CrossMark

RESEARCH ARTICLE

Can Reform Waves Turn the Tide? Some Case Studies using the Synthetic Control Method

Bibek Adhikari[1] · Romain Duval[2] · Bingjie Hu[3] ·
Prakash Loungani[2]

Abstract A number of advanced economies carried out a sequence of extensive reforms of their labor and product markets in the 1990s and early 2000s. Using the Synthetic Control Method (SCM), this paper implements six case studies of well-known waves of reforms, those of New Zealand, Australia, Denmark, Ireland and Netherlands in the 1990s, and the labor market reforms in Germany in the early 2000s. In four of the six cases, GDP per capita was higher than in the control group as a result of the reforms. No difference between the treated country and its synthetic counterpart could be found in the cases of Denmark and New Zealand, which in the latter case may have partly reflected the implementation of reforms under particularly weak macroeconomic conditions. Overall, also factoring in the limitations of the SCM in this context, the results are suggestive of a positive but heterogenous effect of reform waves on GDP per capita.

Keywords Structural reforms · Synthetic control method · Liberalization · Labor and productivity market reforms · Growth

JEL Classification E02 · J08 · O40 · O43

✉ Bibek Adhikari
  badhik1@ilstu.edu

  Romain Duval
  rduval@imf.org

  Bingjie Hu
  Bhu1@worldbank.org

  Prakash Loungani
  ploungani@imf.org

[1] Department of Economics, Illinois State University, Normal, IL, USA

[2] International Monetary Fund, Washington, D.C., WA, USA

[3] World Bank Group, Washington, D.C., USA

1 Introduction

There is an extensive empirical literature on the impact of structural reforms on economic performance. Much of it relies on cross-country regressions of the outcome variable on a proxy of reforms. Such cross-country studies often face a conceptual difficulty in that each reform is highly specific to the country under consideration and that no two reforms, even in the same policy area, are comparable across countries.

Indeed, cross-country empirical studies often suffer from omitted variable biases due to unobserved idiosyncratic characteristics of each country, and measurement errors that are associated with reform indicators. For instance, Rodriguez and Rodrik (2001) express skepticism over whether there is a general unambiguous relationship between trade policy reforms and growth. Bhagwati and Srinivasan (2002) review previous studies on linkages between openness and growth and also critique the cross-country regression methodology; they argue that a more compelling source of evidence is in-depth case studies of specific countries. Case studies, however, also have limitations, particularly the lack of a clearly defined counterfactual. The selection of appropriate comparison units is crucial for the success of case studies (Abadie et al. 2015). The comparison units are required to be sufficiently similar in order for the case study to yield useful results.

A recent econometric method, the Synthetic Control Method (SCM), developed by Abadie et al. (2003), and extended by Abadie et al. (2010, 2015), provides a systematic way to conduct quantitative case studies that combines the careful and meticulous description of case studies with precise numerical estimation and inference of regression studies.. It allows us to demonstrate how we construct the counterfactual, the synthetic comparison unit, in a transparent, data-driven manner. In the Synthetic Control Method, the control unit is selected as the linear combination of all potential comparison units that have the most similar characteristics to those of the case of interest, prior to the treatment. When interpreting the results, we also benefit from the precise quantitative estimation of the dynamic treatment impact given by the SCM. By construction, the pre-reform performance of the synthetic control unit was similar to that of the reformer country. The estimated reform impact is given by the difference in the post-reform values of outcome variable between the treated unit and the synthetic control unit.

In this paper, we apply the Synthetic Control Method to investigate the impact of major product and labor reform episodes on income per capita[1] in advanced economies in the past three decades. We identify major reform episodes based on a review of relevant history and a new database on structural reforms constructed by Duval and others (2018). Drawing from previous literature on the reforms, we designate the beginning year of the reform episode as the treatment date. We focus our study on cases for which we have a sufficiently good pre-treatment fit when implementing the SCM. To assess whether the control group is indeed a good counterfactual, we use a measure of fit developed by Adhikari and Alm (2016).

In the selection of control units, we restrict our pool of candidate countries (the donor pool) to those that are also advanced economies. While this approach shrinks the

---

[1] We have tried to examine the reform impact on alternative outcome variables such as total factor productivity, employment rate, and female labor participation rate. However, we did not obtain meaningful results because given the volatility in these outcome variables, we were unable to find synthetic units that could provide sufficiently good pre-treatment fit for the treated countries. In addition, we would like to focus on the impact of reforms on overall economic performance.

size of the donor pool, it also avoids interpolating across countries with heterogeneous characteristics. To construct an appropriate counterfactual, we also exclude those countries from the donor pool that also had a major reform within five years either before or after the starting year of the reform episode.

Initially, we selected nine cases: two reform episodes in Australia, two in New Zealand, and two in the Netherlands, in the 1980s and 1990s, respectively, one reform episode each in Denmark, and Ireland in the 1990s, and one in Germany in the early 2000s. After the implementation of the SCM, we decide to drop the cases of Australia, New Zealand and the Netherlands in the 1980s, due to poor pre-treatment fit, as it is not feasible to assess the impact of structural reforms in such cases.

Therefore, this paper is based on six case studies, namely those of Australia, New Zealand, Denmark, Ireland, and the Netherlands in the 1990s, and Germany in the early 2000s. All six cases meet pre-treatment fit criteria, although fit is weaker in some cases than in others. Our analysis using SCM suggests that structural reforms had positive impact on income per capita in four out of our six cases, compared to the synthetic control unit. The estimated impact on income per capita varies widely across the four cases but is generally sizeable. Our results suggest that there is much heterogeneity in the size of the estimated effects of structural reforms both across countries and over time. For instance, in the case of New Zealand, the estimated effect of reforms on real GDP per capita is negative until four years after the reform. The income per capita failed to catch up with its synthetic counterpart during the three years following the reforms, which may partly reflect the implementation of reforms under particularly weak macroeconomic conditions.

In contrast, in the case of Germany, the estimated impact of reforms in real GDP per capita increases from 5% in 2004, the year following the reform to 9% in 2008, five years after the reform started. For Denmark, our results show that the estimated reform impact is close to zero.

Our paper not only improves over traditional case studies by using a data-driven approach to the selection of cases and comparators, but also contributes to the broad literature on the impact of structural reforms by improving upon previous studies that use a traditional cross-country regression approach.

First, instead of relying on the average treatment effect from standard OLS regressions, we are able to identify the heterogeneous reform impact of each treated country by looking at the difference in post-reform outcome values between treated and synthetic control units over time. Therefore, we can unveil the reform impact that is specific to the country of interest, which would have been masked in a standard cross-country regression analysis.

Second, we are able to avoid over-extrapolation of the results by carefully choosing our counterfactual, which consists of a linear combination of control units that closely resembles the treated unit in pre-treatment outcome. We also require the donor pool to consist of countries with similar levels of income, and avoid comparing countries with vastly different characteristics.

Third, with the application of SCM, we explicitly obtain the contributions of each comparison unit to the counterfactual of interest. In contrast to the standard regression studies, in which the counterfactual is constructed using the simple average of all control units, our study reports the weights for each comparison unit that contributes to the counterfactual, which enable us to conduct both quantitative and qualitative assessment of the reform impact.

Our paper contributes to a small but growing literature that uses the SCM to assess the impact of policy reforms- (Campos and Kinoshita 2010[2]; Abadie et al. 2015; Billmeier and Nannicini 2013; Abadie et al. 2010; Adhikari and Alm 2016) and non-policy shocks (Abadie and Gardeazabal 2003; Cavallo et al. 2013) on economic outcomes. The SCM employed in our paper allows us to implement a systematic way to select comparison units and a quantitative way to analyze the reform impact, while still allowing qualitative analysis and comparison between the country of study and control units.

There are, however, several limitations that apply to papers such as ours and others in the literature, that seek to estimate the impact of a treatment on country-level macroeconomic outcomes. The method cannot control for other major idiosyncratic shocks that may occur around the treatment date, it does not address all sources of endogeneity such as reverse causality—a limitation to the extent that reforms are partly driven by poor growth prospects, and finding a clean treatment is difficult in the present context—major reforms waves are often phased in over several years, and some control countries may still have implemented minor reforms that could add up to a noticeable impact. The latter two limitations imply that our estimates are more likely to underestimate, rather than to overestimate, the true impact of the treatment—a major wave of structural reforms.

Another challenge is the fact that reforms often come in waves. Since the reforms may come in a sequence, it may be difficult to estimate the impact of each reform within the same wave. In our study, for the feasibility of computation, we identify the treatment date as the beginning of the major reform episodes. Therefore, we do not distinguish the individual reform measures within the same episode. We also acknowledge the fact that we do not examine product and labor market reforms separately, as these two types of reforms tend to take place in the same episode in most of our cases.

The next two sections describe the implementation of the Synthetic Control Method and our data. Section IV discusses the results from the case studies. Section V concludes.

2 Synthetic Control Method

In order to isolate the impact of structural reforms from other influences, we employ the synthetic control method (SCM), a data-driven way of finding the counterfactual (i.e., how the trajectory of outcome variables would evolve had the structural reform not been implemented) in generalized Difference-in-Differences (DID) estimation. DID estimation consists of identifying a specific treatment (structural reform in our case), and then comparing the difference in outcomes before and after the treatment for the treated country to the difference in outcomes before and after the treatment for the untreated countries. The primary motivation to use synthetic control is the belief that the effect of a particular intervention can be empirically assessed only by comparison with the appropriate counterfactual. The SCM was developed by Abadie and Gardeazabal (2003) and expanded by Abadie et al. 2010, 2015). The use of SCM has been growing, and it is now being applied to a very diverse set of topics, including the effects of trade liberalization on growth (Billmeier and Nannicini 2011; Billmeier

---

[2] Campos and Kinoshita (2010) use the synthetic control method to analyze the impact of financial reforms on FDI inflows, as a robustness check of their regression results.

and Nannicini 2013), and the 5economic effect of the 1990 German reunification in West Germany (Abadie et al. 2015).[3]

There are various advantages of using the SCM, and because of these advantages, Athey and Imbens (2017) consider the development of the SCM to be, "… arguably the most important innovation in the policy evaluation literature in the last 15 years". First, in comparative case studies the researcher is allowed to choose the comparison group that resembles the counterfactual. However, Abadie et al. (2010) argue that this introduces substantial ambiguity about how comparison groups are chosen since researchers select comparison groups based on subjective measures of affinity between the treated country and the untreated countries. SCM solves this problem by creating a synthetic control unit that resembles the treated unit in the pre-treatment period, using a weighted average of all other unaffected units from the same income group. The weights are chosen so that the pre-treatment outcomes and the covariates of the synthetic control unit match on average the outcomes and covariates of the treated unit.

Second, SCM obtains the estimates one case at a time, which allows us to explore the heterogeneity of the effects of the reform in a very flexible way as opposed to the DID estimation where, for the most part, only average effects are reported. This is especially important because the design of each reform and the broader environment under which they are carried out all matter for their impact, thus, assuming the impact will be homogenous is not plausible. Indeed, we find very heterogeneous impact of the reforms that an average effect would easily mask.

Third, SCM can substantially reduce any potential endogeneity problem caused by the omitted variables. Abadie et al. (2010) prove that if a synthetic country can be found such that it matches the pre-treatment trajectory of the outcome variable of the treated country, then the size of the bias caused by time varying unobserved confounders in the difference between the post-treatment outcome variable for the treated and the synthetic control countries goes towards zero as the pre-intervention period increases. The intuitive explanation is that only countries that are alike in both observed and unobserved predictors of the outcome variable as well as in the effect of those predictors on the outcome variable would produce similar trajectories of the outcome variable over extended periods of time. To ensure that our estimates substantially reduce the biases caused by the omitted variables, we use pre-treatment fit index to select the cases where pre-treatment trajectory of outcome variable between treated and control countries match very well.

