ECONSTOR
Make Your Publications Visible.

A Service of

ZBW Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Francis, Joseph

Working Paper

A Replication of "Comparative Politics and the Synthetic Control Method" by Abadie et al. (2015)

I4R Discussion Paper Series, No. 271

Provided in Cooperation with:
The Institute for Replication (I4R)

Suggested Citation: Francis, Joseph (2025) : A Replication of "Comparative Politics and the Synthetic Control Method" by Abadie et al. (2015), I4R Discussion Paper Series, No. 271, Institute for Replication (I4R), s.l.

This Version is available at:
https://hdl.handle.net/10419/330835

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.

WWW.ECONSTOR.EU

Mitglied der Leibniz-Gemeinschaft

INSTITUTE for REPLICATION

No. 271
DISCUSSION PAPER SERIES

A Replication of "Comparative Politics and the Synthetic Control Method" by Abadie et al. (2015)

Joseph Francis

This paper received a response:

Abadie, A., A. Diamond, J. Hainmueller. 2025. Response to Francis (2025). *I4R Discussion Paper Series* No. 272. Institute for Replication.

November 2025

INSTITUTE for REPLICATION

I4R DISCUSSION PAPER SERIES

I4R DP No. 271

A Replication of "Comparative Politics and the Synthetic Control Method" by Abadie et al. (2015)

Joseph Francis

NOVEMBER 2025

Any opinions in this paper are those of the author(s) and not those of the Institute for Replication (I4R). Research published in this series may include views on policy, but I4R takes no institutional policy positions.

I4R Discussion Papers are research papers of the Institute for Replication which are widely circulated to promote replications and meta-scientific work in the social sciences. Provided in cooperation with EconStor, a service of the ZBW – Leibniz Information Centre for Economics, and RWI – Leibniz Institute for Economic Research, I4R Discussion Papers are among others listed in RePEc (see IDEAS, EconPapers). Complete list of all I4R DPs - downloadable for free at the I4R website.

I4R Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

Editors

Abel Brodeur
University of Ottawa

Anna Dreber
Stockholm School of Economics

Jörg Ankel-Peters
RWI – Leibniz Institute for Economic Research

E-Mail: joerg.peters@rwi-essen.de
RWI – Leibniz Institute for Economic Research
Hohenzollernstraße 1-3
45128 Essen/Germany
www.i4replication.org

ISSN: 2752-1931

A Replication of "Comparative Politics and the Synthetic Control Method" by Abadie *et al.* (2015)

Joseph Francis

September 28, 2025
(v4.0)

Abstract

Abadie, Diamond, and Hainmueller's (2015) analysis of the costs of German reunification is a flagship study for the "synthetic control method" (SCM). Yet three issues can be identified. First, Abadie *et al.* remove countries that performed poorly in the 1990s from their donor pool, potentially biasing their results to amplify the negative impact of reunification on West Germany. Second, Abadie *et al.* use nominal GDP per capita to measure growth but report using real GDP per capita in 2002 USD. When the reported series is used, their results become more fragile: they remain statistically significant at $p < 0.10$, but the results of other robustness tests that Abadie *et al.* include are weaker. Third, even those fragile results reflect an arbitrary choice of reference year, given the instability that results from the spatial Gerschenkron effect: when Abadie *et al.*'s study is replicated using real GDP per capita series that are re-referenced using every year from 1960 to 2003, only 8 out of 44 synthetic controls generate treatment effects with $p < 0.10$.

Keywords: econometrics, German reunification, Gerschenkron effect, growth, synthetic control method (SCM)

JEL codes: C12, C22, C52, C80

Alberto Abadie, Alexis Diamond, and Jens Hainmueller (2015) have sought to demonstrate how the Synthetic Control Method (SCM) can be applied to

---

Joseph Francis is an economic historian from Wales. He is an honorary research fellow at the University of Birmingham and can be contacted at joefrancis505@gmail.com. Discussions with Daniel B. Klein aided the writing of this paper. It has also benefited from Abel Brodeur's feedback, as well as a response to an earlier draft from Alberto Abadie, Alexis Diamond, and Jens Hainmueller. Supplementary materials are available at https://github.com/joefrancis505/Francis_Comparative_Politics.

comparative politics, using the effects of reunification on West Germany's GDP per capita as a case study. An algorithm builds a synthetic West Germany out of a pool of suitable comparison countries—the so-called "donors"—to estimate the treatment effect. This technique, Abadie *et al.* (2015, 495–496, 500) state, allows them to address "a widespread consensus [that] has emerged about the necessity of establishing bridges between quantitative and qualitative approaches to empirical research in political science." Crucially, the SCM "provides a systematic way to choose comparison units in comparative case studies." In doing so, it "opens the door to the possibility of precise quantitative inference." It allows, moreover, for "a quantitative comparison between the distribution of placebo effects and the synthetic control estimate [that] can be operationalized through the use of p-values." Abadie *et al.* thereby imply that they are making comparative politics more scientific.

Yet three issues are worthy of discussion. First, the exclusion of some countries from the donor pool is questionable. Second, Abadie *et al.* make the unusual choice of selecting nominal GDP per capita to measure German reunification's effect on growth, while reporting that they used real GDP per capita. Third, even when real GDP per capita is used as the outcome variable, the results are highly unstable due to the Gerschenkron effect—a known index number problem that makes it impossible to accurately compare the growth rates of different countries.

1. Donor Pool Selection

Donor pool selection is a significant (but often overlooked) problem because of how it affects the permutation tests that are often used to calculate p-values in the SCM (Francis 2025a). As Abadie (2021, 403–405) explains, causation is inferred by estimating placebo effects for each unit in the donor pool: the treatment is iteratively reassigned to them, allowing a synthetic control to be constructed for each. The estimated effect for the actual treated unit is then compared to the distribution of these placebo effects. If the effect for the treated unit is larger in magnitude than most placebo effects, it is considered significant. Hence, a p-value is calculated as:

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

where p is the proportion of units in the entire sample—the treated unit plus all donors (J)—for which the calculated ratio of post-intervention Root

Mean Squared Prediction Error (RMSPE) to pre-intervention RMSPE (r_j) is greater than or equal to the same ratio observed for the actual treated unit (r_1). The donor pool construction matters, then, because all donors are included in the denominator of the *p*-value calculation, even if they are not included in the synthetic control itself.

