \sum_{t=1}^T (d_{it} - \bar{d}_i)^2}} \right)^2, \quad t = 1, \dots, T \quad (14)$$

$$\rho_{ijt}^{nb} = \frac{1}{2} * \ln \left(\frac{1 + \frac{\rho_{ijt}}{2T-3}}{1 - \rho_{ijt}} \right) \quad (15)$$

$$\rho_{ijt}^{cor} = \tanh(\rho_{ijt}^{nb}) \quad (16)$$

Cerqueira (2013) shows that ρ_{ijt} is bounded to $[3-2T, 1]$, while ρ_{ijt}^{nb} is (un)bounded to $(-\infty, \infty)$ and so ρ_{ijt}^{cor} is bounded to $[-1, 1]$. While all of these measures are indicators of the degree of co-movement of two series, only ρ_{ijt}^{cor} allows the usual interpretation of correlation, and so ρ_{ijt}^{cor} is chosen to measure business cycle correlation.¹² To calculate the business cycle, the HP filter is applied to the GDP per capita series (transformed into logarithms, sample range from 1970 to 2019) with $\lambda = 100$, as suggested for annual data. The synchronization of the business cycles is evaluated as the mean of the correlations between the CEA8 and the 12 early adopters (EA12), so that the synchronization of the business cycle for CEA8 i at time t is measured as follows

$$\bar{\rho}_{it}^{cor} = \frac{1}{M-1} \sum_{j=1}^N \rho_{ijt}^{cor}, i \neq j, i = 1, \dots, N, j = 1, \dots, M, t = 1999, \dots, 2019 \quad (17)$$

where $N = 8$ and $M = 12$. If the hypothesis of Frankel & Rose (1997, 1998) holds, the sign of $\bar{\rho}_{it}^{cor}$ should be positive.

Finally, two control variables – public debt and competitiveness – are added to the regressions. Public debt is added to control for the effect of fiscal stance and to observe its effect during the financial crisis and, in particular, the euro crisis, for which public debt is often considered the main driver. In general, a country that does not share its currency with other countries can increase its public debt to dampen the effects of a shock. The country can then create inflation, i.e., print money to repay its debt. This is not possible if a country is a member of a monetary union with a single central bank. Then rising public debt can increase investor distrust in the country in question, and interest rates on government debt rise. Instead of cushioning the shock, this can make it worse (De Grauwe 2020, pp. 10ff.). This consideration is the basis for Kenen's (1969) criterion of fiscal integration. Asymmetric shocks could be cushioned by fiscal transfers rather than by an increase in public debt. However, since EMU is not a fiscal union, fiscal stance becomes important to avoid sliding into a sovereign debt crisis. Public debt data are provided by Eurostat and measured as a percentage of GDP. The expected sign of government debt is negative, and the impact of government debt should be greater in the period of the euro crisis.

Currency revaluations and devaluations affect the competitiveness of an economy because they change the domestic price level relative to the foreign price level. In a monetary union, the domestic price level is determined by the aggregate price level of the monetary union. An economy joining a monetary union must compete in international markets with the price level set by the monetary union. Once an economy can no longer compete with this price level, for example because it is hit by a shock and unemployment rises, it has no way to change its competitiveness through currency devaluation, and sharing a currency becomes costly. This point is made by Sinn (2014). As a result, a loss of competitiveness is expected to increase losses from EMU membership. To estimate the impact of competitiveness, real unit labour costs from

¹² Note that $mean(\rho_{ijt}) = cor(d_i, d_j)$, but $mean(\tanh(\rho_{ijt}^{nb})) \approx cor(d_i, d_j)$ because $\tanh$ is not linear.

the AMECO database are included in the regression. The base year for real unit labour costs is 1998, so real unit labour costs in 1998 are 100. The higher the real unit labour cost, the lower the competitiveness, so the expected sign of the real unit labour cost is negative.