Fourth, by taking the weighted average of control countries, SCM makes explicit the contribution of each comparison unit to the counterfactual of interest. This allows us to combine quantitative and qualitative techniques to analyze similarities and differences between the treated unit and the synthetic control. Contrast this with regression technique like conventional DID where the counterfactual is constructed using the simple average of

---

[3] Abadie and Gardeazabal (2003) use the approach to assess the negative impact of the violent conflict in the Spanish Basque Country on economic growth. Abadie et al. (2010) use it to estimate the impact of a large anti-tobacco initiative in California on the per capita sales of cigarettes, the impact of the terrorist attacks on electoral outcome (Montalvo 2011); the effects of relaxing restrictions on home equity lending on retail spending by households (Abdalah and Lastrapes 2012); the impact of natural disasters on economic growth (Cavallo et al. 2013); the effect of civil conflict on economic growth (Dorsett 2013); the impact of nutrition policies on dietary behavior and childhood obesity (Bauhoff 2014); the effect of immigration laws on demographic composition (Bohn et al. 2014); the impact of decrease in police enforcement on traffic fatalities and injuries (DeAngelo and Hansen 2014); and the impact of major natural resource discoveries on economic growth (Smith 2015).

all control countries (conditional on covariates). However, these weights are implicitly calculated and are not reported in practice, which makes qualitative techniques very difficult. Moreover, the weighted average of similar countries can provide much credible counterfactual than the simple average of all control countries.[4]

Finally, Bertrand et al. (2004) show that standard errors in the papers that use DID estimation with many years of data severely understate the standard deviation of the estimators due to serially correlated errors. Our use of SCM allows us to use placebo experiments to draw valid inference in the presence of correlated errors. Since the placebo experiments do not make parametric assumptions about the error structure, it does not suffer from the over-rejection bias of the standard t-test.

There are also some limitations of the SCM. First, the method does not address other sources of endogeneity such as reverse causalisty. As most of the structural reforms are motivated by expectation of future growth prospects, this would bias the estimates obtained from SCM as long as growth expectations are not captured by the unobservable heterogeneity included in the estimation.

Second, the ideal control group for the SCM estimation consists of countries that don't experience any major idiosyncratic shocks to their economy that can affect the outcome of interest in the sample period, especially in the post-reform period. For practical purposes, we only exclude other major reformers from the donor pool. Since we do not exclude countries that might have had moderate or minor reforms in other structural areas or other growth enhancing reforms. There might be cases in which some control countries have had many minor reforms and their effects on the outcome variable add up to become meaningful. Thus, in this sense, our estimates most likely underestimate the true impact of the structural reforms. However, if some of our control countries experienced idiosyncratic negative shocks in the post-reform period, that can overestimate the true impact of structural reforms.

Third, the SCM assumes no anticipation effects, which implies that key variables do not change in anticipation of the future structural reform, before the structural reform actually takes place. In our case, if firms anticipate structural reforms and delay their hiring, investment, or expansion decisions until the passage of the reform, it can contaminate our estimated treatment effects.[5]

Finally, the SCM estimates only the reduced-form aggregate impact of structural reforms on economic growth, thus, this paper cannot explain the underlying mechanisms through which such reforms affect economic growth.

In the next section, we provide more details about the synthetic control methods.

2.1 Synthetic Control Method

Suppose that we have $J + 1$ countries, where country 1 (i.e., a treated country) adopts structural reform at time $T_0 + 1$ and the remaining J countries act as potential controls called the donor pool. Let T_0 be the number of pre-intervention periods, with $1 \leq T_0 \leq T$. Also, let Y_{it}^{NR} be the outcome variable observed for country i at time t with no reform

---

[4] See Abadie, Diamond, Hainmueller (2015) for details on the comparison of SCM and traditional regression techniques.

[5] The firms may also make their decision in advance, in anticipation of the reforms. In either case, the anticipation effects make it difficult to identify the effective treatment date.

(NR), and Y_{it}^R be the outcome variable with reform (R). The observed outcome variable can be written as:

$$Y_{it} = \begin{cases} Y_{it}^{NR} & \textit{in the absence of reform} \\ Y_{it}^R \equiv Y_{it}^{NR} + \tau_{it} D_{it} & \textit{in the presence of reform} \end{cases}$$

where $\tau_{it} = \left(Y_{it}^R - Y_{it}^{NR}\right)$ is the effect of the reform for country i at time t and $D_{it} = 1$ if $t > T_0$ and $i = 1$, and $D_{it} = 0$ otherwise.

For any treated country, we can observe Y_{it}^R. However, we need to estimate the counterfactual Y_{it}^{NR}, which is the outcome variable of the country that adopted structural reform had the country not adopted it. In order to estimate the counterfactual, we use the linear factor model of the form:

$$Y_{it}^{NR} = \propto_t + \theta_t Z_i + \lambda_t \mu_i + \varepsilon_{it}$$

where $\propto_t$ is an unknown common factor with constant factor loadings across countries, Z_i is a vector of observed covariates with coefficients θ_t, μ_i is a $(F \times 1)$ vector of unknown parameters, λ_t is a $(1 \times F)$ vector of unobserved common factors, and ε_{it} are idiosyncratic error terms with zero mean. Note that this specification allows the effects of confounding unobserved characteristics to vary with time $(\lambda_t \mu_i)$, unlike in conventional difference-in-differences that allows for the presence of unobserved confounders but restricts the effects of those confounders to be to be time invariant $(\lambda \mu_i)$.

Define a synthetic control unit as a weighted average of the units in the donor pool. That is, a synthetic control can be represented by a $J \times 1$ vector of weights, $W = (w_2, \ldots, w_{J+1})'$ such that $w_j \geq 0$ for $j = 2, \ldots, J+1$ and $w_2 + \ldots + w_{J+1} = 1$, where vector W represents a potential synthetic control. Then the outcome variable for each potential synthetic control unit is given by:

$$\sum_{j=2}^{J+1} w_j Y_{jt} = a_t + \theta_t \sum_{j=2}^{J+1} w_j Z_j + \lambda_t \sum_{j=2}^{J+1} w_j \mu_i + \sum_{j=2}^{J+1} w_j \varepsilon_{jt}$$

Now suppose that there are $\left(w_2^*, \ldots, w_{J+1}^*\right)'$, such that the following holds.

$$\sum_{j=2}^{J+1} w_j^* Y_{j1} = Y_{11}, \ldots, \sum_{j=2}^{J+1} w_j^* Y_{jT_0} = Y_{1T_0}, \sum_{j=2}^{J} w_j^* Z_j = Z_1$$

Thus, the treatment effect at time $t \in \{T_0 + 1, \ldots, T\}$ can be estimated by:

$$\hat{\tau}_{1t} = Y_{1t} - \sum_{j=2}^{J+1} w_j^* Y_{jt}$$

To find the optimal weights, let the $(T_0 \times 1)$ vector $K = (k_1, \ldots, k_{T_0})'$ define a linear combination of pre-intervention outcomes, $\overline{Y}_j^K = \sum_{s=1}^{T_0} k_s Y_{js}$, where j belongs to $\{1, \ldots, J+$

1}, $j = 1$ denotes treated country and $j \neq 1$ denote donor countries. Consider M of such linear combinations defined by the vectors $(K_1, \ldots, K_M)$. Let X_1 be $(Z_1', \overline{Y}_1^{K_1}, \ldots, \overline{Y}_1^{K_M})'$, the vector of pre-treatment variables that we aim to match as closely as possible for the treated country. Let X_0 be the matrix where each column of the matrix is a vector of same pre-treatment variables for each potential donor country. The synthetic control algorithm chooses W^* to minimize the distance $\|X_1 - X_0 W\|$ $V = \sqrt{(X_1 - X_0 W)'} \, 'V(X_1 - X_0 W)$, where V is a symmetric, positive semi-definite and diagonal matrix such that the mean square prediction error (RMSPE) of the outcome variable is minimized for the pre-intervention periods. Note that the inferential procedure is valid for any choice of V; however, by minimizing the RMSPE, the algorithm assigns larger weights to those pre-treatment variables that have the highest predictive power.[6]

2.2 Pre-Treatment Fit Index

To assess whether the comparison country created using SCM is a good counterfactual, we need some measure of how well it resembles the treated country before the treatment. Abadie et al. (2010) use root mean square prediction error (RMSPE) of the outcome variable to measure fit or lack of fit between the path of the outcome variable for treated country and its synthetic counterpart, defined as:

$$RMSPE = \sqrt{\frac{1}{T_0} \sum_{t=1}^{T_0} \left(Y_{1t} - \sum_{j=2}^{J+1} w_j^* Y_{jt} \right)^2}$$

However, we use "pre-treatment fit index" developed by Adhikari and Alm (2016) to assess the overall quality of the pre-treatment fit. There are two main advantages of using pre-treatment fit index. First, pre-treatment fit index normalizes RMSPE, which makes it possible to compare the fit between the synthetic control method across different outcome variables and different countries, such as when GDP per capita varies quite significantly across advanced economies and low income economies. Second, this approach provides an index number that makes assessing the quality of fit very intuitive.

Define the benchmark RMSPE as the RMSPE obtained from the zero fit model, or:

$$Benchmark\ RMSPE = \sqrt{\frac{1}{T_0} \sum_{t=1}^{T_0} (Y_{1t})^2}$$

Then the pre-treatment fit index is defined as the ratio of the RMSPE and the benchmark RMSPE:

$$Fit\ Index = \frac{RMSPE}{benchmark\ RMSPE}$$

---

[6] See Appendix V in Abadie and Gardeazabal (2003) for details.

A fit index of $[0, X]$ means that the RMSPE is equivalent to the RMSPE obtained when the difference between the treated and the synthetic unit is X percent for each pre-treatment year. If the RMSPE is 0, then the fit index will be 0, indicating a perfect fit; if the RMSPE is equal to the benchmark RMSPE, then the fit index will be 1, indicating that the fit is equal to that created by a zero fit model. The fit index will be greater than 1 in cases where the outcome variable of the treated unit is bigger (or smaller) by a magnitude of two or more (or half or less) than that of the synthetic control. A fit index greater than 1 indicates a poor fit, such that this synthetic unit should be discarded, especially when the outcome variable is GDP per capita where it is safe to say that a synthetic country with twice (or half) the GDP per capita of the treated country cannot reasonably be a good counterfactual.

2.3 Inference

In a comparative case study where identification of the treatment effect arises from the change in policy by a small group of countries and where data are usually of small sample size, standard large-sample approximations that are typically used for inference are not appropriate. However, placebo experiments as used in Abadie and Gardeazabal (2003), Bertrand et al. (2004), and Abadie et al. (2010) can be used to evaluate the significance of treatment effects.

The essence of placebo experiments is to test whether the estimated impact of the structural reform could be driven entirely by chance. Specifically, we conduct a series of placebo experiments by iteratively estimating the "placebo" treatment effect for each country in the donor pool (i.e., untreated countries) by first assuming that these countries implemented structural reform in the same year as our country of interest and by then running the synthetic control method. In selecting a synthetic control group for the countries in the placebo pool, we omit the treated countries.

