

Abadie, Alberto; Diamond, Alexis; Hainmueller, Jens

Working Paper

Response to Francis (2025)

I4R Discussion Paper Series, No. 272

Provided in Cooperation with:

The Institute for Replication (I4R)

Suggested Citation: Abadie, Alberto; Diamond, Alexis; Hainmueller, Jens (2025) : Response to Francis (2025), I4R Discussion Paper Series, No. 272, Institute for Replication (I4R), s.l.

This Version is available at:

<https://hdl.handle.net/10419/330836>

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.



No. 272
DISCUSSION PAPER SERIES

Reponse to Francis (2025)

Alberto Abadie

Alexis Diamond

Jens Hainmueller

This paper responds to:

Francis, J.. 2025. A Replication of “Comparative Politics and the Synthetic Control Method” by Abadie et al. (2025). *I4R Discussion Paper Series* No. 271. Institute for Replication.

November 2025

I4R DISCUSSION PAPER SERIES

I4R DP No. 272

Response to Francis (2025)

Alberto Abadie¹, Alexis Diamond², Jens Hainmueller³

¹Massachusetts Institute of Technology (MIT), Cambridge/USA

²Minerva University, San Francisco/USA

³Stanford University, Stanford/USA

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](https://www.zbw.eu/), and [RWI – Leibniz Institute for Economic Research](https://www.rwi-essen.de/), 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

Response to Francis (2025)

Alberto Abadie, MIT
Alexis Diamond, Minerva University
Jens Hainmueller, Stanford University

Response to Francis (2025) initial submission

We thank Joseph Francis for pointing out the mislabeling of the per-capita GDP variable in the empirical application of Abadie, Diamond, and Hainmueller (2015). The article reported the variable in PPP 2002 USD. The correct units are PPP current USD. The error stems from legacy labels carried over from an earlier draft.

While the change in units does not affect the validity of the analysis, it does alter the interpretation of the results of the empirical application. Specifically, absolute differences in per-capita GDP between West Germany and its synthetic control should be interpreted in PPP adjusted current USD prices, not constant 2002 USD prices.

To our knowledge, no harmonized real GDP series at constant PPP units exists for West Germany and the donor pool during the sample period. To approximate real per-capita GDP, one can divide both the West Germany and synthetic control series (current GDP in PPP units) by a common deflator. This changes the units of measurement, but leaves the estimated percentage effects unchanged. Alternatively, one can deflate per-capita GDP before constructing the synthetic control. Results from both approaches are nearly identical. We report the results of both approaches in a supplementary appendix to Abadie, Diamond, and Hainmueller (2025).

Francis (2025) also offers a critical re-analysis of the empirical illustration in our article. In it, Francis purports to estimate a synthetic control for West Germany using data on real per-capita GDP. We raise two key objections to his analysis.

First, the pre-1990 GDP and population series Francis draws from the World Development Indicators refer to unified Germany, meaning the combined totals for West and East Germany, not West Germany alone. This error invalidates the analysis in the Francis replication. The Maddison Project data he also uses suffers from the same problem.

Second, the credibility of synthetic control studies hinges on the quality of the pre-treatment fit of the predictors (see Abadie, 2021), and much diligence is exerted in synthetic control studies to achieve proper fit. Francis (2025) ignores this requirement. The resulting synthetic controls (especially, the preferred specification in Figure 8) exhibit large discrepancies relative to Germany's per-capita GDP before reunification.

Despite these flaws in the analysis, Francis (2025) arrives at the same conclusions as Abadie, Diamond, and Hainmueller (2015): large negative estimates of the German reunification's im-

This response reproduces some of the material in an erratum, Abadie, Diamond, and Hainmueller (2025), submitted by the authors to the *AJPS*.

pact (Figure 3(b)) and extreme post-reunification gaps for Germany (Figure 6(b)). Contrary to Francis’s assertion, his results do not support “an argument that is the opposite to [Abadie, Diamond, and Hainmueller (2015)]”. Francis’s own results support the conclusions in Abadie, Diamond, and Hainmueller (2015).

Finally, Francis questions our choice of donor pool. We stand by our selection, which we explain in detail in the article. Excluding units affected by large idiosyncratic shocks to avoid distortions in estimation *and inference* is standard practice, not just for synthetic control methods (Abadie, 2021), but more broadly for policy evaluation, irrespective of the method used for the analysis.