The potential issue is that countries that performed poorly after 1990 may have been removed from the donor pool. Abadie *et al.* (2015, 501n14) describe how they "started with the 23 OECD member countries in 1990 (excluding West Germany)," but then began to exclude donors. Luxembourg and Iceland were excluded "because of their small size and because of the peculiarities of their economies." And Turkey was dropped due to its low GDP per capita. Ireland is in turn taken out because it "experienced a rapid Celtic Tiger expansion period in the 1990s." And Canada, Finland, and Sweden are excluded as a result of the "profound financial and fiscal crises at the beginning of the 1990s." They have decided they are inappropriate based on what happens after the treatment begins, potentially predetermining the significance of the permutation tests. As a result, it is unclear whether any causality can be robustly inferred from their research design.

A detour into economic history illustrates why the justification given for their decision is questionable from a theoretical perspective. Abadie *et al.* (2015, 500) note that "units that may have suffered large idiosyncratic shocks to the outcome of interest during the study period should also be excluded if such shocks would have not affected the treated unit in the absence of the treatment." Yet there is compelling evidence that the excluded countries' difficulties were not idiosyncratic but rather manifestations of systematic pressures that Germany itself faced. The Nordic crises emerged from the same financial liberalization and property bubble dynamics affecting multiple European economies (Jonung et al. 2009). More fundamentally, all European countries to varying degrees felt the same pressures from integration, labor market rigidities, technological disruption, and demographic transitions—precisely the forces affecting West Germany beyond the specific costs of reunification (Berend 2010). Rather than representing idiosyncratic shocks, then, the excluded countries' poor performance in the 1990s resulted in large part from the same general pressures, making them potentially ideal counterfactuals for isolating reunification's specific effects on West Germany. By excluding them, Abadie *et al.* eliminated precisely the controls that could distinguish reunification's impact from broader 1990s structural changes. And they did so

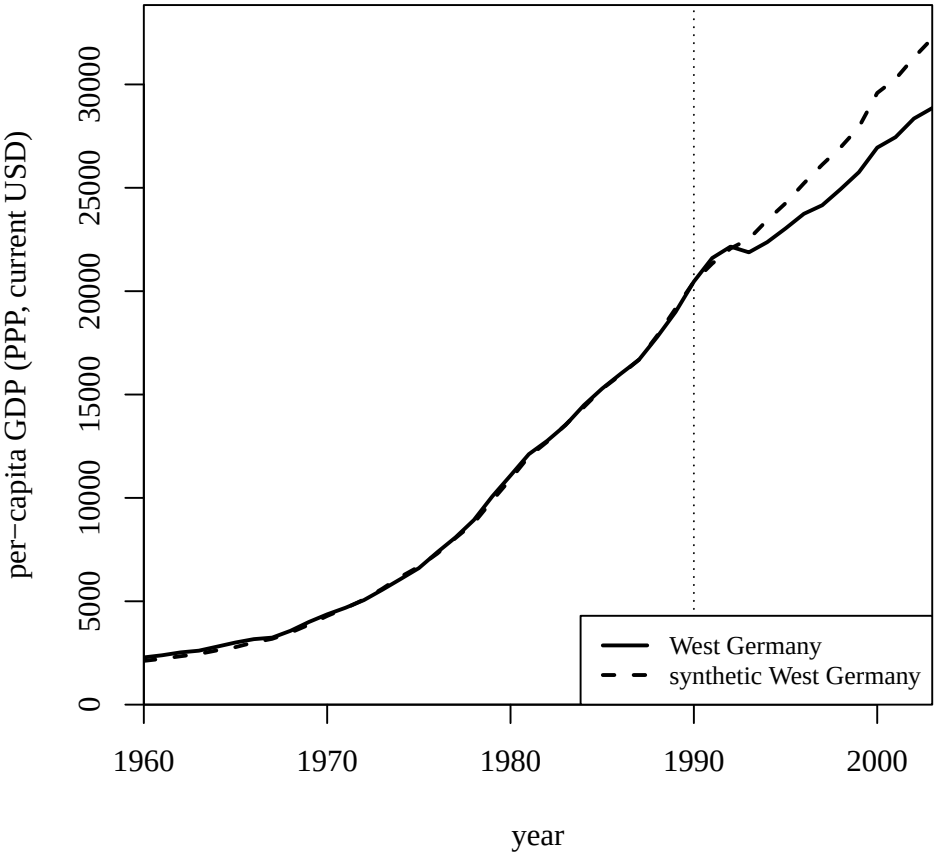

Figure 1
West Germany versus Synthetic West Germany (PPP, current USD)

Notes: The dotted line represents the treatment year. Calculated from data underlying Abadie *et al.* (2015, 503, Figure 2).

in a way that may have inadvertently biased their results to show Germany performing worse than the rest of the OECD, leading to smaller *p*-values calculated from the permutation tests.

2.The Outcome Variable

Abadie *et al.* use nominal GDP per capita as the outcome variable, despite routinely reporting it as real GDP per capita in 2002 PPP USD throughout their article. A visual inspection of Figure 1 illustrates that this cannot have been the case. Abadie *et al.*'s main result shows West Germany's GDP per capita increasing from around $2,300 in 1960 to $20,500 in 1990, implying a compound annual growth rate of 7.6 percent, which is far too fast. In fact, it represents the growth in nominal GDP per capita in PPP USD, which is the variable that Abadie *et al.* actually used in their article and then described as real GDP per capita. Such a rapid increase in West Germany's

GDP per capita is therefore primarily a result of inflation rather than a reflection of what economists normally consider to be "growth."