4.4.2 Regression analysis

To test the derived hypotheses, several panel regressions are performed. The analysis mainly relates to two benchmark models. The first benchmark model covers the entire period after the introduction of the euro. In the second benchmark, the effect of several variables is split using cross terms with step dummies to extract the effect of the variables in different time periods. These dummies are defined to divide the period into a pre-crises period, a post-financial crisis period, a post-euro crisis period, and a post-crises period:

$$\begin{aligned} precrises_t &= \begin{cases} 1, & \text{if } 1999 \leq t \leq 2007 \\ 0, & \text{else} \end{cases} \\ fcrisis_t &= \begin{cases} 1, & \text{if } 2008 \leq t \leq 2009 \\ 0, & \text{else} \end{cases} \\ ecrisis_t &= \begin{cases} 1, & \text{if } 2010 \leq t \leq 2013 \\ 0, & \text{else} \end{cases} \\ postcrises_t &= \begin{cases} 1, & \text{if } 2014 \leq t \leq 2019 \\ 0, & \text{else} \end{cases} \end{aligned} \quad (18)$$

Table 8 provides information on the different variables used in the regression models. Column 1 informs about the variables in the regression analysis with their abbreviations in parentheses, column 2–8 inform about their scaling, some descriptive statistics, the number of NAs and the expected sign. As can be seen, net migration is the only variable with missing values. These missing values occur for France in 1999–2005, Belgium in 2008–2009, and Spain in 1999–2001. In addition to the variables listed in Table 8, a cross term of $GDP_{pc98} \times RULC_{98}$ is added to the regressions to control for the association that poorer economies are generally less competitive. Adding this cross term separates the effects of convergence (poorer economies gain more) and competitiveness (less competitive economies gain less). The explanatory variables are mean-centred to ensure that the estimated time fixed effects can be interpreted meaningfully. Mean-centring allows their parameters to be interpreted as the euro effect for an average EMU member at time t . Comparing the parameters of the time fixed effects with the estimated gap for the EA19 aggregate provides some cross-validation of the regression results and the estimate of the synthetic STM controls.

The results of the regressions are reported in Table 9. The first two regressions are estimated for the entire sample after the introduction of the euro. Model (1) is the benchmark with all variables. Model (2) excludes net migration to control for the effect of excluding 12 observations. For model (3), the effects of business cycle correlation, competitiveness, trade openness, public debt, net migration and employment

Table 8 Overview of the variables used for the regression analysis

Variable	Transformation	Min	Max	Mean	Var	NA's	Exp. sign
Estimated Gap (Gap)	—	− 0.2623	0.2010	− 0.0273	0.0070	0	
Business cycle correlation (BCS)	—	− 0.5137	0.9128	0.6075	0.0643	0	+
Real unit labour costs (RULC98)	Base year: 1998 = 1	0.8983	1.0940	0.9916	0.0014	0	—
Trade Openness(TO)	$\frac{TO}{GDP}$	0.4462	1.6542	0.8560	0.1352	0	+
Public Debt (Debt)	$\frac{Debt}{GDP}$	0.3580	1.8120	0.8783	0.1042	0	—
Net migration (Net — migr)	$\frac{net_migr}{population} * 100$	− 0.6020	1.6167	0.3277	0.1272	12	+
Employment Protection (Emp — prot)	—	1.6380	3.6110	2.5748	0.2275	0	—
GDP per capita in 1998 (GDPpc98)	per 1000 USD	21.9027	43.0192	34.2762	44.2732	0	—

Abbreviations of the variables are given in parentheses in column one. Variable *Gap* is the gap estimated by the full models discussed in Sect. 4.2

protection are partitioned into the pre-crisis, financial crisis, euro crisis and post-crises periods by multiplying them by the mentioned step dummies in (18). All three models include time fixed effects and the capturing of outliers ($|std. residuals| > 2.75$). These occur only in the residuals of Greece in 2004, 2006, 2007, and 2012 in models (1) and (2) and in 1999 in model (3). Moreover, the variance of the Greek residuals is also visibly higher. Given this, the models exhibit some difficulty in predicting Greece, suggesting that Greece's post-1999 performance is influenced, at least to some extent, not only by the euro adoption but also by an individual growth path.

A first analysis of models (1) and (2) shows that the main drivers of gains and losses are real convergence processes, trade openness, public debt, competitiveness and net migration. All of these factors are highly significant and have the correct sign. Business cycle synchronization and employment protection are not significant. The results of model (2) show that the exclusion of net migration does not change the sign, magnitude, and significance of the other variables, leading to the conclusion that the addition of net migration and the accompanying loss of 12 observations is not a trade-off. Moreover, $\bar{R}^2$ indicates that model (1) is slightly better and therefore the inclusion of net migration may be the better choice to explain the gains and losses.