Response to Francis (2025) revised submission

Following our response to Francis’s (2025) initial submission, Joseph Francis submitted a revised report, as permitted under I4R procedures. This section addresses the new material in that second report.

1. In our response to Francis’s (2025) initial submission (see previous section), we pointed out that the pre-1991 World Development Indicators (WDI) data Francis used in his report covered the sum of West and East Germany, not West Germany alone. Although we have sought to eliminate this misunderstanding, the revised version of Francis (2025) repeats the same error, treating the WDI series as if it referred only to West Germany before reunification. This mistake invalidates the replication attempts in Francis (2025), because the pre-reunification data in Francis (2025) do not measure GDP per-capita in West Germany, as Francis believes.

Verifying this assertion is straightforward. Figure 1 plots population, real GDP, and land area of Germany, extracted from the same WDI files cited in Francis (2025).¹ If the series referred to West Germany before 1991 and to unified Germany afterward, each of the three measures would display a sharp increase around the time of reunification. Instead, population and land area remain at levels consistent with unified Germany (about 80 million people and 350,000 sq. km), and real GDP shows no increase in 1991 large enough to suggest the addition of East Germany’s output. We contacted the World Bank and obtained confirmation via email that the WDI series for Germany report data for the sum of West and East Germany in years before 1991.

The WDI real GDP per-capita series used in Francis (2025) is identical to the series in Figure 1, obtained by dividing real GDP by population.² Table 1 confirms this identity. Column (1) reports GDP per-capita values for 1981-1990 extracted with the Francis (2025) code and labeled as West Germany in his document. Column (2) reports GDP per-capita values from the data underlying Figure 1. The comparison shows that the series Francis (2025) attributes to West Germany is in fact data for unified Germany.

Francis’s (2025) post-reunification data are also invalid. He constructs them by taking regional accounts data for West Germany after reunification and scaling by the ratio of

¹WDI series codes SP.POP.TOTL, NY.GDP.MKTP.KD and AG.LND.TOTL.K2, respectively.

²WDI series code $NY.GDP.PCAP.KD = NY.GDP.MKTP.KD / SP.POP.TOTL$.

	GDP per-capita West Germany in Francis (2025)	GDP per-capita Germany from data in Figure 1
Year	(1)	(2)
1981	24067.48	24067.48
1982	23995.27	23995.27
1983	24436.55	24436.55
1984	25213.38	25213.38
1985	25858.06	25858.06
1986	26437.42	26437.42
1987	26766.96	26766.96
1988	27651.04	27651.04
1989	28507.16	28507.16
1990	29747.69	29747.69

Table 1: GDP per-capita series, 1981–1990

real per-capita GDP in the WDI database in 1991 to the corresponding value from the German regional accounts. Because the scaling factor uses WDI values for unified Germany, Francis’s procedure cannot produce correct post-reunification data for West Germany. The procedure is numerically equivalent to extrapolating the 1991 WDI GDP figure for unified Germany forward in time using West Germany’s growth rates from the regional accounts.

2. In our response to Francis’s initial submission, we say our results “should be interpreted in PPP adjusted USD prices, not constant 2002 USD prices.” In the revised version of the report, Francis is not satisfied with this solution, claiming that the aim of our paper is to make comparisons in time. This is incorrect. The primary aim of our paper is to make comparisons across cross-sectional units: West Germany and the synthetic control composed by a weighted average of the units in the donor pool.
3. Francis (2025) objects to our conversion of PPP GDP into constant dollars. Our procedure is routine and yields easily interpretable magnitudes. PPP GDP is reported in current US dollars. We convert these figures to constant dollars using the US GDP deflator.

Equations (4) and (7) in the appendix show that Francis’s construction is algebraically equivalent to ours within a PPP benchmark.

Where does the difference come from then? The derivations in the technical appendix clarify this point. Although equations (4) and (7) show that Francis’s proposal and ours are algebraically equivalent within a benchmark, the construction of the series based on equation (2) and (5) rely on two very different implementations. Our formulation in (5) produces one single series, calculated over all values of GDP_{jt}^{PPP} , the OECD re-based current PPP GDP series, which stitches successive PPP benchmarks together. By contrast, Francis reconstructs a new series for each possible reference year using a single observation of current PPP GDP as an anchor.