This use of nominal GDP per capita invalidates Abadie *et al.*'s research design precisely because it is supposed to be about growth. "We apply the synthetic control method to estimate the economic impact of the 1990 German reunification on West Germany," Abadie *et al.* (2015, 496) write. "This example illustrates the potential gains derived by using synthetic controls in comparative case studies," they continue. "We show that no single country is able to closely approximate the values of economic growth predictors for West Germany before the reunification." The benefit of the synthetic control method is thus supposed to be that it can be used to estimate the effects of reunification on West Germany's growth. Hence, when describing Figure 1, Abadie *et al.* (2015, 503) describe how "the two lines diverge substantially" from 1992 onward. "While per capita GDP growth decelerates in West Germany, for the synthetic West Germany per capita GDP keeps ascending at a pace similar to that of the prereunification period." It is clear, then, that Abadie *et al.* are supposed to be measuring the impact of reunification on growth. And yet they use nominal GDP per capita.

This is, of course, an error. Simply put, economists should not use nominal GDP per capita to measure growth. Following the pioneering work of Simon Kuznets (1934), economists have long known that nominal GDP estimates need to be deflated into real values. And Milton Gilbert and Irving B. Kravis (1954) developed the first systematic estimates of real GDP per capita in PPP prices, allowing international comparisons to be made—a research agenda that would subsequently be institutionalized via the Penn World Tables and the Maddison Project. There is therefore a standard methodology for producing real GDP per capita and there are various datasets that can be used.

Crucially, this error is not merely a technical detail but fundamentally undermines the study's credibility. A replication using real GDP per capita in 2002 USD—the reported variable—reveals that Abadie *et al.*'s remarkably close pre-treatment fit largely reflected the use of nominal data, which mechanically captures common inflation trends across countries rather than genuine economic similarity. This can be seen in Figure 2, where the nominal GDP per capita data are replaced by an estimate of real GDP per capita in 2002 USD—the stated unit in their article.[1] As can be seen,

[1] For all countries except West Germany, Abadie *et al.*'s nominal GDP per capita data for 2002 were extended using constant 2015-price GDP from the World Bank's *World*

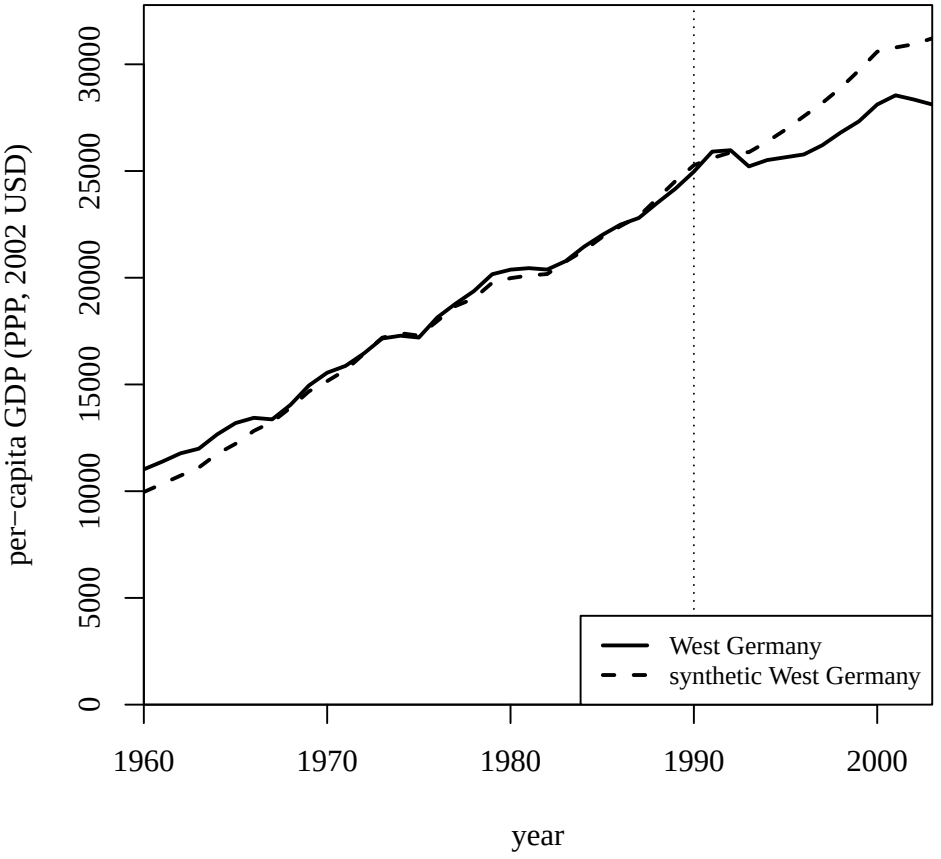

Figure 2
West Germany versus Synthetic West Germany (PPP, 2002 USD)

Notes: The dotted line represents the treatment year. Calculated from data underlying Abadie *et al.* (2015, available at https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/24714), Destatis (2025a, Sheet 8.3; 2025b, Sheet 81000-117), and the World Bank's *World Development Indicators*.

once inflation is accounted for, the pre-treatment fit deteriorates. This deterioration reveals that Abadie *et al.*'s remarkably close pre-treatment fit was largely reflected inflation rather than reflecting genuine pre-treatment similarity between West Germany and its synthetic control. Figures 3 to 6 show, moreover, that there is also a significant deterioration in the other robustness tests when real GDP per capita is used: in Figure 3, there is clearly now a divergence between the actual and synthetic West Germany in the

Development Indicators. For West Germany, the quantity index of real GDP per capita is calculated from the series provided by Destatis (2025a, Sheet 8.3; 2025b, Sheet 81000-117). As discussed in Francis (2025b), care must be taken when using Germany's GDP statistics due to the need to splice the series in 1991. The full construction of the series used here can be found in the file West_Germany_gdp.ods in the replication package accompanying this paper.

pre-treatment period; in Figure 4, there is now a spurious treatment effect in 1975; in Figure 5, Germany's post/pre-period RMPSPE ratio shrinks from 16.3 to 3.7, even if the permutation test continues to produce $p = 0.06$, as in Abadie *et al.*; and in Figure 6, the closeness of fit is considerably more vulnerable to leaving donors out of the pool. The results thus look more fragile when the outcome variable is the reported PPP, 2002 USD.