This iterative procedure provides a distribution of estimated placebo treatment effects for the countries where no intervention took place. If the placebo experiments create enough placebo treatment effects of magnitude greater than the one estimated for the treated country, then we conclude that there is no statistically significant evidence of an effect of reform in the treated country. If the placebo experiments show that the treatment effect estimated for the treated country is unusually large relative to placebo treatment effects for countries that did not implement structural reforms within our sample period, then we conclude that there is a statistically significant evidence of an impact of reform in the treated country. Note that one limitation of the placebo test is that it can't rule out the chance that the estimated impact of the structural reforms could be driven by idiosyncratic shocks or other policy changes; it can only test the rarity of obtaining the effect of such magnitude. Thus, it is important to carefully construct the control group to minimize the presence of idiosyncratic shocks.

3 Data and Selection of Cases

3.1 Data

To study the growth effects of structural reforms, we create a cross-country panel data set for the period between 1960 and 2011 by combining multiple data sources. The

outcome of interest is log of real GDP per capita. Following endogenous growth theory, the vector of covariates includes the initial level of GDP per capita (log), stock of physical capital per capita (log), stock of human capital per capita, democracy, trade openness, and population growth. These variables are meant to capture, in a broad sense, the impact of institutions, demography, and macroeconomic conditions on top of traditional growth accounting variables such as the stock of physical and human capital.

The primary data source is Penn World Table (PWT) version 8.1, which is supplemented the World Development Indicators dataset, and the World Economic Outlook dataset from the IMF. We use democracy indicators from Polity IV. Following Persson et al. (2007), we classify a country as democratic if the "polity2" in the Polity IV data set is strictly positive.

3.2 Selection of the Treated Countries and Control Groups

In this exercise, we define the treatment as a major reform wave in the area of product and labor market institutions. In the determination of reform dates, our main source of information are the OECD country surveys and the structural reforms dataset underlying Chapter 3 of the April 2016 World Economic Outlook (Duval et al. 2018; IMF 2016). We define the treatment date as the starting year of the reform episode.

We require pre-treatment observations of GDP per capita to be available to calibrate the synthetic control. We restrict the sample period to 10 pre-treatment years to calibrate the synthetic unit and 5 post-treatment periods to evaluate the impact of the treatment. Thus, for each case the sample window consists of 16 years.

In order to select a comparable control group that can provide reasonable counterfactual (i.e., how GDP per capita would have evolved in the treated country if there had been no treatment), we restrict the pool of donor countries in several ways.

First, from the donor pool, we remove those countries that also had a major reform within a window of five years before or after the treatment year. We do so because the synthetic unit is meant to reproduce the level and the trend in the outcome variable that would have been observed for the treated unit in the absence of treatment. Thus, including countries in the donor pool that also underwent major reform during the sample period implies that the synthetic unit is not reproducing the potential outcome in the absence of treatment.

Second, we remove any country that was under major civil or ethnic violence during the sample period. The data on ethnic and civil violence comes from the Correlates of War project.

Third, if there are any missing values for the outcome variable in the sample window then that country is dropped from the donor pool.

Finally, we restrict the donor pool to the countries that belonged to the same income group during the year of treatment, where countries belong to one of two income groups: Advanced economies or Non-Advanced economies (i.e., emerging and low income). The income classification closely follows International Monetary Fund's classification. However, for some early years this information is missing. In that case, we label the countries to the income class they belong when the data first becomes available.

There are two main advantages of restricting potential controls to the same income group. First, it avoids biases caused by interpolating across countries with very different characteristics. That is, even if we are able to find a synthetic unit with good pre-treatment, interpolation biases may be large if the linear factor model used in the estimation of the synthetic control doesn't hold over the entire set of regions in any particular sample. This happens if the relationship between the outcome variables and the predictors is highly nonlinear and the combination of two extreme donor units is used to construct a synthetic unit that has average value of the covariate. In a similar spirit, it also controls for unobservable characteristics associated with the level of economic development and any other secular changes over time that might affect countries from different income groups differently. The next section describes the implementation of the synthetic control method. In the next section we introduce our cases of study, briefly describe the major reforms in labor and product markets, demonstrate our results using the synthetic control method, and discuss the implications of our findings.

4 Cases and Results

4.1 Structural Reforms in New Zealand in the 1990s

4.1.1 The Relative Decline in Economic Performance Triggered the Reforms

The primary goal of the reforms was to reverse the decline in New Zealand's economic performance compared with other OECD countries, which showed in a growing GDP per capita gap vis-à-vis the OECD average (OECD 1991). Economic performance was seen as being hampered by distortive economic policies, including high tariffs, import licensing and quotas, complex regulations and subsidies. In response to such ineffi-ciencies, the government implemented a wave of structural reforms in the 1980s, which aimed to reduce state control and re-orient the economy toward market-based resource allocation.

Starting around 1991, a second wave of significant reforms took place. The mea-sures included enhancing labor market flexibility and strengthening competition in product markets through domestic deregulation and further reduction in trade protection.

4.1.2 Deregulation of the Labor Market

- *Collective wage bargaining*. The Employment Contracts Act of 1991 radically altered industrial relations by introducing a decentralized wage-bargaining structure that replaced the centralized system under which wages were determined by an arbitration court. Under the new law, each employee could choose to negotiate either an individual employment contract or to be bound by a collective contract. Likewise, an employer can choose to negotiate an individual contract with an employee or a collective contract. In negotiations the parties could choose their own bargaining agent. No contracts were allowed to require any person to join, not join or leave a union.

- *Unemployment benefits*. In December 1990 the government implemented a package of measures to increase work incentives. Unemployment benefits were reduced and eligibility criteria were tightened. The wait period before benefit entitlement commences was increased from 6 to 26 weeks.
- *Active labor market policies*. A work-for-benefit scheme (the "Community Taskforce") was established in 1991, with the aim of providing work experience and also work-testing benefit beneficiaries. Under this scheme beneficiaries could volunteer, or be required, to work three days per week on a community project, earning an increment on their unemployment benefit. In July 1992, the work-for-benefit schemes were expanded to incorporate training assistance, with a focus on the long term unemployed.
- *Pension system*. In July 1991, the standard eligibility age for retirement benefits was raised from 60 to 65, to be phased in gradually between 1992 and 2001. Benefits were frozen for two years as opposed to inflation indexation.

4.1.3 Opening the Economy to Greater Competition

- *Trade policy*. In March 1990 the government announced a post-1992 tariff program in which most tariffs would fall to a maximum level of 10% between 1993 and 1996. Further programs of tariff reductions were announced and implemented in the second half of the 1990s that led to major tariff reductions on many goods and a rationalization of the tariff structure.
- *Public ownership*. In the broader context of an overall reduction in direct government assistance to industry (Evans et al., 1996), the early 1990s witnessed a wave of full or partial privatization of state-owned enterprises, including in electricity, telecommunications, railways and hotels.
- *Barriers to entry*. All statutory protections of the telecommunications monopoly were removed, and other firms were able to provide network services in competition with the incumbent (Telecom) in the course of 1989. Taxi licensing was liberalized by the Ministry of Transport in 1990. Fares were allowed to fluctuate based on market conditions, and geographic restrictions on the number of licenses were removed. The 1992 Electricity Act introduced potential competition between energy retailers by allowing sale to any consumer using existing distribution lines. The reform was implemented in two steps in April 1993 (normal consumers) and April 1994 (large consumers). In July 1994 the government separated electricity transmission (Transpower) from generation (Electricorp). The 1992 Gas Act and 1997 Information Disclosure Regulations fostered competition in the gas market through the termination of exclusive franchise arrangements, information disclosure regulation, and a removal of price controls.

4.1.4 Macroeconomic Context

The wave of reforms was initiated during a downturn—the 1990–1991 recession, in part in order to regain external competitiveness. Macroeconomic policy was restrictive. Policy rates stood above 10% until 1991, as the Reserve Bank sought to achieve rapid

disinflation in order to meet the ambitious inflation objectives set by the 1989 Reserve Bank Act that made the central bank independent and introduced inflation targeting. The real effective exchange rate had appreciated sharply in the late 1980s, contributing to a further tightening of financial conditions. Fiscal policy, which had been accommodative earlier, also became restrictive around 1991, contributing to an improvement in the cyclically-adjusted balance of over 3 percentage points between 1991 and 1993.

4.1.5 Impact of the Reforms

While real GDP per capita grew by 2.5% over the five-year period 1985–1990, it increased by 14.3% during 1991–1996—by 9% during 1990–1996, due to the large drop during the 1991 recession. In contrast, average real GDP per capita across all advanced economies grew by 16.4% during 1985–1990, versus 9.4% over 1991–1996. This would seem to suggest that the reforms had a significant impact on New Zealand's growth.

However, our analysis using the synthetic control method does not confirm a positive impact of the reforms. Relative to its synthetic control group, New Zealand's GDP per capita dropped sharply in 1991—the assumed treatment date, reflecting the recession that hit in 1990–91 on the back of tight monetary conditions. Indeed, our results show that the real GDP per capita in New Zealand was about 9% lower than that of its synthetic counterpart. While New Zealand's GDP per capita level subsequently caught up to that of the synthetic control group, it did not exceed even five years after the reform. Its real GDP per capita was still 3% lower than that of synthetic New Zealand in 1996. In addition, our estimated dynamic p-values[7] are volatile over time, which suggests the estimated reform impact is insignificant in the case of New Zealand.

Therefore, on the basis of our exercise, it is unclear whether the reforms gave a significant boost to living standards, which may reflect the fact that they were implemented in times of recession and restrictive macroeconomic policies (Fig. 1).

Previous studies of this reform episode in New Zealand also failed to reach a consensus on whether the reforms of the 1990s improved economic performance. Evans and others (1996) argue that the inadequate macroeconomic policy mix and sequencing of reforms partly explain why reform was not as successful as expected, but emphasize the "broadly positive macroeconomic outcomes". Glyn (2005) point out New Zealand's poor productivity performance in the 1990s relative to other advanced economies such as Australia and Ireland, which is not indicative of a strong positive effect of reforms. Other scholars (McCann 2003; Boulhol et al., 2008) have hinted at New Zealand's lack of agglomeration effects and geographical distance from key markets as a possible contributing factor that preventing reforms from unleashing major efficiency gains through economies of scale and technology absorption. A study by the New Zealand Treasury (Galt 2000) suggests that growth benefits from the reforms might be slow to appear and not fully measured. However, it is unclear why this measurement issue would have worsened after the reforms relative to comparator countries.