When analysing model (3), it is important to keep in mind that splitting the effects of the variables across the pre-crisis, financial crisis, euro crisis and post-crises periods drastically reduces the degrees of freedom for estimating the respective parameters.

Table 9 Regression results

Variable	(1)	(2)	(3)			
	Full sample			Precrisis	Fcrisis	Ecrisis
BCS	-0.0099 (-0.4890)	-0.0196 (-0.8561)	0.0665*** (2.9569)		-0.0294 (-1.0812)	-0.1065*** (-3.3981)
GDPpc98	-0.0089*** (-9.0246)	-0.0086*** (-8.9405)	-0.0123*** (-15.4076)			
RULC98	-0.3027** (-2.6038)	-0.2846** (-2.3901)	-0.0319 (-0.1960)		0.3289 (0.7605)	-0.1805 (-0.7708)
TO	0.0967*** (11.7004)	0.0985*** (12.9523)	0.1505*** (13.5976)		0.1127*** (3.5570)	-0.0301** (-2.0008)
Debt	-0.1538*** (-7.7166)	-0.1561*** (-7.5490)	-0.0550*** (-4.4673)		0.0152 (0.5660)	-0.1942*** (-9.2763)
Net-migr	0.0386*** (4.8846)		0.0152 (1.3314)		-0.0267 (-1.6120)	0.0245 (1.4281)
Emp-prot	-0.0023 (-0.3362)	-0.0020 (-0.2676)	0.0562*** (7.0280)		-0.0940*** (-3.8873)	-0.1020*** (-5.1634)
GDPpc98 × RULC98	-0.1280*** (-8.1787)	-0.1144*** (-6.6409)	-0.1135*** (-10.9114)			
Time FE	Yes	Yes	Yes			
Country FE	No	No	No			
R ²	0.7135	0.6675	0.8936			
NT	156	168	156			
DF	123	136	108			

Abbreviations of the variables are as given in column 1 in Table 8. Dependent variable in all models is *Gdp*. In model (3) *GDPpc98* and *GDPpc98* × *RULC98* are the only variables that are not multiplied by any of the step variables given in (18), so their parameters are estimated with the full sample. t-statistics based on Newey–West errors with lag = 2 are reported in parentheses. ***, **, * denote significance at 1, 5, and 10%, respectively.

Since there are 123 degrees of freedom to estimate 8 variables in model (1), extracting the effects for different periods reduces the degrees of freedom to 108 to estimate 26 variables in model (3). Moreover, the observations for estimating the parameters in different periods reduce to 8×9 for the pre-crises period, 8×2 for the financial crisis period, 8×4 for the euro crisis period, and 8×6 for the post-crises period. Thus, the analysis of model (3) is not so much about significance, but about the signs of the parameters and the comparison of their magnitudes with model (1). Thus, a comparison of the results of model (3) with those of model (1), focusing mainly on the sign and magnitude of the parameters, provides some interesting information. The first interesting observation is that $\bar{R}^2$ is much higher than in model (3), even though the same variables are used. This suggests that considering different slopes in different sample periods not only provides interesting information about the variables, but also makes sense from an econometric point of view.

The other interesting information concerns the variables. Instead of being insignificant and close to zero in model (1) and (2), the business cycle correlation actually affects profits and losses, given the different slopes in the different periods. While its impact is positive in the pre-crises period, its slope becomes negative thereafter. A simple interpretation of this change in slope is that ECB policy was optimal in the pre-crises period but detrimental thereafter. The difficult task is to explain why the effect of ECB policies changes. As a possible explanation, I examine in Fig. 6 the cycle movements of the various EA12 countries and $\bar{\rho}_{it}^{cor}$ of the CEA8 from 1990 to 2019. In general, the cycles of the EA12 countries coincide very well until 2007. Some exceptions to the general co-movement are Germany in the early 1990s and around 2005, the Netherlands during the turn of the century, and Greece, whose cycle starts to diverge around 1998. The only negative correlation until 2007 is for Greece in 1999. Looking now at the period after 2007, the picture is different. In particular, the