3. The Gerschenkron Effect

An index number problem may also affect the SCM more generally when referenced data—such as real GDP—are used. As estimates of national output began to be produced, Alexander Gerschenkron (1947) soon realized that there was a problem. The growth rate of Soviet industrial output varied dramatically depending on which reference year was chosen. Methodologists responded with the chain-linking approach, which resolved the problem of measuring growth across time. Yet the spatial Gerschenkron effect has persisted, making it ultimately impossible to measure growth rates across countries in constant PPP USD, despite the best efforts of the economists working on the Penn World Table and the Maddison Project (Deaton and Heston 2010; Bolt et al. 2018).

The mechanics of the spatial Gerschenkron effect are well known. The basic extrapolation method used to estimate GDP in constant PPP USD is:

$$Y_{i,t}^{ref} = Y_{i,ref}^{PPP} \times \frac{Y_{i,t}^{volume}}{Y_{i,ref}^{volume}} \tag{2}$$

where i indexes countries, t indexes years, and *ref* denotes the reference year. $Y_{i,t}^{ref}$ represents country i's GDP per capita in year t using reference year *ref*, $Y_{i,ref}^{PPP}$ is country i's PPP-adjusted GDP per capita in the reference year, and $Y_{i,t}^{volume}$ and $Y_{i,ref}^{volume}$ are the volume indices of GDP per capita from national accounts for years t and *ref* respectively. This approach anchors all countries to a common PPP benchmark but then allows them to grow at their own national accounts rates. But, as Gerschenkron observed, the choice of reference year is not neutral. Consequently, the measured income gap between any two countries depends systematically on which year's price structure is used as the reference:

$$\frac{Y_{WG}^{ref1}}{Y_{US}^{ref1}} \neq \frac{Y_{WG}^{ref2}}{Y_{US}^{ref2}} \tag{3}$$

where Y_{WG} and Y_{US} represent West Germany and the United States respectively, and $ref1 \neq ref2$ are different reference years. This occurs because

Figure 3

Per Capita GDP Gap between West Germany and Synthetic West Germany

Notes: The dotted line represents the treatment year. See Figure 2 for details.

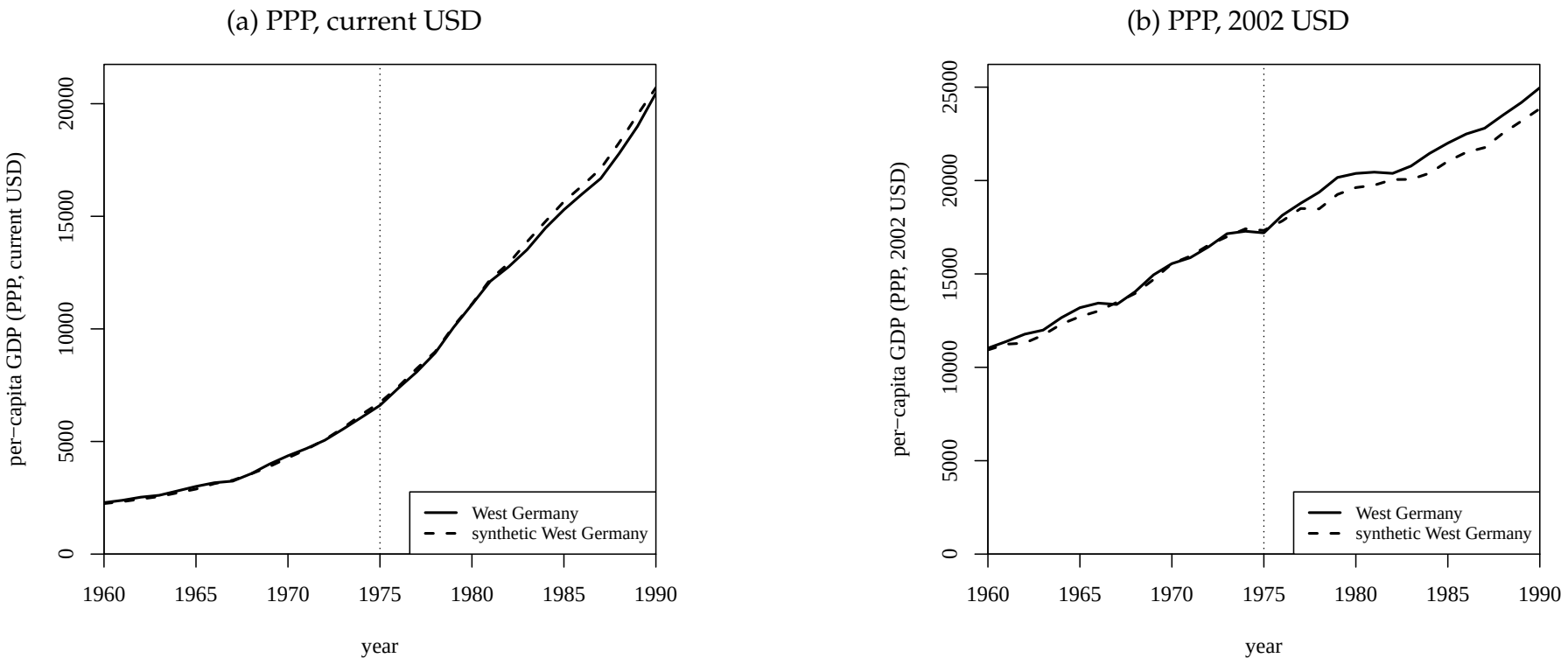

Figure 4

Placebo Reunification 1975–Trends in per Capita GDP: West Germany versus Synthetic West Germany

Notes: The dotted line represents the treatment year. See Figure 2 for details.