---

[7] Calculated using "in space" placebo experiments. See, for instance, Abadie et al. (2015). The dynamic p-value indicates the likelihood of obtaining an estimate at least as large as the one obtained for the treated unit at each post-reform period analyzed. The placebo experiment graphs for each of the treated countries are provided in the Appendix (Figs. 7, 8, 9, 10, 11 and 12).
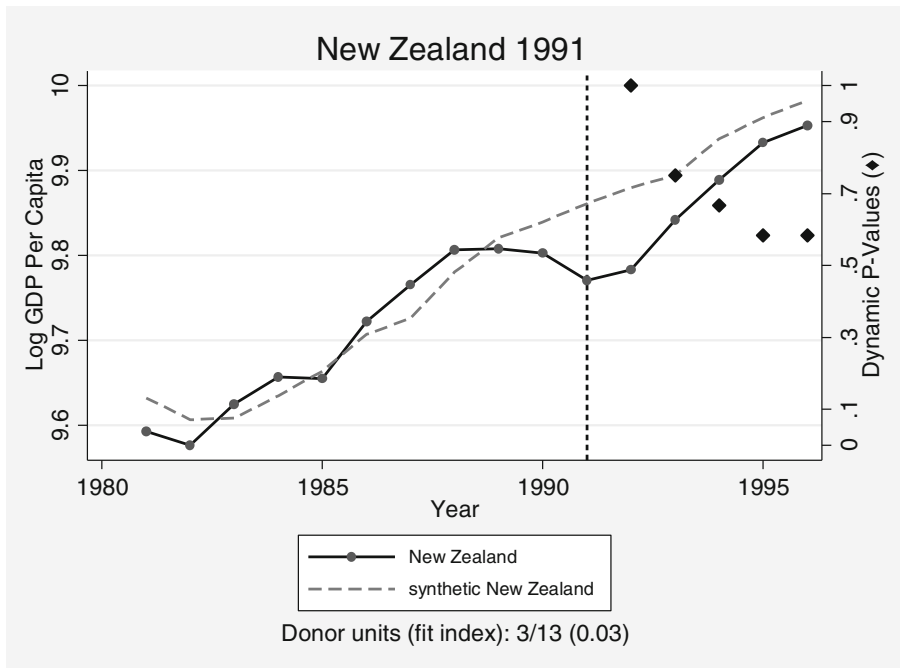
Fig. 1 GDP per capita in New Zealand and the synthetic unit over time. Note: Synthetic New Zealand consists of Greece (38%), Portugal (22%) and the United States (40%)

4.2 Labor and Product Market Reforms in Australia in the 1990s

4.2.1 High Unemployment and Poor Growth Performance Triggered a (New) Wave of Reforms

The Australian economy entered recession in the early 1990s and suffered from slow growth and high unemployment, which rose above 10%. This took place against the background of disappointing underlying growth performance despite a first wave of reforms in the 1980s, with a steady decline in the ranking of Australia's income per capita among OECD countries.

In response, the Australian government implemented a (new) wave of major reforms to the labor and product markets in the mid-1990s in order to promote growth and reduce unemployment.

4.2.2 Labor Market Reforms

- *Collective bargaining.* The wage system started to transition away from long-standing centralized bargaining (Briggs and Buchanan 2000) toward firm-level bargaining. The 1992 Industrial Relations Legislation Amendment Act removed the requirement for certified enterprise-level agreements to meet a "public interest" test. On March 30, 1994, the 1993 Industrial Relations Reform Act became law. For the first time since 1904, provision was made for registering collective non-union agreements. Incentives and penalties were put in place for single-employer

bargaining, and non-union agreements could be formally registered. The role of the central government was reoriented away from settling disputes toward facilitating enterprise-level agreements. Following the shift, union density declined significantly, while the growth in non-union collective agreements picked up.

- *Active labor market policies.* In response to high unemployment rates, the Australian government launched the Working Nation program in May 1994 to ensure that the long-term unemployed could benefit more from the economic recovery. The key initiative, the Job Compact program, focused on persons who had been receiving unemployment income support for at least 18 months, offering job placement—including wage subsidies—for 6 to 12 months and intensified job-search assistance afterward. Labor market assistance was also strengthened for non-job-compact workers, with particular emphasis on those at risk of becoming long-term unemployed. Another major initiative was the restructuring of the Commonwealth Employment Service (CES) in 1998 in relation to the provision of case management services. These services were opened to competition in order to provide a more effective and responsive service to both employers and the unemployed. Incentives for recipients of social security payments (including unemployment income support) to seek additional income were also improved by lowering the withdrawal rate for benefits from 100% to 70% of each additional dollar earned. Finally, under the Youth Training Initiative, unemployed people under the age of 18 were provided with case management assistance and subsidies to find suitable work, training or education placement.

4.2.3 Strengthening Competition in Product Markets

- *Barriers to entry.* In August 1994, the Council of Australian Governments (COAG) agreed to a package of reforms that established: i) a regime to provide access to essential facilities such as electricity grids, gas pipelines, airports, rail networks, postal delivery services, communication channels and seaports; ii) the application by individual jurisdictions of agreed principles on structural reform of public monopolies, competitive neutrality between the public and private sectors where they compete, and a program for the review of regulations restricting competition. In July 1993 the COAG agreed to reforms to create a competitive market for bulk electricity in southern and eastern Australia from July 1995. In June 1995, an access price system for the Commonwealth's rail track was announced, allowing competition between the National Rail Corporation and private operators. In the following month, heads of state governments agreed that a competitive electricity market would commence in July 1995. In June 1994, more competition was allowed in the gas transmission sector, when the Moomba-Sydney Pipeline was sold and the sale legislation incorporated a series of provisions to encourage competition in gas transmission. In July 1997, full and open competition began in telecommunications.
- *State ownership.* In the railways sector, the business units of Australian National were privatized in mid 1997. The privatization of Australia's largest airline company Quantas started in March 1993, and was completed in late 1995. The government also reduced its share in the Commonwealth Bank of Australia from 70 to 51% in 1994.

- *Competition policy.* The August 1994 package of reforms agreed by the COAG introduced a system in each jurisdiction to carry out surveillance of prices charged by utilities and other corporations with high levels of monopoly power, as well as the establishment of the Australian Competition Commission (ACC) and the National Competition Council (NCC) to make recommendations in relation to access and pricing surveillance issues and to advise on matters to be determined by governments. In April 1995, the COAG agreed to implement the National Competition Policy (NCP) package. A key element of the package was the Commonwealth's Competition Policy Reform Act (1995). It extended the coverage of the competitive conduct rules to unincorporated and government business enterprises, and amended the Price Surveillance Act to extend price oversight to State- and Territory-owned enterprises. It took effect in November 1995.

4.2.4 Macroeconomic Context

This wave of labor and product market reforms that started roughly around 1994 was implemented in the aftermath of a recession in the early 1990s. As a result, the macroeconomic policy stance was eased in 1993; the Australian government provided a small-scale fiscal stimulus to support the growth recovery, the central bank cut the policy rate by 1 percentage point, and the exchange rate depreciated. However, with the exception of this initial recession-driven impulse, macroeconomic policy was mostly restrictive throughout the reform period. On the fiscal policy front, priority was given to consolidation, helping bring down the cyclically-adjusted primary deficit from about 4% of potential GDP in 1993 to virtually zero in 1997. Monetary policy gave high priority to locking in over the medium term the low inflation already achieved, and as the recovery proceeded the policy rate went up from 4.75% in 1993 to 7.5% in 1995.

4.2.5 The Impact of the Reforms

Income per capita grew by 15.6% over the five-year post-reform period (1994–1999), compared to just 5.5% in the five years prior to the reforms (1988–1993). Advanced economies experienced a more modest pick-up from 8.4 to 12.8% between these two periods. The stronger growth pick-up in Australia tentatively suggests that the mid-1990s reforms had a positive impact.

Likewise, while prior to the reforms, the GDP per capita path of Australia was close to that of the synthetic Australia, it became moderately steeper in the post-reform (1995–1999) period. Indeed, our analysis finds that GDP per capita levels was 2% higher in Australia than in the synthetic control group five years after the reform, while levels were similar in 1994. At the same time, the placebo experiments are not conclusive, consistent with the small size of the growth acceleration. Half of the 12 placebo experiments in the potential controls yield a treatment effect below that of Australia. However, given the small number of placebo countries, it is difficult to draw strong conclusions regarding the statistical significance of the reform impact (Fig. 2).

Overall, the results tentatively suggest that Australia's growth might have benefited from the product and labor market reforms during the second half of the 1990s. Furthermore, Australia's performance was strong throughout the 2000s, which could have partly reflected delayed effects of the 1990s wave of reforms. This conclusion is
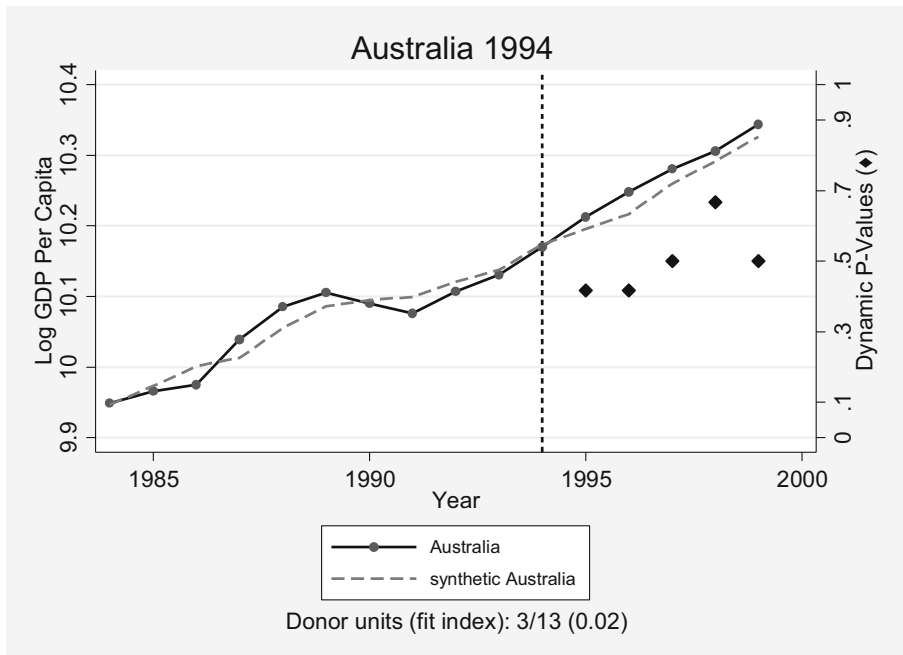
Fig. 2 GDP per capita in Australia and the synthetic unit over time. Note: Synthetic Australia consists of Greece (29%) the United States (66%), and Belgium (5%)

broadly consistent with that of Parham (2004), who argues that Australia's reforms played a key part in the productivity surge of the 1990s. IMF (2015) finds positive effects of the 1998–2001 wave of fiscal reforms using the synthetic control method. Others such as Quiggin (2001, 2004) have been more skeptical.

4.3 Labor and Product Market Reforms in the Netherlands in the 1990s

4.3.1 Rising Unemployment Triggered the Reforms

The Netherlands experienced an economic downturn during the early 1990s, as a result of which unemployment rate rose steeply from about 6% in 1991 to 10% in 1994. Against this background, the government took major reform measures to enhance wage flexibility and boost job search and creation incentives. The government also took measures to boost the competition in the domestic market by reducing state control of the economy and opening up major industries to competition, such as gas, electricity and telecommunications.

4.3.2 Deregulation of the Labor Market

- *Minimum wage.* Starting from 1994 administrative extensions of labor agreements no longer applied to the lower end of the market. In October 1995, representatives of employers and workers' organizations agreed to reduce the gap between the legal minimum wage and minimum wages set in collective labor agreements. Another important policy step toward increasing the flexibility of wage agreements was through

the promotion of "opening clauses" that allowed firms, under certain circumstances to negotiate with their workers pay levels below the minima set in collective contracts.

- *Active labor market policies and benefit conditionality.* Active labor market policies were scaled up, and their design was enhanced on several grounds, including by strengthening benefit conditionality. The TBA Act (Restriction of Claims on the Disability Benefit Regulations) of August 1993 reduced the generosity of the disability scheme and introduced more stringent conditions for initial and continued access to the disability scheme. In January 1995 the government introduced three programs to increase the creation of temporary jobs for the long-term unemployed, and in June the Minister of Social Affairs approved 57 subsidized job programs in the private sector, also targeted at this category of unemployed. To reduce sickness-related absence from work, the government abolished the public sickness benefit scheme in March 1996. In August, a new law banned unemployment benefits in cases where unemployment was deemed to be voluntary. Refusal to accept a suitable job offer would result in a complete and permanent withdrawal of an unemployment benefit. In early 1997, the government introduced a law that included the obligation of Public Employment Services to inform the unemployment benefit agencies of insufficient job search efforts by benefit claimants.