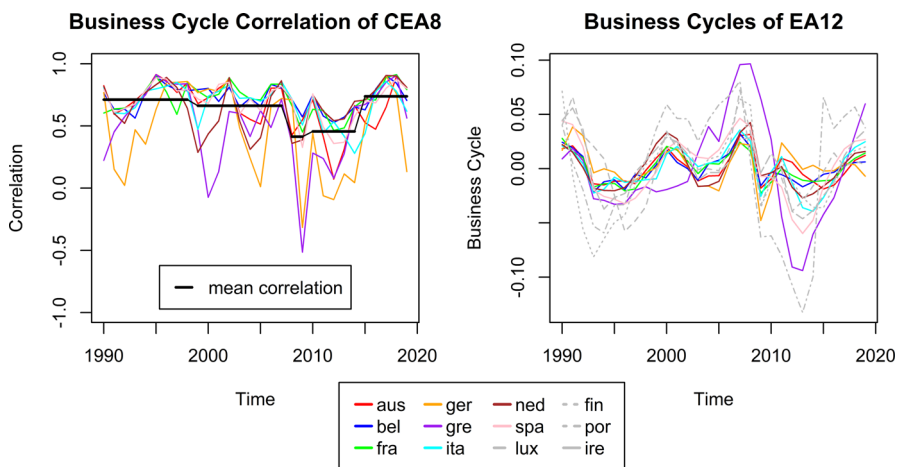


Fig. 6 Business cycle correlation and mean correlation calculated for 1990–1998, 1999–2007, 2008–2009, 2010–2014 and 2015–2019 for the various CEA8 with the EA12 (left) and extracted business cycles of the EA12 (right)

crisis causes the cycles to diverge and the correlation to become lower, which changes again after the euro crisis around 2016/2017. Therefore, in the pre-crisis period, it was more or less easy for the ECB to pursue an optimal business cycle policy for all EMU members, as the business cycles were synchronous. However, during and after the crisis, the cycles were no longer so synchronous, and the ECB may have chosen its policy to dampen the shock for the heaviest "losers" of the crisis, namely Greece, Italy and Spain (and indeed quantitative easing, debt relief for Greece and low interest rates suggest that the ECB did so). In turn, ECB policy was more optimal for the less correlated EMU members, leaving out the more correlated countries. However, this effect needs further research to fully explain the negative sign since the financial crisis.

While the real unit labour cost parameter is negative and significant in models (1) and (2), the impact of competitiveness seems to be a driving force especially during the euro crisis and in the post-crisis period. As expected, its sign is negative at this time. During the financial crisis, on the other hand, lower competitiveness appears to be beneficial. However, the t-statistics of real unit labour costs are very small in each period. Net migration is an important factor in dampening the shock of the euro crisis, while the fiscal stance is important in the pre-crisis and post-crisis periods. As expected, the negative impact of fiscal instability is much larger during and after the euro crisis than in the pre-crisis period. Trade openness is beneficial until the financial crisis, but it does not help dampen the shock of the euro crisis. On the contrary, a more open economy further increases the losses from EMU membership during and after the euro crisis. Employment protection has a positive effect in the pre-crisis period, but as expected, a rigid labour market increases losses during and after the crisis.

Overall, the empirical results of the three regressions paint a picture that is largely consistent with the hypotheses, and examination of the crisis channel shows that some of the main drivers do indeed exert their effects during a crisis. Lastly, Fig. 7 shows the plot of the estimated gap for the EA19 aggregate together with the time fixed effects of the three regression models. The estimates for an average EMU member (the EA19 aggregate is a weighted combination of the 19 EMU members and can therefore be considered an EMU average) are very similar despite different approaches, further cross-validating the estimation results of the HT-SCM and the regression. Only in the 2004–2007 period are the time fixed effects of models (1) and (2) significantly lower than those of model (3) and those of the estimated EA19 gap of the full model, while the time effects of model (3) are significantly higher in the 2008–2010 period. When capturing outliers in models (1)–(3), the time fixed effects are even closer to the gap, which is due to the fact that capturing outliers isolates Greece's observations from the regression at the corresponding points in time. Therefore, when comparing different models, it may be more appropriate not to include outliers in the regressions.