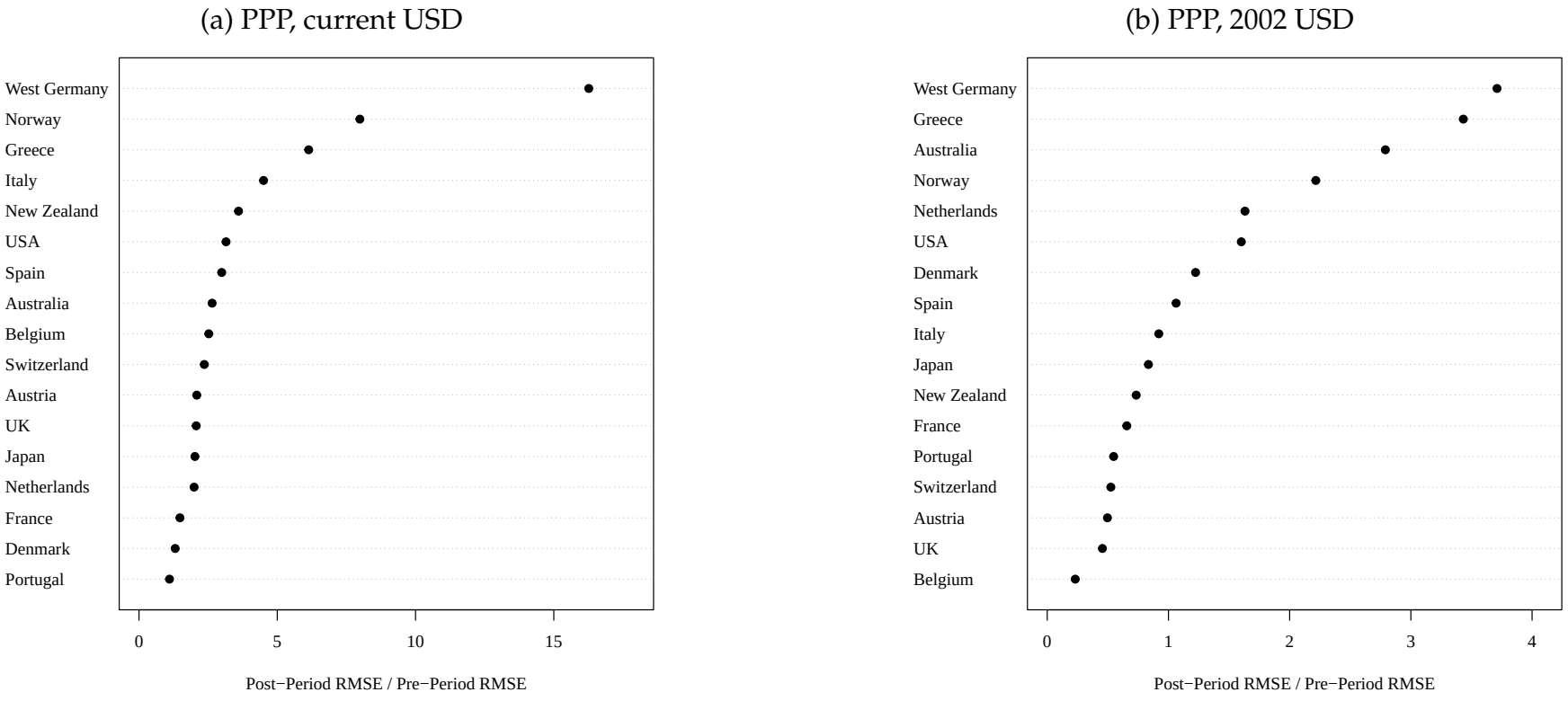

Figure 5

Ratio of Postreunification RMSPE to Prereunification RMSPE: West Germany and Control Countries

Notes: The dotted line represents the treatment year. See Figure 2 for details.