- *Tax wedges.* The government cut labor tax wedges on low-wage workers. In January 1996, the authorities introduced a special measure to reduce employers' social contributions for workers with wages that were less than 115% of the legal minimum wage. The costs of hiring a long-term unemployed person at wages up to 130% of the legal minimum were reduced by about 13%.

- *Employment protection legislation.* In December 1995, the government decided to shorten dismissal procedures. According to the new policy, an employer could dismiss his employee at the same time or even before seeking permission from the director of the Public Employment Service. A new law on fixed-term contracts, probation periods and dismissal procedures came into effect in early 1998. The law removed restrictions on the renewal of fixed-term contracts and slightly eased dismissal procedures for permanent workers. In January 1999, the Flexibility and Security Law came into force. It promoted the use of flexible working contracts and also streamlined somewhat dismissal procedures.

4.3.3 Product Market Liberalization

Starting from the mid-1990s, the Dutch government implemented a series of measures that drastically strengthened competition including by reducing barriers to entry in various product markets.

- *Barriers to entry.* Air transport was deregulated with the implementation of the corresponding EU Directive in 1993. In postal services, the government sold a 30% stake in KPN to the public in 1994, and another 22% stake in 1995, reducing the government's share to 48%. The railway reform initiated in 1995 made competition possible and several new operators appeared starting from the following year. In retail trade, the Shop Hours Act of 1996 led to a major liberalization of existing rules and regulations regarding shop opening hours. In 1997, two new national telephone operators were allowed to enter the market, and in mobile telephony the government

announced plans to auction additional frequencies after those it had already auctioned in late 1995. Finally, in the late 1990s the government implemented a wave of reforms to open up including electricity, telecommunications, pharmaceuticals and gas. The Postal Directive of the European Commission was implemented in 1998 to open some segments of mail services, while the Dutch electricity market was opened to competition in 1999—licenses were no longer required for the production of electricity and large users were allowed to choose among suppliers. Likewise, the 1999 Gas Act started a gradual liberalization process in gas distribution.

- *Competition policy.* Three decrees based on the existing 1956 Economic Competition Act were introduced, each of which dealt with a specific group of competition-restricting agreements—for instance the first decree prohibited horizontal price agreements from July 1993, with effective implementation over 1994 especially when the government decided to refuse 12 requests for an exception to the general ban on such agreements. In addition, in September 1993 amendments of the Economic Competition Act were submitted to Parliament, the most important of which concerned improved possibilities of legal action against informal restrictive practices. This effectively broadened the scope of the law to cover professional services.
- *Administrative burdens on firms.* In January 1996, the Liberalized Establishment Law came into force, and decreased the number of establishment licenses from 88 to 8. The aim was to reduce barriers to entry and lower administrative costs for startups.

4.3.4 Macroeconomic Context

The reforms were introduced against the background of high unemployment in the aftermath of the economic recession of the early 1990s. The macroeconomic policy mix turned accommodative around the start of the reform wave in 1994. Strong confidence in the Dutch guilder throughout the 1993 European Exchange Rate Mechanism crisis allowed the Netherlands to keep the lowest interest rates in the European Union, and to cut policy rates by over 400 basis points between 1992 and 1994 that continued afterward. The corresponding easing of monetary conditions was substantial, although it was somewhat dampened by the appreciation of the real effective exchange rate. On the fiscal policy front, the drastic consolidation of the early 1990s came de facto to a halt in 1994–1995, with a small weakening of the cyclically-adjusted balance.

4.3.5 What Was the Impact of the Reforms on Income and Growth?

The Dutch economy recovered more strongly than the average European or advanced economy from the downturn. Real GDP per capita grew by 18.3% over 1994–1999, versus 10.8% during the 1988–1993 (pre-reform) period. The pick-up in average real GDP per capita across advanced economies was comparatively milder—from 8.4% over 1988–1993 to 12.8% during 1994–1999.

Real GDP per capita in the Netherlands was very close to that of its synthetic control group in 1994. After the start of structural reform episode, that the estimated impact of the reforms on real GDP per capita increased from 4% in 1995 to 6 percent in 1996. The placebo test finds that less than one-fifth of the control countries have a larger estimated treatment effect than that of the Netherlands—although, again, it is difficult to draw

strong conclusions from such tests given the small number of units. Overall, there is tentative evidence that the major labor and product market reforms carried out in the Netherlands starting from the mid-1990s had some positive impact on GDP per capita growth in the five years after the start of the episode (Fig. 3).

4.4 Labor and Product Market Reforms in Denmark in the 1990s

4.4.1 Slow Growth and Rising Unemployment Motivated the Reforms

Denmark experienced sluggish economic growth and rising unemployment in the early 1990s. In response, the government implemented a reform package that involved an overhaul of the unemployment benefit system and active labor market policies—the build-up of Denmark's "flexi-security"—and a sizeable decline in labor tax wedges, as well as product market deregulation.

4.4.2 Labor Market Reforms

• *Active labor market policy.* In June 1993, the parliament approved a labor market reform effective from 1994. A key element of the reform was the strengthening of activate labor market policies, with particular focus on the long-term unemployed. Participation in training or education and taking up suitable job offers became mandatory after four years of unemployment. In 1996, additional measures were implemented that further strengthened activation. Compulsory full-time
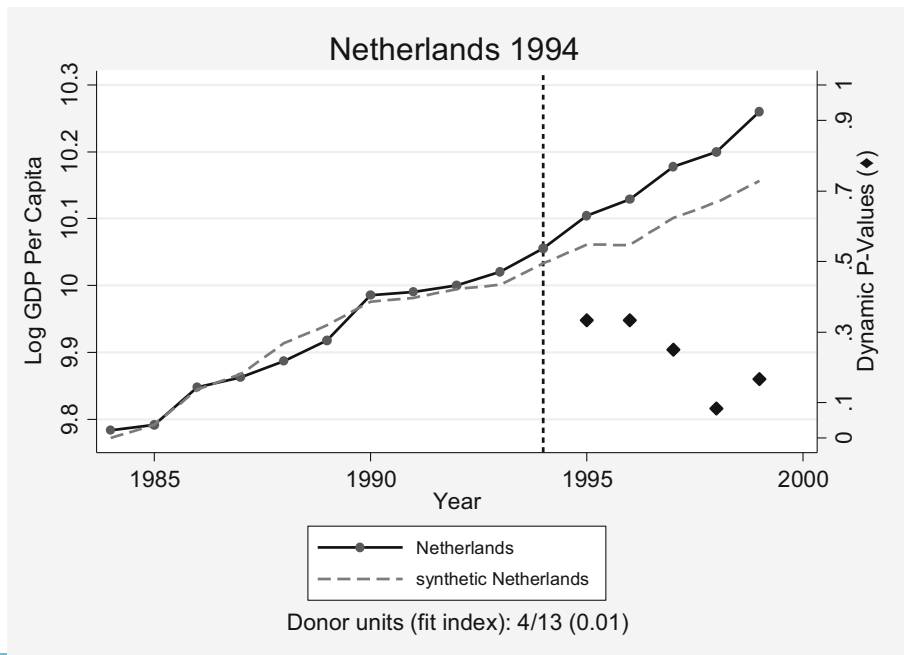
Fig. 3 GDP per capita in Netherlands and the synthetic unit over time. Note: Synthetic Netherlands consists of Belgium (28%), Greece (33%), Switzerland (33%), and the United States (6%)

participation in activation measures was brought forward by two years to take effect after two years of unemployment. Young people below 25 years of age who were on unemployment benefits had to take 18 months training with benefits at 50% of the normal level after six months of unemployment. The new policy entered into force in April 1996 and was fully phased in by April 1997.

- *Unemployment benefit duration and eligibility.* In 1994–95 the maximum benefit period was limited to seven years, and could be extended by two years through participation in paid-leave schemes. From 1996, the maximum effective benefit period was reduced to five years of which the last three involved full-time activation. Furthermore, eligibility criteria for unemployment benefits were tightened. Access to unemployment benefit required 52 weeks of non-subsidized work (formerly 26 weeks) over the last three years. Paid leave for education could no longer extend the benefit period.
- *Tax Reforms.* Labor tax wedges were cut. The government announced in May 1993 that marginal tax rates on personal incomes for all income groups would be gradually reduced by 8 to 14 percentage points over the 1994–1998 period. However, in order to partially maintain revenue-neutrality, payroll taxes were introduced to finance labor-market policy measures. The tax base was also broadened by eliminating special privileged tax arrangements for some income types. Overall, the combined marginal rate of personal income tax and social security contributions was lowered by 6 to 9 percentage points for most employees.

4.4.3 Product Market Deregulation

- *Public ownership.* State enterprises were transformed into limited liability companies, in industries such as air transportation, postal services, telecommunications, and insurance. Statsanstalten for Livsforsikring (State Life Insurance Company) was sold to a private company. In 1993 the government sold shares in state companies, including 25% shares in Copenhagen Airport and the Postal Giro and 49% of the shares in TeleDanmark, the Telecom holding company. The privatization of Danish Telecom was completed in 1997.
- *Retail trade regulation.* The government introduced new regulations on shop opening hours in March 1994, clarifying existing rules and extending standard opening hours. Opening hours became unregulated from 8 a.m. Monday to 5 p.m. Saturday, and shops with a particularly low turnover were allowed to stay open on Sundays.
- *Barriers to entry in network industries.* Along with the dissolution of Danair, the domestic airline industry was deregulated in 1995. Some liberalization of the telecommunications sector took place in 1996. The parliament adopted a framework for improving competition in the rail transport sector in 1997, when the operation of railway transport was separated from track maintenance. The 1999 Energy Supply Act introduced competition in electricity.

4.4.4 Macroeconomic Context

The reforms were initiated as economic growth was starting to recover from the recession of the early 1990s. Monetary policy was accommodative—after a short

stress period during the 1993 crisis of the European Exchange Rate Mechanism, policy rates were reduced drastically alongside Germany's over 1993–1994. On the fiscal front, the new government that took office in 1993 put forward a medium-term strategy to improve public finances through reduced public expenditure, with an operational target of eliminating the general government deficit by 1997. In the short term, however, a fiscal stimulus of about 1.5 percentage points of GDP was implemented against the backdrop of high unemployment, with consolidation to take place later on when the economy would recover. Indeed, the cyclically-adjusted general government deficit widened in 1995 before gradually shrinking in the following years.

4.4.5 The Impact of the Reforms

Real GDP per capita grew by 12.7% in the five years after the start of the reform episode (1994–1999), compared to just 4.5% in the pre-reform period (1988–1993). GDP per capita growth also picked up in advanced economies, from 8.4% over 1988–1993 to 12.8% over 1994–1999.