5 Conclusion

This study assesses the impact of euro adoption on GDP per capita for eight early adopters and the EA19 aggregate by applying a new SCM based on STM and cointegration. The results show that France and Italy are clear losers from the euro. Italy in particular suffers drastic losses. While Greece gains a lot until the financial crisis,

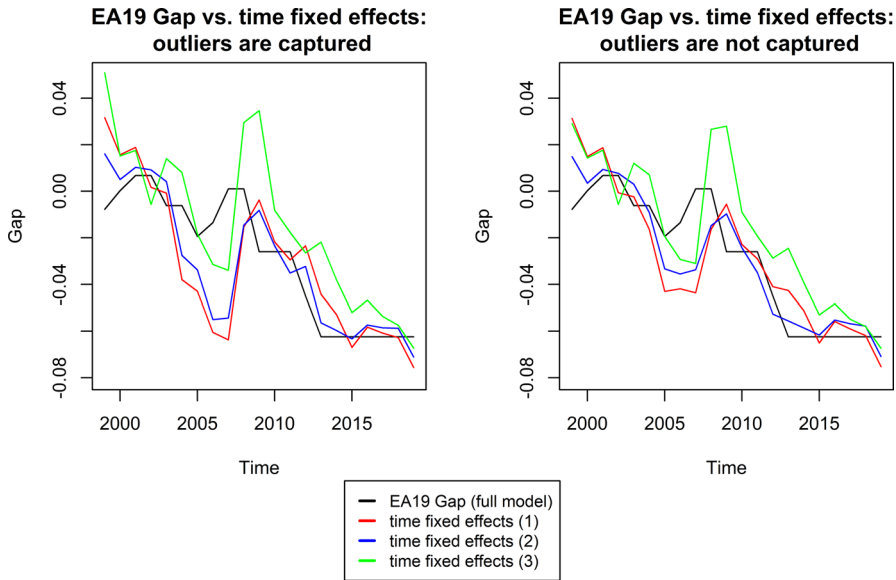


Fig. 7 Estimated EA19 gap of the full HT-SCM model and time fixed effects of the regression models when regression outliers are captured (left) and when regression outliers are not captured (right). The red, blue and green line visualize the time fixed effects of regression model (1), (2), and (3), respectively

its losses thereafter are also remarkable. Spain also gains until 2009, but loses thereafter. After the euro crisis, Spain increasingly catches up with its synthetic estimate. Germany loses initially, but is able to catch up during the financial and euro crises. The Netherlands, Austria and Belgium are hardly affected with slight gains or losses until 2009, but then lose due to the euro crisis. The results clearly show that, with the exception of Germany, all CEA8 slip into the crisis in 2010.

Identifying an ex-post benchmark for the net euro effect opens the possibility to test hypotheses from the OCA literature. Running panel regressions shows that real convergence, trade openness, and net migration as a proxy for labour mobility are the main drivers of gains from EMU membership, while fiscal instability and low levels of competitiveness are the main drivers of losses. Thus, the regression confirms the thoughts and theories of the OCA literature.

Both the OCA literature and the estimated synthetic controls suggest that the crisis channel plays an important role. While the OCA literature concludes that the main costs stem from the inability to respond to asymmetric shocks with independent central bank policy, the estimated gaps between the CEA8 and their synthetic controls show that the euro crisis in particular leads to a larger negative gap. Examining the effects of the main drivers in the pre-crisis, financial crisis, euro crisis and post-crises periods, I find that the effects of the main drivers unfold during the crises. Taking into account different slopes for these periods, I find that business cycle synchronization has a negative impact during and after the crises, while it is beneficial in the pre-crises period. In contrast, its effect turns out to be insignificant when different slopes are not taken into account. This suggests that ECB policy was not optimal for the more correlated EMU

members and focused mainly on the less correlated and more suffering economies of Greece, Spain and Italy. Employment protection, as an indicator of wage and labour market rigidity, is negative during and after the crises, thus acting mainly through the crisis channel. Net migration has its positive effect during the euro crisis. Openness is beneficial in the pre-crisis period, but its effect is negative thereafter. Finally, fiscal instability is found to be the main driver of losses during and after the euro crisis, confirming that the euro crisis is a sovereign debt crisis. Lower competitiveness may also have a negative effect in the period after the financial crisis.