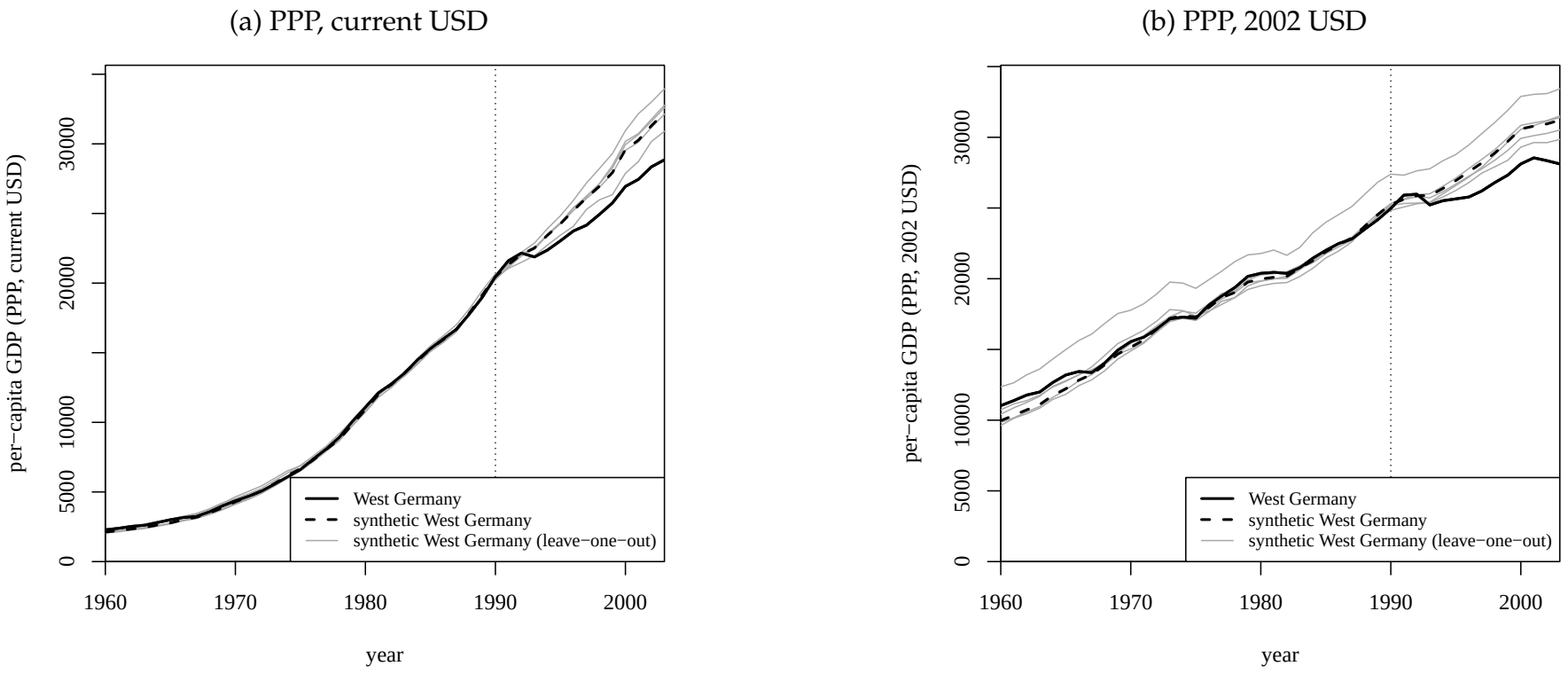

Figure 6

Leave-One-Out Distribution of the Synthetic Control for West Germany

Notes: The dotted line represents the treatment year. See Figure 2 for details.

countries with different economic structures will see their relative positions systematically altered when different price structures are applied to value their output.

For the SCM, the issue has the potential to be particularly acute. The problem is that the estimated treatment effect becomes:

$$\hat{\delta}_t^{ref} = Y_{WG,t}^{ref} - \sum_{j \in \mathcal{J}} w_j^{ref} Y_{j,t}^{ref} \tag{4}$$

where $\hat{\delta}_t^{ref}$ is the estimated treatment effect in year t using reference year ref, $Y_{WG,t}^{ref}$ is West Germany's GDP per capita, $Y_{j,t}^{ref}$ is donor country j's GDP per capita (both in year t using reference year ref), w_j^{ref} represents the weight assigned to donor country j, and $\mathcal{J}$ is the set of donor countries. Crucially, both the pre-treatment matching (which determines the weights w_j^{ref}) and the post-treatment gap calculations depend on the reference year choice through the underlying GDP measures. The result is a fundamental ambiguity that undermines claims to precision:

$$\hat{\delta}_t^{ref1} \neq \hat{\delta}_t^{ref2} \text{ for } ref1 \neq ref2 \tag{5}$$

Since there is no objective basis for preferring one reference year over another, any specific quantitative estimate becomes arbitrary. This arbitrariness extends, moreover, beyond the treatment effect itself to contaminate the entire statistical inference framework. The p-value calculation in Equation 1 depends on comparing RMSPE ratios (r_j) across all units in the sample. Yet these ratios are themselves calculated from further synthetic controls calculated as indices from the reference-year-dependent GDP estimates:

$$r_j^{ref} = \frac{\text{RMSPE}_{j,\text{post}}^{ref}}{\text{RMSPE}_{j,\text{pre}}^{ref}} \tag{6}$$

where both the post-treatment and pre-treatment RMSPE calculations rely on the same GDP measures $Y_{j,t}^{ref}$ that suffer from reference year dependency. Consequently, not only do the treatment effects vary arbitrarily across reference years, but so too do the statistical significance tests:

$$p^{ref1} \neq p^{ref2} \text{ for } ref1 \neq ref2 \tag{7}$$

The SCM thus suffers from what might be termed a compound index number problem: both the substantive estimates and their statistical evaluation depend on index number comparisons that are corrupted by the spatial

Gerschenkron effect. Unlike other econometric methods that might use price indices as inputs, the SCM's core methodology—from weight determination through treatment effect calculation to significance testing—is fundamentally built on layered index number comparisons, making it particularly vulnerable to the arbitrariness that Gerschenkron identified.

This arbitrariness can be demonstrated through multiple replications of Abadie *et al.*'s research design, with real GDP per capita referenced to all available years. Using PPP, 2002 USD, as in Figure 2, the treatment effect—the gap between the actual and synthetic West Germany as a percentage of the latter—is -9.9 percent, with $p = 0.06$. Yet Figure 7 illustrates how the results fluctuate wildly depending on the reference year: at one extreme is the PPP, 2001 USD version, with a –10.5 percent gap, while at the other is PPP, 1966 USD version, with a –2.2 percent gap. Abadie *et al.*'s (2015, 495–496, 500) claim to be able to achieve "precise quantitative inference" therefore seems erroneous. Furthermore, Figure 8 shows large fluctuations in the *p*-values calculated using the permutation tests from Equation 1: just 8 of the 44 versions of the SCM for West Germany calculated with reference years from 1960 to 2003 returned $p < 0.10$. Whether or not the analysis seems valid depends, then, on the arbitrary choice of reference year for real GDP per capita.[2]

The spatial Gerschenkron effect is such that there is no way to accurately measure the impact of German reunification on West Germany's growth relative to a hypothetical counterfactual made up of a weighted average of other countries. Figures 7 and 8 indicate that the results of Abadie *et al.*'s research design are unstable and their significance depends on the reference year used. A similar conclusion would, moreover, apply to other attempts to use the SCM to assess the impact of an event on growth rates in a particular country. Indeed, the issue will arise when using the SCM with any kind of referenced data. Going forward, then, researchers should check that their results are consistent across multiple reference years.