Analysis using the synthetic control method suggests an insignificant impact of the reforms on GDP per capita in the first few years after the start of the reform episode. The GDP per capita gap between Denmark and its synthetic equivalent remained roughly constant over the 10 years around the treatment date (1994).
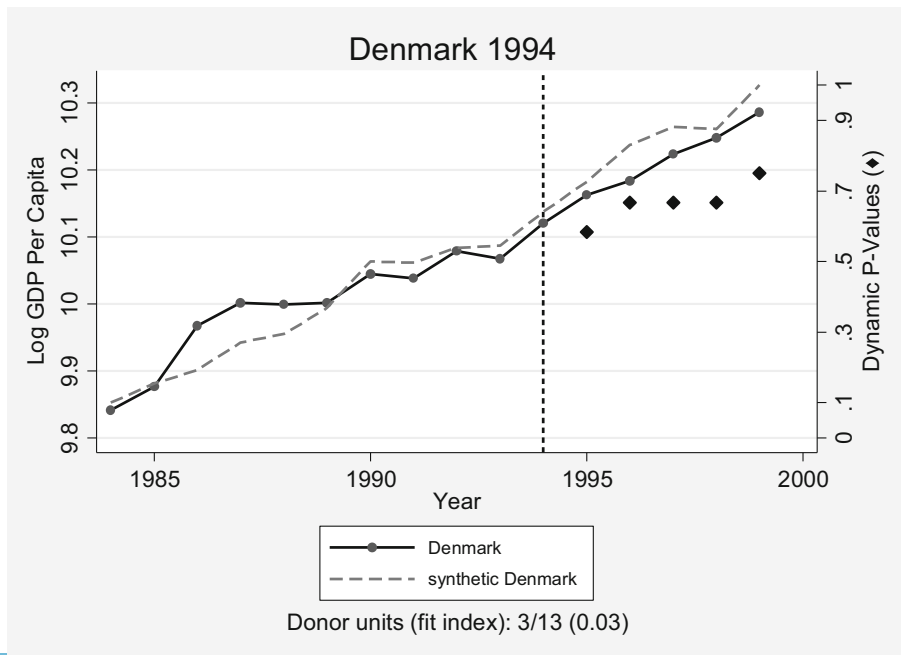
Fig. 4 GDP per capita in Denmark and the synthetic unit over time. Note: Synthetic Denmark consists of Belgium (17%), Italy (40%), and Norway (43%)

Accordingly, the placebo experiments do not point to a significant impact of reform (Fig. 4).

4.5 Labor and Product Market Reforms in Ireland in the 1990s

4.5.1 Poor Performance of the Labor Market Induced the Reforms

In the early 1990s, Ireland witnessed the most rapid growth among all EU countries.

Despite the recession in continental Europe, Ireland's growth remained robust, driven by both rising exports and recovering domestic demand. Nevertheless, the persistently high long-term and youth unemployment rates became a major concern for Irish policy makers, who responded with a package of labor market reforms (OECD 1995). The authorities also pushed through product market reforms in order to strengthen competition.

4.5.2 Labor Market Reforms

- *Unemployment benefits.* In the early to mid-1990s, about half of the total unemployed had been without work for more than a year. In 1994, the government introduced important changes to the benefit system that lowered replacement rates for the long-term unemployed, with the view to strengthening monetary incentives to take up job offers. The average replacement rate was reduced from 42 to 35% for a single person and from 70 to 64% for a married person with two children. Furthermore, the Back to Work Allowance Scheme allowed the long-term unemployed to accept jobs while partly retaining their social welfare benefits. The 1995 Budget further reduced replacement ratios for the unemployed, and also narrowed the gap between child benefits paid to the unemployed and those paid to the employees.
- *Active labor market policies.* Starting from 1994 the government scaled up its active labor market policies. Compared with the earlier programs it replaced, the new Community Employment Programme effectively doubled the number of long-term unemployed participants in activation activities—including part-time and public-sector jobs as well as training. Participants in this scheme accounted for about 3% of total employment. The 1997 the government further strengthened active labor market policies, including through job subsidies and greater scope for the unemployed to combine wage income with unemployment benefit receipt. Fourteen local employment services offices were established to help integrate the long-term unemployed back into the labor market.
- *Collective bargaining and labor taxation.* Wage moderation, which had started with a national pay agreement in 1987, was pursued further through two centralized wage agreements—the Programme for Competitiveness and Work over 1993–1996 and the Partnership 2000 for Inclusion, Employment and Competitiveness over 1997–1999. These agreements were facilitated by a sharp accompanying cut in labor income taxation that boosted workers' purchasing power.

4.5.3 Reforms of the Product Market

- *Barriers to entry.* The 1993 EU Air Transport Liberalization Directive was swiftly implemented, contributing to significant liberalization from 1994. Barriers to entry in the gas sector were lowered, and in 1998 legislation was published to partially deregulate electricity supply. In 1998 the government also liberalized telecommunications, and an independent regulator was set up.
- *Competition policy.* In 1994 the government introduced amendments to the 1991 Act—which had established the Competition Authority—to improve the enforcement of competition policy. The amended law would allow the Authority to initiate investigations and to take court actions either on its own initiative or as a result of third party complaints. Enforcement was also strengthened by providing additional resources to the Competition Authority in 1995.
- *Public ownership.* The privatization and restructuring of state-owned enterprises also strengthened competition and efficiency. The government privatized Irish Life Plc. in 1995, and Irish Steel Ltd. In 1996. At the end of 1996, the government sold 20% of its share in Telecom Eireann. A number of loss-making state-owned enterprises were also restructured, including e.g. the Bord na Mona (the Irish Peat Board) and Aer Lingus.

4.5.4 Macroeconomic Context

Unlike the other case studies in this paper, reform in Ireland was carried out under robust economic growth. Macroeconomic policy was supportive at the start of the reform episode. Following the 10% devaluation of the Irish pound within the European Monetary System, exchange rate pressures were alleviated, and the central bank sharply lowered policy rates in 1993–1994 as these also fell in "core" countries. Monetary policy was subsequently tightened in 1995–1996 amid rapid credit growth and rising asset prices. Fiscal policy was also fairly supportive. In 1994, the government lowered the personal income tax burden on low-income earners, increased social welfare benefits and expenditure on new community employment schemes. Overall, despite strong revenue growth, the cyclically-adjusted balance weakened over 1994–1996.

4.5.5 Impact of Reforms

There is strong evidence from the synthetic control method that reforms paid off. Real GDP per capita grew by about 21% over the five years (1988–1993) prior to the reforms, whereas it rose by more than 54% during 1994–1999. This acceleration was much stronger than observed in other advanced economies. Our analysis using the synthetic control method finds that real GDP per capita in Ireland was 6% higher than in the synthetic unit in 1994, and 32% higher in 1999.[8] The placebo experiment

---

[8] We acknowledge that in the case of Ireland, the synthetic control method may not be able to disentangle the effects of other developments in the 1990s that may have contributed to the success of the reforms, such as a favorable external environment and rising confidence of international investors associated with financial globalization and its EU membership.

confirms a significant effect; the estimated impact of reforms is estimated to be larger in Ireland than in all of the twelve placebo units (Fig. 5).

At the same time, as pointed out for instance by Walsh (2003), several other exogenous factors, including a boom in FDI inflows, particularly from the United States, owing to the low corporation tax regime applied to foreign manufacturing firms, might have played a role in supporting growth. Ireland might also have benefitted with a lag from the important labor market reforms it had carried out in the late 1980s (see e.g. Tille and Yi, 2001). Together these other factors might explain why Ireland had already started to diverge from its synthetic control around 1992–1993.

4.6 Labor Market Reforms in Germany in the 2000s

4.6.1 Weak Economic Growth and High Unemployment Triggered the Reforms

In response to relatively low growth and rising and persistent unemployment in the early 2000s, the German government implemented a wide range of labor market reforms, the so-called Hartz reforms, which aimed at increasing job matching efficiency and raising work incentives. The Hartz reforms were also preceded, and accompanied by an increase in the decentralization of the bargaining process that set wages, hours and working conditions, from the industry- and region- levels to firm or individual levels.
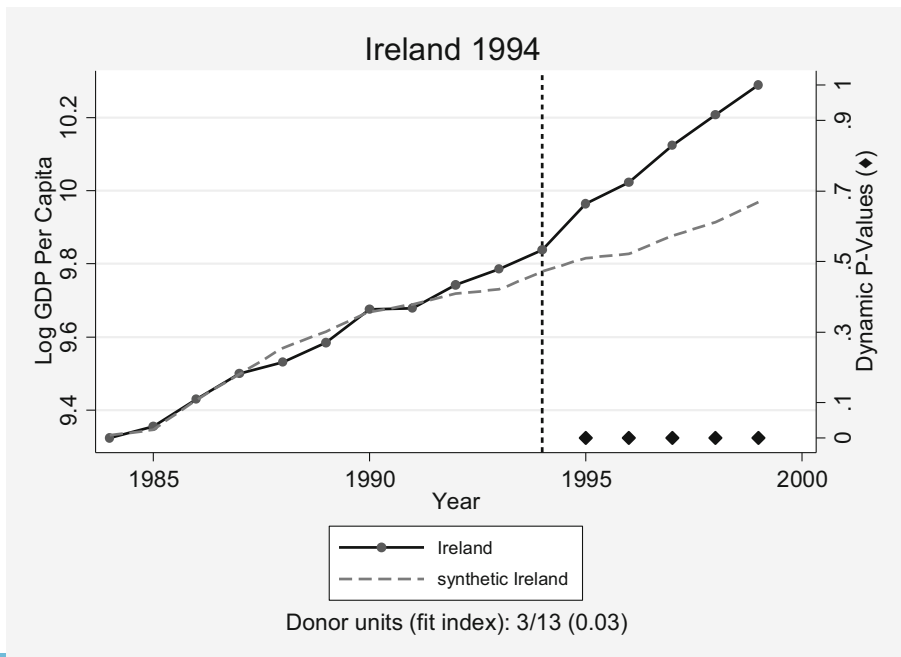
Fig. 5 GDP per capita in Ireland and the synthetic unit over time. Note: Synthetic Ireland consists of Belgium (23%), Portugal (64%), and the United States (13%)

4.6.2 The Hartz Reforms

- *Overview.* The first three parts of the reform package, Hartz I-III, were implemented in 2003 and 2004, while the fourth part was implemented in 2005. Hartz I-III mainly focused on creating new types of employment opportunities (Hartz I), introducing additional wage subsidies (Hartz II), and restructuring the Federal Employment Agency (Hartz III). The specific reform measures were based on a three-part reform strategy: improving the efficiency of employment services and policy measures, activating the unemployed and fostering employment demand (Jacobi and Kluve 2006). The reforms modified the already existing measures of Active Labor Market Policy and implemented new measures to facilitate the expansion of the temporary work sector.
- *Unemployment benefits.* The fourth stage of the reform (Hartz IV) resulted in a significant cut in the unemployment benefits for the long-term unemployed, as these were merged with, and thereby reduced to the lower level of welfare benefits. The reform package also included a significant shortening of the eligibility period for unemployment insurance benefits for older employees and a tightening of job search requirements.

4.6.3 Macroeconomic Context

Growth was weak and unemployment was on the rise when reform implementation began. The macroeconomic policy stance was mixed, with accommodative monetary policy but rather restrictive fiscal policy. The European Central Bank cut its main refinancing rate by 125 basis points over 2002–2003, and left it at 2% during the period 2003–2005. Nonetheless, due to comparatively lower inflation, real short-term interest rates were higher in Germany compared with other countries in the euro area. Meanwhile, despite high unemployment, the government gave priority to gradual fiscal consolidation, which was to be achieved in part through public sector expenditure reforms.