In summary, the results of this study draw a rather pessimistic picture of the economic impact of a euro adoption, and both the STM synthetic controls and the time fixed effects of the regression models suggest that an average EMU economy has to expect losses from an adoption. However, both measures also suggest that the euro crisis was responsible for a large share of these losses. Thus, understanding the crisis channel is important for future research and for policymakers. The finding that fiscal instability is a major problem of EMU opens the debate on whether EMU members should integrate further into a fiscal union and become more and more the "United States of Europe" or whether countries should be allowed to leave EMU and return stronger after regaining their competitiveness and fiscal stance. Aside from the question of how to adjust the EMU system in detail, the results of this study show that it must be adjusted to be successful in the future.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s00181-022-02257-x>.

Funding Open Access funding enabled and organized by Projekt DEAL.

Declarations

Conflict of interest There is no conflict of interest to declare for this study.

Code availability The empirical analysis of this study is performed in R (version 4.0.4). All codes used in the analysis are available upon request.

Availability of data and material All data and materials used for this study are publicly available. All sources to procure the data are named in the manuscript. To facilitate replication, data are also available upon request.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Abadie A, Gardeazabal J (2003) The economic costs of conflict: a case study of the basque country. *Am Econ Rev* 93(1):112–132. <https://doi.org/10.1257/000282803321455188>
- Abadie A, Diamond A, Hainmueller J (2010) Synthetic control methods for comparative case studies: estimating the effect of california's tobacco control program. *J Am Stat Assoc* 105(490):493–505. <https://doi.org/10.1198/jasa.2009.ap08746>
- Abadie A, Diamond A, Hainmueller J (2015) Comparative politics and the synthetic control method. *Am J Polit Sci* 59(2):495–510. <https://doi.org/10.1111/ajps.12116>
- Alesina A, Barro RJ, Tenreiro S (2003) Optimal currency areas. *NBER Macroeconomics Annual* 17:301–356. <https://doi.org/10.1086/ma.17.3585292>
- Baldwin R, Wyplosz C (2019) *The Economics of European Integration*, 6th edn. McGrawHill, London
- Bayoumi T, Eichengreen B (1993) Shocking aspects of monetary unification. In: Giavazzi F, Torres F (eds) *Adjustment and growth the european monetary union*. McGrawHill, London, pp 193–235
- Bayoumi T, Eichengreen B (1997) Ever closer to heaven? an optimum-currency-area index for european countries. *Eur Econ Rev* 41(3–5):761–770. [https://doi.org/10.1016/s0014-2921\(97\)00035-4](https://doi.org/10.1016/s0014-2921(97)00035-4)
- Bekaert G, Harvey CR, Lundblad CT, Siegel S (2017) *Economic and financial integration Europe*. ifo DICE Report. 15(1):36–42