Postscript on an Erratum

In an Erratum in the *American Journal of Political Science*, Abadie *et al.* (2025) have sought to correct one of the issues discussed here. They recognize that they used nominal GDP per capita, stating that the "error stems from legacy labels carried over from an earlier draft." By their own account, then,

[2] It seems likely that the instability is aggravated by the cross-validation technique used by Abadie *et al.* (2011) in the *synth* software package, as Klößner *et al.* (2018) previously highlighted.

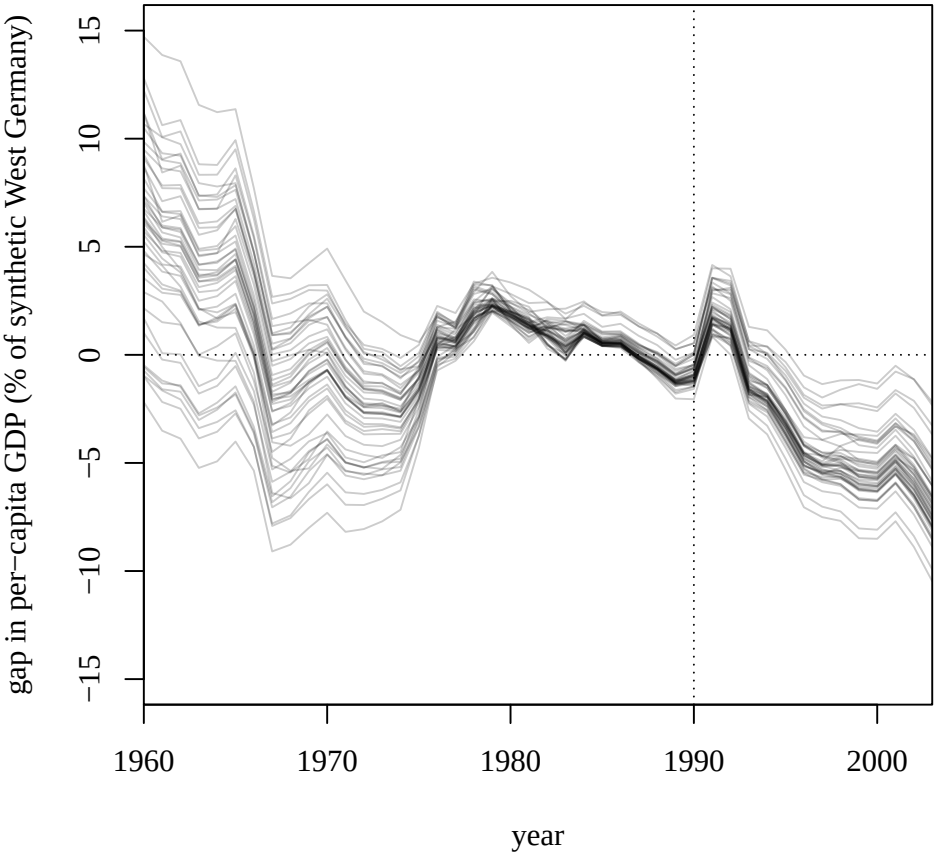

Figure 7
Per Capita GDP Gap

Notes: The dotted line represents the treatment year. The gaps between 44 synthetic controls for West Germany and actual West Germany are shown as a percentage of the former. The 44 synthetic controls were calculated as in Figure 2, but with real GDP per capita in PPP USD referenced to each of the years from 1960 to 2003.

they used nominal GDP per capita in a study of growth, but then accidentally reported having used real GDP per capita. "To our knowledge," they continue, "no harmonized real GDP series at constant PPP units exists for West Germany and the donor pool during the sample period." Yet Abadie *et al.* obtained their nominal GDP per capita from the OECD, which has long published them alongside real PPP GDP per capita for West Germany and the donor countries. The World Bank similarly publishes real GDP per capita for all these countries from 1960 onward, which can then easily be converted to a PPP basis using Equation 2, as has been done in this paper.

In their Erratum, Abadie *et al.* (2025) go on to use two unconventional approaches to estimate real GDP per capita. First, they continue to use nominal data but "divide both the West Germany and synthetic control