4.6.4 The Impact of the Reforms

Considering 2003 as the starting date for the reform episode, real GDP per capita in Germany grew by about 11% in the five-year period (2003–2008) after the reforms began, compared to 8.3% during the five years prior to the reforms (1997–2002). In contrast, the growth of average real GDP per capita in all advanced economies slowed down from 10.6% in the pre-reform period to 8.7% in the post-reform period. Our analysis using the synthetic control method also points to an impact of reform, with the real GDP per capita gap between Germany and its synthetic counterpart rising from 2% in 2003 to 6% in 2008. The placebo experiment is somewhat less conclusive, with a third of the placebo units showing a larger estimated impact of reform than Germany (Fig. 6).

There is only limited consensus in the literature regarding how big a role the Hartz reforms played in lowering unemployment and raising GDP per capita in the 2000s. Using a calibrated model, Krebs and Scheffel (2013) find that the entire
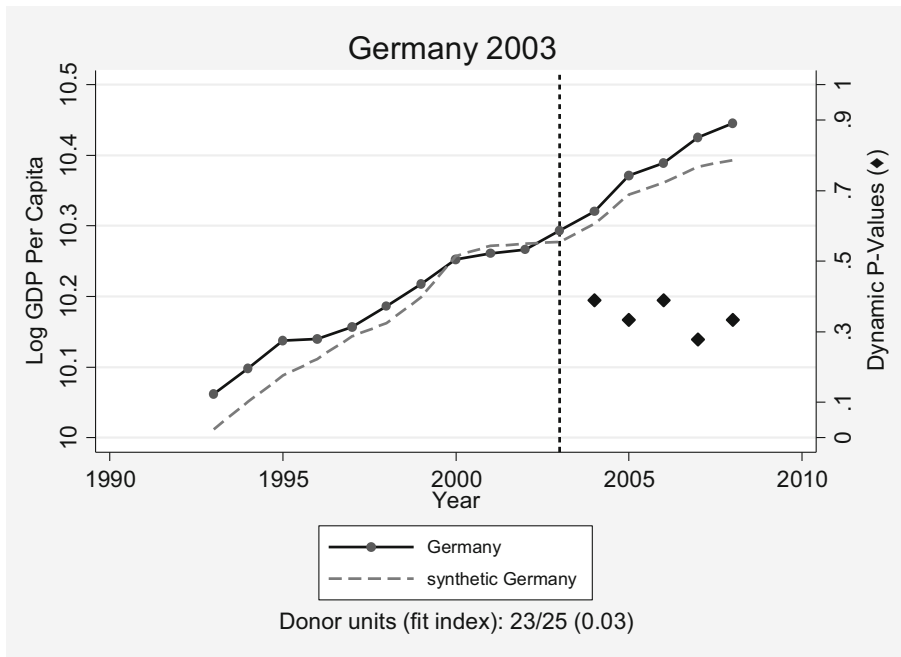
Fig. 6 GDP per capita in Germany and the synthetic unit over time. Note: Synthetic Germany consists of Australia, Austria, Belgium, Canada, Cyprus, Denmark, Finland, France, Greece, Ireland, Israel, Italy, Japan, Latvia, Lithuania, the Netherlands, New Zealand, Norway, Portugal, South Korea, Spain, Sweden, Switzerland, the United Kingdom and the United States

Hartz reform package led to a permanent reduction in the German unemployment rate of almost three percentage points. By contrast, Dustmann and others (2014) argue that while the Hartz reforms contributed to some reduction in long-term unemployment, specific features of German system of industrial relations, which facilitated a decentralization of wage determination from the industry to the firm level, played a dominant role in enhancing the competitiveness of German economy. Finally, the Hartz reforms were implemented at a time when strong global trade growth and Germany's specialization toward highly-demanded goods also helped the recovery, possibly more than it did in the synthetic control group.

5 Concluding Remarks

Our paper contributes to the large quantitative and qualitative literature on the impact of structural reforms through individual case studies of major past reform episodes in advanced economies cutting across labor and product market areas. Compared with traditional case studies, the main advantage of our approach is to carefully construct a counterfactual using the synthetic control method, taking as a treatment date the starting year of the reform wave considered. Once data and methodological constraints are taken into account, we are left to analyze six individual cases, namely the product and labor market reform packages in New Zealand, Australia, The Netherlands, Denmark,

and Ireland in the 1990s, and the labor market (so-called Hartz) reforms in Germany in the 2000s. In four out of these six cases, there is some evidence of a pick-up in the GDP per capita path relative to the counterfactual already in the five years after the start of the reform episode. Exceptions include Denmark—no estimated effect relative to the counterfactual, although GDP per capita growth did pick up—and New Zealand— where the pick-up was initially smaller than under the no-reform counterfactual, although some catch-up took place later on, which could reflect that New Zealand was the only country amongst the six cases reviewed here where macroeconomic policy was clearly restrictive around the start of the reform period. At the same time, with the exception of Ireland, the placebo experiments do not enable us to claim statistical significance of the results. Taken together, these results point to the following two tentative policy conclusions. First, major reform packages are typically implemented gradually over several years, and also often take time to pay off, making it difficult to identify systematically large effects in their immediate aftermath. Second, supportive macroeconomic policies seem to enhance the short-term pay-off from reforms—New Zealand, where such support was lacking when major reforms were implemented in the early 1990s, was a case in point.

Appendix: Figures for placebo experiments
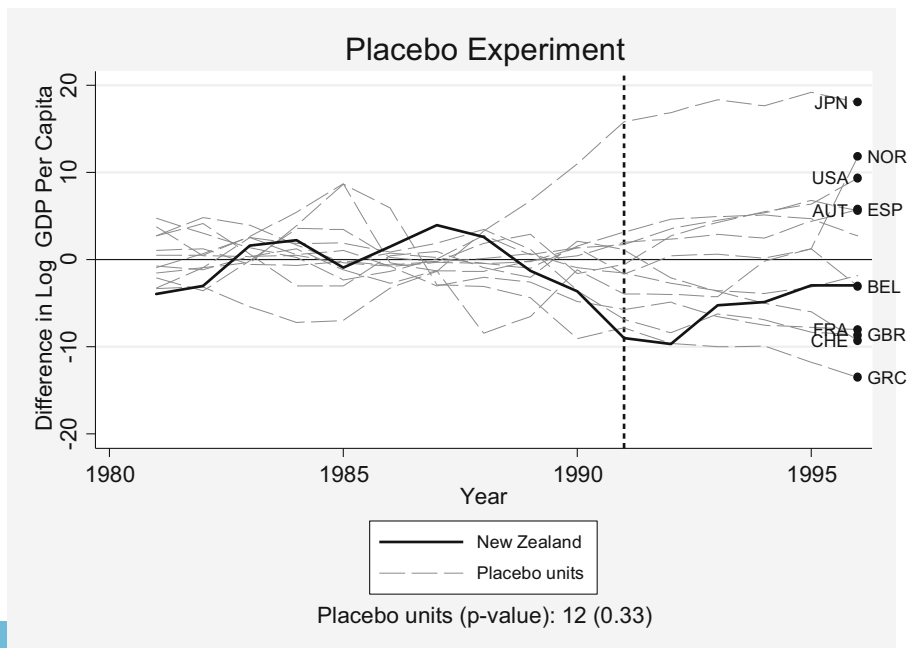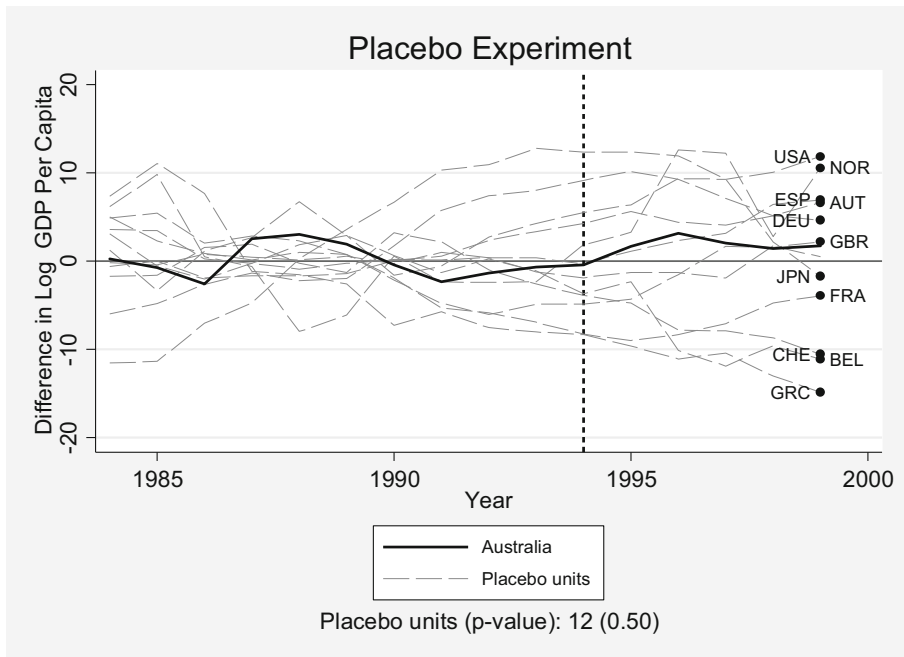

Fig. 7 New Zealand (1991)
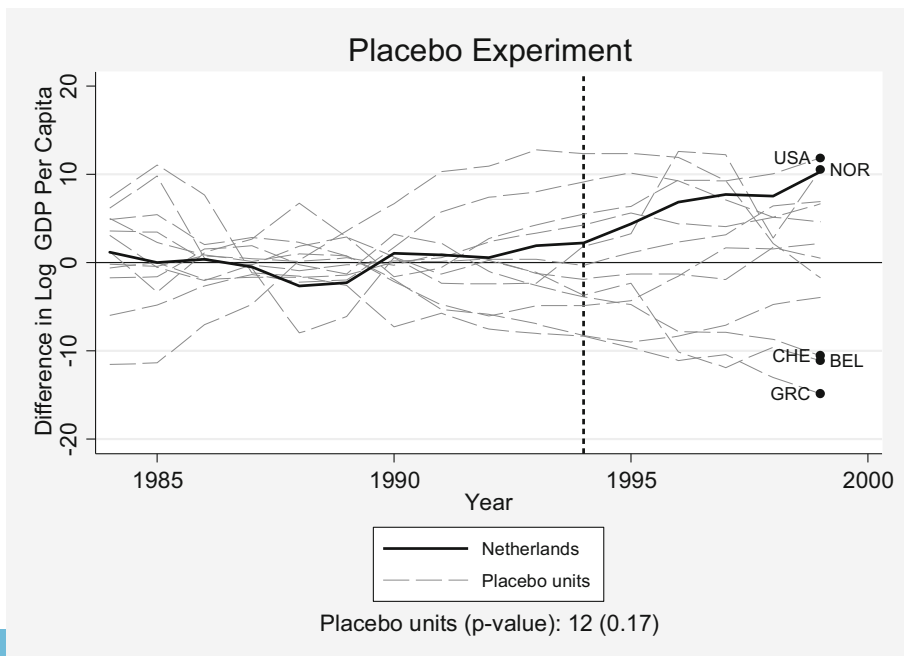
Springer

Fig. 8 Australia (1994)