- Broz T (2005) The theory of optimum currency areas: a literature review. *Privredna Kretanja I Ekonomska Politika* 104(2005):53–78
- Campos NF, Fidrmuc J, Korhonen I (2017) Business cycle synchronisation in a currency union: taking stock of the evidence. *Bank of finland research discussion* pp. 28
- Campos NF, Macchiarelli C (2016) Core and periphery the european monetary union: bayoumi & eichengreen 25 years later. *Econ Lett* 147:127–130. <https://doi.org/10.1016/j.econlet.2016.07.040>
- Carvalho VM, Harvey AC (2005) Convergence in the trends and cycles of euro-zone income. *J Appl Econ* 20:275–289. <https://doi.org/10.1002/jae.820>
- Cerqueira PA (2013) A closer look at the world business cycle synchronization. *Int Econ Econ Pol* 10:349–363. <https://doi.org/10.1007/s10368-013-0233-z>
- Cerqueira PA, Martins R (2009) Measuring the determinants of business cycle synchronization using a panel approach. *Econ Lett* 102(2):106–108. <https://doi.org/10.1016/j.econlet.2008.11.016>
- Cesarino F (2006) The origins of the theory of optimum currency areas. *Hist Polit Econ* 38(4):711–731. <https://doi.org/10.1215/00182702-2006-017>
- Cimadomo J, Furtuna O, Giuliodori MM (2017) *Private and Public Risk Sharing the Euro Area*. Tinbergen Institute Discussion Paper No. 17–064/VI.
- Commandeur JF, Koopman SJ (2007) *An Introduction to State Space Time Series Analysis*. Oxford University Press, New York
- De Grauwe P (2020) *Economics of Monetary Union*, 13th edn. Oxford University Press, Oxford
- Dellas H, Tavlas GS (2009) An optimum-currency-area odyssey. *J Int Money Fin* 28(7):1117–1137. <https://doi.org/10.1016/j.jimonfin.2009.06.001>
- Dreyer JK, Schmid PA (2016) Does the Euro enhance Economic Growth? In: Paleckova I, Szarowska I (eds.) *Proceedings of the 15th International Conference on Finance and Banking*, pp. 72–83
- Durbin J, Koopman SJ (2000) Fast filtering and smoothing for non-stationary time series models. *J Am Stat Assoc* 92:1630–1638. <https://doi.org/10.1080/01621459.1997.10473685>
- Durbin J, Koopman SJ (2003) Filtering and smoothing of state vector for diffuse state space models. *J Time Anal* 24(1):85–98. <https://doi.org/10.1111/1467-9892.00294>
- Durbin J, Koopman SJ (2012) *Time Series Analysis by State Space Methods*, 2nd edn. Oxford University Press, Oxford
- Ferman B, Pinto C, Possebom V (2020) Cherry picking with synthetic controls. *J Pol Anal Manag* 39(2):510–532. <https://doi.org/10.1002/pam.22206>
- Fernández C, García Perea P (2015) The impact of the euro on euro area GDP per capita. *Documentos de Trabajo* 1530, Banco de Espana
- Ferrari A, Picco AR (2016) International risk sharing the EMU. *European stability mechanism working paper series*, No. 17
- Fidrmuc J, Korhonen I (2006) Meta-analysis of the business cycle correlation between the euro area and the CEECs. *CESifo Working* pp. 1693.
- Fingleton B, Garretsen H, Martin R (2015) Shocking aspects of monetary union: the vulnerability of regions euroland. *J Econ Geogr* 15(5):907–934