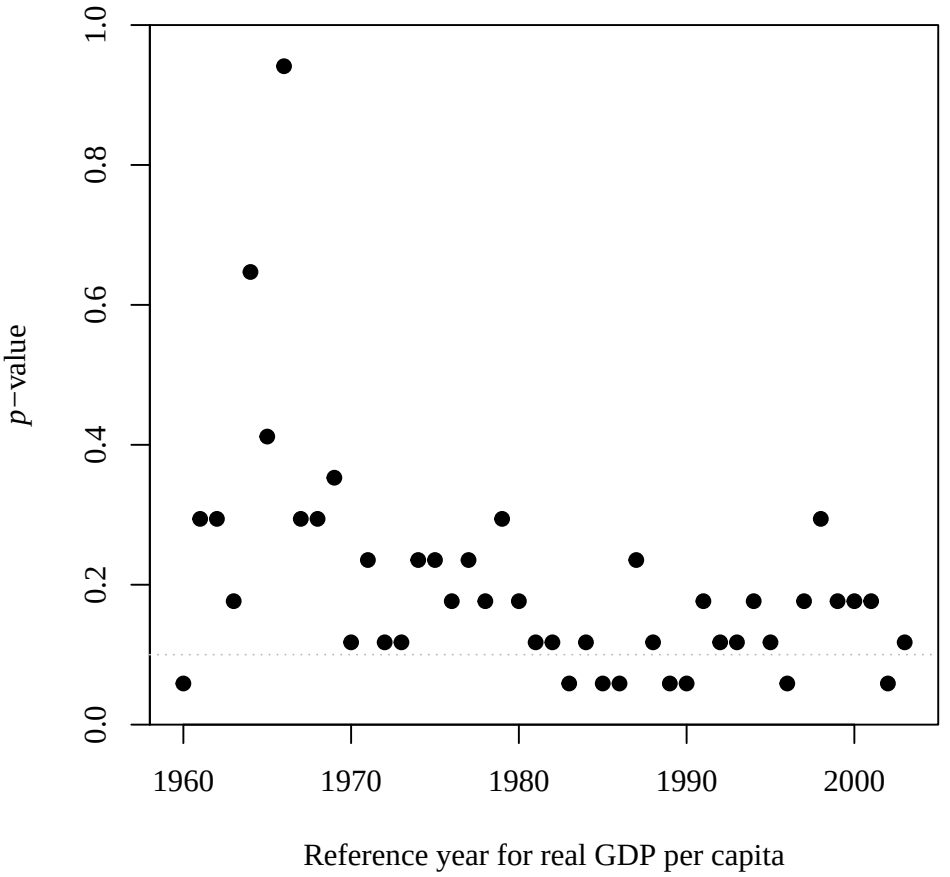

Figure 8
p-Values from Permutation Tests (PPP, Constant USD)

Note: *p*-values for the 44 synthetic controls in Figure 7 were calculated using the permutation tests, as in Equation 1.

series (current GDP in PPP units) by a common deflator"—that of the United States. Second, they "deflate per-capita GDP before constructing the synthetic control," again by the US GDP deflator, which seems to be based on the belief that the conversion from nominal to real PPP GDP per capita should be performed as:

$$Y_{i,t} = Y_{i,t}^{PPP,current} \times \frac{D_{US,base}}{D_{US,t}} \tag{8}$$

where $Y_{i,t}$ is real GDP per capita for country i at time t, $Y_{i,t}^{PPP,current}$ is nominal PPP GDP per capita for country i at time t, $D_{US,t}$ is the US GDP deflator at time t, and $D_{US,base}$ is the US GDP deflator at the base year. This approach incorrectly applies only the US deflator to all countries, ignoring country-specific inflation rates. Assuming the United States is still the reference country and all values have been converted to US dollars,

the correct conversion from nominal to real PPP GDP per capita should be performed as:

$$Y_{i,t} = Y_{i,t}^{PPP,current} \times \frac{(D_{US,t}/D_{US,base})}{(D_{i,t}/D_{i,base})} \tag{9}$$

where $D_{i,t}$ is the GDP deflator for country i at time t, and $D_{i,base}$ is the GDP deflator for country i at the base year. This formula properly accounts for relative price level changes between countries by adjusting for both the reference country's and the individual country's inflation rates. In effect, Equation 9 is functionally equivalent to Equation 2, which has been used in this paper.

References

Abadie, Alberto (2021) "Using Synthetic Controls: Feasibility, Data Requirements, and Methodological Aspects." *Journal of Economic Literature*. 59(2), 391–425.

Abadie, Alberto and Alexis Diamond (2025) "Erratum: Comparative Politics and the Synthetic Control Method." *American Journal of Political Science*. forthcoming.

Abadie, Alberto, Alexis Diamond, and Jens Hainmueller (2011) "*synth*: An R Package for Synthetic Control Methods in Comparative Case Studies." *Journal of Statistical Software*. 42, 1–17.

— (2015) "Comparative Politics and the Synthetic Control Method." *American Journal of Political Science*. 59(2), 495–510.

Berend, Iván T. (2010) *Europe since 1980*. Cambridge: Cambridge University Press.

Bolt, Jutta, Robert Inklaar, Herman de Jong, and Jan Luiten van Zanden (2018) "Rebasing 'Maddison': New Income Comparisons and the Shape of Long-Run Economic Development." Unpublished paper, GGDC Research Memorandum 174, Jan. https://www.rug.nl/ggdc/html_publications/memorandum/gd174.pdf.

Deaton, Angus and Alan Heston (2010) "Understanding PPPs and PPP-based National Accounts." *American Economic Journal: Macroeconomics*. 2(4), 1–35.

Destatis (2025a) "Bruttoinlandsprodukt, Bruttowertschöpfung | Statistikportal.de." Mar. 28. https://www.statistikportal.de/de/veroeffentlichungen/bruttoinlandsprodukt-bruttowertschoepfung.

— (2025b) "Statistischer Bericht: Volkswirtschaftliche Gesamtrechnungen: 2. Vierteljahr 2025." Sept. 2. https://www.destatis.de/DE/Themen/Wirtschaft/Volkswirtschaftliche-Gesamtrechnungen-Inlandsprodukt/Publikationen/Downloads-Inlandsprodukt/statistischer-bericht-2180120.html.

Francis, Joseph (2025a) "A *p*-Hackers Guide to the Synthetic Control Method." Unpublished paper, v1.0, June 2. https://github.com/joefrancis505/Francis_A_p-Hackers_Guide.

Francis, Joseph (2025b) "Germany's GDP per Capita before 1991 A Replication from Destatis to the World Bank." Unpublished paper, Sept. 27.

Gerschenkron, Alexander (1947) "The Soviet Indices of Industrial Production." *Review of Economics and Statistics*. 29(4), 217–226.

Gilbert, Milton and Irving B. Kravis (1954) *An International Comparison of National Products and the Purchasing Power of Currencies: A Study of the United States, the United Kingdom, France, Germany and Italy*. Paris: Organisation for Economic Co-operation and Development.

Lars Jonung, Jaakko Kiander, and Pentti Vartia (eds.) (2009) *The Great Financial Crisis in Finland and Sweden: The Nordic Experience of Financial Liberalization*. Cheltenham, UK: Edward Elgar.

Klößner, Stefan, Ashok Kaul, Gregor Pfeifer, and Manuel Schieler (2018) "Comparative Politics and the Synthetic Control Method Revisited: A Note on Abadie et al. (2015)." *Swiss Journal of Economics and Statistics*. 154(1), 11.

Kuznets, Simon (1934) "National Income, 1929–1932." NBER Bulletin 49. https://www.nber.org/books-and-chapters/national-income-1929-1932/national-income-1929-1932.