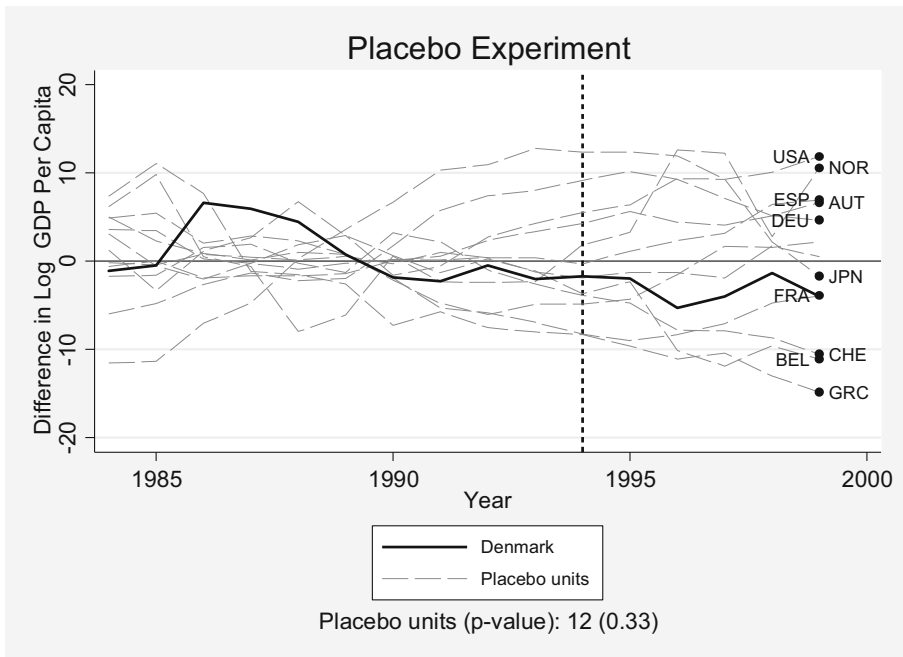**Fig. 9** Netherlands (1994)

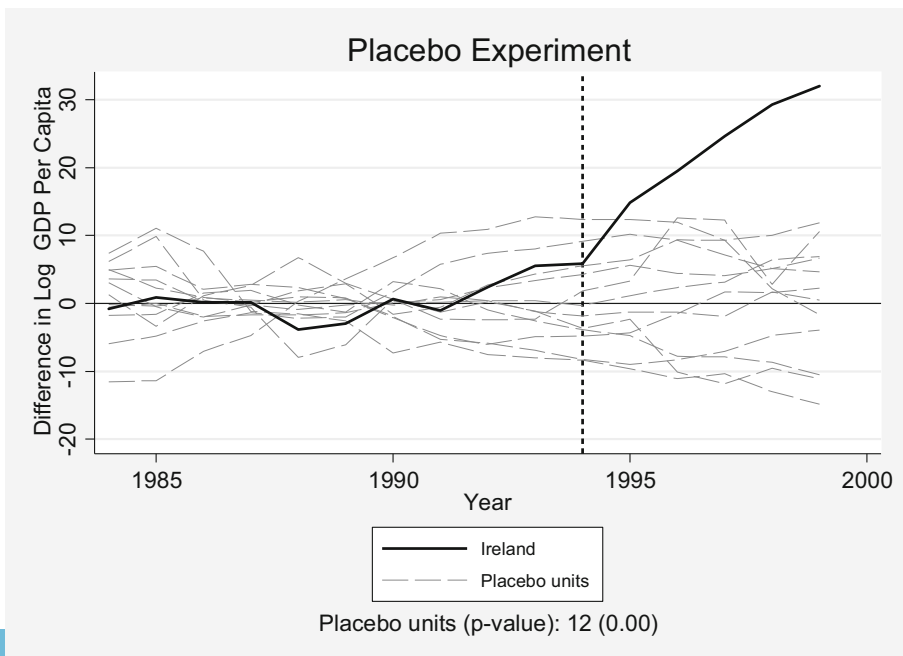**Fig. 10** Denmark (1994)

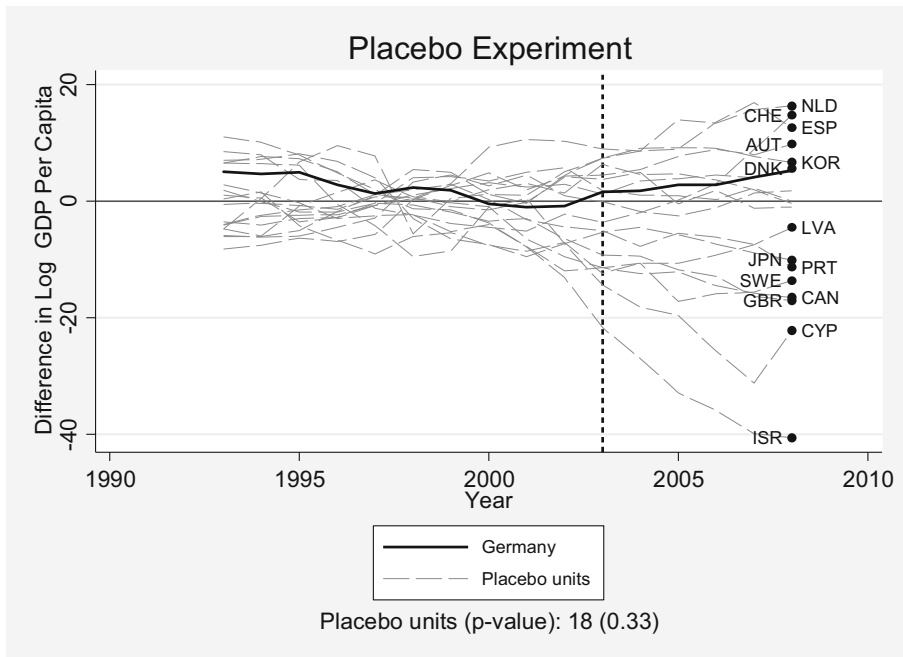**Fig. 11** Ireland (1994)

Springer

Fig. 12 Germany (2003)

References

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

Abadie A, Diamond A, Hainmueller J (2015) Comparative politics and the synthetic control method. Am J Polit Sci 59(2):495–510

Abadie A, Gardeazabal J (2003) The economic costs of conflict: a case study of the Basque Country. Am Econ Rev 93(1):113–132

Abdallah CS, Lastrapes WD (2012) Home equity lending and retail spending: evidence from a natural experiment in Texas. Am Econ J Macroecon 4(4):94–125

Adhikari B, Alm J (2016) Evaluating the economic effects of flat tax reforms using synthetic control methods. Southerm Economic Journal 83(2):437–463

Athey S, Imbens GW (2017) The state of applied econometrics: causality and policy evaluation. J Econ Perspect 31(2):3–32

Bauhoff S (2014) The effect of School District nutrition policies on dietary intake and overweight: a synthetic control approach. Economics & Human Biology 12:45–55

Bertrand M, Duflo E, Mullainathan S (2004) How much should we trust differences-in-differences estimates? Q J Econ 119(1):249–275

Bhagwati J, Srinivasan TN (2002) Trade and poverty in the poor countries. Am Econ Rev 92(2):180–183

Billmeier A, Nannicini T (2011) Economies in transition: How important is trade openness for growth? Oxf Bull Econ Stat 73(3):287–314

Billmeier A, Nannicini T (2013) Assessing economic liberalization episodes: a synthetic control approach. Rev Econ Stat 95(3):983–1001

Bohn S, Lofstrom M, Raphael S (2014) Did the 2007 legal Arizona workers act reduce the State's unauthorized immigrant population? Review of Economics and Statistics 96(2):258–269

Boulhol, Herve, Alain De Serres and Margit Molnar. 2008. "The contribution of economic geography to GDP per capita." OECD Economic Journal: Economic Studies Volume 2008

Briggs C, Buchanan J (2000) Labor market deregulation: a critical assessment. In: Research paper no 21. Information and Research Services, Department of Parliamentary Library, Canberra, ACT, Australia, pp 1999–2000

Smith B (2015) The resource curse exorcised: evidence from a panel of countries. J Dev Econ 116:57–73

Campos N, Kinoshita Y (2010) Structural reforms, financial liberalization, and foreign direct investment. IMF Staff Pap 57(2):326–365

Cavallo E, Galiani S, Noy I, Pantano J (2013) Catastrophic natural disasters and economic growth. Rev Econ Stat 95(5):1549–1561

DeAngelo G, Hansen B (2014) Life and death in the fast lane: police enforcement and traffic fatalities. Am Econ J Econ Pol 6(2):231–257

Dorsett R (2013) The effect of the troubles on GDP in Northern Ireland. Eur J Polit Econ 29:119–133

Duval R, Furceri D, Hu B, Jalles J, Nguyen H (2018) A narrative database of major labor and product market reforms in advanced economies. IMF working paper No. 18/19. International Monetary Fund, Washington D.C.

Dustmann C, Fitzenberger B, Schonberg U, Spitz-Oener A (2014) From sick man of Europe to economic superstar: Germany's resurgent economy. J Econ Perspect 28(1):167–188

Evans L, Grimes A, Wilkinson B, Teece D (1996) Economic reform in New Zealand 1984-95: the pursuit of efficiency. J Econ Lit 34(4):1856–1902 http://www.jstor.org/stable/2729596

Galt, David. 2000. "New Zealand's economic growth". New Zealand Treasury Working Paper No. 00/09

Glyn, Andrew. 2005. Labour market success and labour market reform: lessons from Ireland and New Zealand. In *Fighting Unemployment: The Limits of Free Market Orthodoxy*: 197, edited by David Howell, Oxford University Press, New York

International Monetary Fund (IMF) (2016) World economic outlook: April 2016. Washington, D.C.

International Monetary Fund (IMF) (2015) Fiscal policies and long-term growth. Washington, D.C.

Jacobi L, Kluve J (2006) Before and after the Hartz reforms: the performance of active labour market policy in Germany. In: IZA discussion paper no. 2100; RWI discussion paper no. 41. Available at SSRN http://ssrn.com/abstract=900374

Krebs T, Scheffel M (2013) Macroeconomic evaluation of labor market reform in Germany. IMF Economic Review 61(4):664–701

McCann P (2003) Geography, trade and growth: problems and possibilities for the New Zealand economy. New Zealand Treasury Working Paper 03(/03)

Montalvo JG (2011) Voting after the bombings: a natural experiment on the effect of terrorist attacks on democratic elections. Review of Economics and Statistics 93(4):1146–1154

Organization for Economic Co-operation and Development (OECD) (1991) OECD economic surveys: New Zealand 1990/1991. OECD Publishing, Paris

OECD (1995) OECD economic surveys: Ireland 1995. OECD Publishing, Paris

Parham, Dean. 2004. Australia's 1990s productivity surge and its determinants. In *Growth and Productivity in East Asia,* edited by Takatoshi Ito and Andrew rose. Chicago: University of Chicago Press

Persson T, Roland G, Tabellini G (2007) Electoral rules and government spending in parliamentary democracies. Quarterly Journal of Political Science 2(2):155–188

Quiggin J (2001) The Australian productivity miracle: a Sceptical view. Agenda 8(4):333–348

Quiggin J (2004) Looking back on microeconomic reform: a skeptical viewpoint. In: Australian public policy program working paper 1/P03, vol 15. The University of, Queensland, pp 1–25

Rodriguez, Francisco, and Dani Rodrik. 2001. "Trade Policy and Economic Growth: a Skeptic's Guide to the Cross-national Evidence." In *NBER Macroeconomics Annual* 2000, Volume 15, pp. 261–338. MIT Press

Tille, Cedric, and Kei-Mu Yi, 2001. Curbing unemployment in Europe: are there lessons from Ireland and the Netherlands? *Current Issues in Economics and Finance* 7(5), Federal Reserve Bank of New-York

Walsh, Brendan, 2003, "When Unemployment Disappears: Ireland in the 1990s". CESifo Working Paper Series No. 856. Available at SSRN: http://ssrn.com/abstract=382944