- Frankel JA, Rose AK (1997) Is EMU more justifiable ex post than ex ante. *Eur Econ Rev* 41(3–5):753–760. <https://doi.org/10.1093/jeg/lbu055>
- Frankel JA, Rose AK (1998) The endogeneity of the optimum currency area criteria. *Econ J* 108(449):1009–1025. <https://doi.org/10.1111/1468-0297.00327>
- Franks F, Barkbu B, Blavy R, Oman W, Schoelermann H (2018) Economic convergence the euro area: coming together or drifting apart? IMF working pp. 10
- Frydrych J, Burian S (2017) OCA indexes and convergence process Europe. *Sci Ann Econ Bus* 62(2):187–197
- Gabriel RD, Pessoa AS (2020) Adopting the euro: a synthetic control approach. *MPRA* pp. 99622.
- Harvey AC (1989) Forecasting, structural time series models and the kalman filter. Cambridge University Press, Cambridge
- Harvey AC, Koopman SJ (1992) Diagnostic checking of unobserved-components time series models. *J Bus Econ Stat* 10(4):377–389. <https://doi.org/10.1080/07350015.1992.10509913>
- Harvey AC, Koopman SJ (1997) Multivariate structural time series models. In: Heij C, Schumacher H, Hanzon B, Praagman K (eds) *System dynamics in economic and financial models*. Wiley and Sons, Chichester
- Harvey AC, Thiele S (2021) Cointegration and control: assessing the impact of events using time series data. *J Appl Econ* 36:71–85. <https://doi.org/10.1002/jae.2802>
- Harvey AC, Koopman SJ, Penzer J (1998) Messy time series. *Adv Econ* 13:103–143
- Helske J (2017) KFAS: Exponential family state space models in R. *J Stat Softw.* 78(10):1–39
- Hoffman M, Maslov E, Sorensen BE, Stewen I (2018) Are Banking and capital markets union complements? evidence from risk channels of risk sharing the eurozone. *CEPR discussion* pp. 13254.
- Horvath R, Komarek L (2003) Optimum currency area indices: evidence from the 1990s. *Warwick economic research* pp. 665.
- Horvath R (2005) Exchange rate variability, pressures and optimum currency area criteria: lessons for the central and eastern European countries. *Czech national bank working* pp. 8
- Ingram JC (1969) Comment: The Currency Area Problem. In: Mundell RA, Swoboda AL (eds) *Monetary Problems of the International Economy*. University of Chicago Press, Chicago, pp 95–100
- Ingram JC (1973) The case for European monetary integration. *essays international finance*. 98.
- Kaul A, Klößner S, Pfeifer G, Schieler M (2015) Synthetic control methods: never use all pre-intervention outcomes together with covariates. *MPRA* pp. 83790.
- Kenen PB (1969) The Theory of Optimum Currency Areas: An Eclectic View. In: Mundell RA, Swoboda AK (eds) *Monetary Problems of the International Economy*. University of Chicago Press, Chicago, pp 41–60
- Koopman SJ, Shephard N, Doornik JA (1999) Statistical algorithms for models state space using ssfpack2.2. *Econometrics J.* 2(2):107–160
- Koopman SJ, Harvey AC, Doornik JA, Shephard N (2008) *STAMP 8.2 Structural Time Series Analyser Modeller and Predictor*. London: Timberlake Consultants Ltd.
- Krugman P (1993) Lessons of massachusetts for EMU. In: Torres F, Giavazzi F (eds) *Adjustment and growth the european monetary union*. Cambridge University Press, New York, pp 241–266
- Kunroo HK (2015) Theory of optimum currency areas: a literature survey. *Rev Mark Integr* 7(2):87–116. <https://doi.org/10.1177/0974929216631381>
- Kwiatkowski D, Phillips PCB, Schmidt P, Shin Y (1992) Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root? *J Econ* 44:159–178. [https://doi.org/10.1016/0304-4076\(92\)90104-y](https://doi.org/10.1016/0304-4076(92)90104-y)
- Lin PC, Chen LY (2017) Assessing the growth effect of common currency adoption: synthetic control approach. *J Finance Econ* 5(6):269–280
- McKinnon RI (1963) Optimum currency areas. *Am Econ Rev* 53(4):717–725
- Mongelli FP (2002) new views on the optimum currency area theory: What is EMU telling Us. *European central bank, working* pp. 138.
- Mongelli FP (2008) European economic and monetary integration, and the optimum currency area theory. *European comission, European economy - Economic* pp. 302
- Mundell RA (1961) A theory of optimum currency areas. *Am Econ Rev* 51(4):657–665
- Mundell RA (1973) Uncommon arguments for common currencies. In: Johnson HG, Swoboda AK (eds) *The economics of common currencies*. Allen & Unwin Ltd, London, pp 114–132
- Pentecôte JS, Huchet-Bourdon M (2012) Revisiting the core-periphery view of EMU. *Econ Model* 29(6):2382–2391. <https://doi.org/10.1016/j.econmod.2012.06.036>

- Puzzello L, Gomis-Porqueras P (2018) Winners and losers from the Euro. *Eur Econ Rev* 108:129–152. <https://doi.org/10.1016/j.euroecorev.2018.06.011>
- Rose AK (2016) Why do estimates of the EMU effect on trade vary so much? *Open Econ Rev* 28(1):1–18. <https://doi.org/10.1007/s11079-016-9420-1>
- Sinn HW (2014) Austerity, growth and inflation: remarks on the Eurozone's unresolved competitiveness problem. *World Econ* 37(1):1–13. <https://doi.org/10.1111/twec.12130>
- Versteegen L, van Groezen B, Meijdam L (2017) Benefits of EMU participation. Estimates using the synthetic control method. CentER discussion paper. 2017–032
- Wortmann M, Stahl M (2016) One size fits some: a reassessment of EMU's core-periphery framework. *J Econ Integr* 31(2):377–413
- Zellner A (1963) Estimators for seemingly unrelated regression equations: some exact finite sample results. *J Am Stat Assoc* 58:977–992. <https://doi.org/10.1080/01621459.1963.10480681>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.