Equations (3) and (6) clarify the difference between the two approaches. Consider first the approach in Francis (2025), described in equation (3). This procedure amounts to taking GDP in local currency units at time t , deflating it to local currency units at time r , and then applying the PPP exchange rate at time r . That is, it imposes the same PPP exchange rate at time r for all periods after deflating. It should not then come as a surprise that this approach produces series that are very sensitive to the choice of r .

In contrast, our approach as described in equation (6) applies the PPP exchange rate at time t to convert GDP at time t in local currency units to US dollars, and then deflates the resulting values to time r . This produces a single series that is not sensitive to changes in r . In our method, changing the reference year r only rescales the series by a proportionality factor common to all countries and years. This adjustment merely changes the units of measurement to reflect a different base year for the US dollar, and therefore leaves our inferential exercises unaffected.

Our approach only extrapolates between benchmark years. This is because our approach employs the entire series of current PPP values, which are extrapolated between benchmark years (see equation (1)). By contrast, each series in Francis (2025) extrapolates across the entire time span, using the PPP exchange rate of a single year to infer PPP values for all other years, even decades apart.

It is worth noting that single-benchmark extrapolation is precisely the method used in older versions of the Penn World Table (through version 7). In those releases, one ICP benchmark (e.g. 1996 or 2005) anchored the entire dataset, and all other years were extrapolated using relative inflation. As the PWT developers themselves recognized, this approach made results highly sensitive to benchmark choice. Starting in 2013 with PWT 8, they abandoned the single-benchmark method for comparisons between countries and adopted a chain-linked, multi-benchmark approach. Our implementation aims to follow current practice to the extent that is possible with the available data, whereas Francis's procedure revives the limitations of the older, single-benchmark method.³

To sum up, Francis (2025) does not uncover a flaw in the synthetic control method, but instead illustrates a well-known limitation of single-benchmark PPP series. By re-anchoring at every possible reference year, his procedure amplifies the benchmark-sensitivity of PPP data, whereas the re-based series we use minimizes this effect and provides a robust basis for inference.

Moreover, Francis PPP extrapolations incorrectly use data for the unified Germany as opposed to West Germany. As we explained in our initial response, to our knowledge no harmonized series of real GDP at constant PPP units exists for West Germany and the donor pool during the sample period (see point 4).

³See R. C. Feenstra, Inklaar, and M. P. Timmer, 2015 and Inklaar, M. Timmer, and R. Feenstra, 2013. As they emphasize: "In version 8.0, we use all available purchasing power parity benchmark data under the principle that the best estimate of real GDP in, say, 1980 is based on purchasing power parity data for 1980 rather than on extrapolations from 2005."

Francis (2025) further claims that our procedure ignores country-specific inflation rates. This assertion is incorrect. Our procedure begins with current PPP data. Country inflation is incorporated in the construction of PPP exchange rates.

4. Francis (2025) asserts that, “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.” This claim is incorrect. The OECD database we used in our paper provides only PPP GDP per capita, not real PPP GDP per capita. He further asserts that “The World Bank similarly publishes real GDP per capita for all these countries from 1960 onward...,” which is also incorrect. As demonstrated above, the World Bank WDI data refer to unified Germany throughout.

Response to Francis (2025), third round

After receiving our response to the second version of his document, Francis submitted a third version to I4R. This section briefly discusses additional material introduced in the third version of Francis’s document and identifies errors that remain uncorrected from the second version.

1. In a note accompanying the most recent version of his document, Francis acknowledges that the pre-1991 WDI data for Germany that he used in the previous versions of his document do not refer to West Germany. We are glad to be in agreement on this point. The estimates in his latest version seem to no longer treat pre-1991 WDI data for Germany as data for West Germany.
2. The latest version of Francis (2025) continues to rely on single-benchmark extrapolation. As we already explained in our previous response, Francis’s approach maximizes sensitivity to the choice of base year and has long been abandoned in cross-country comparisons. The variability in Francis’s results across base years generates only from his reliance on single-benchmark extrapolation, and not from a flaw in the synthetic control method.

Even under single-benchmark extrapolation, and ignoring that in Francis’s analysis some base year choices fit the pre-treatment period poorly, his corrected dataset yields negative effects of German reunification for every choice of base year (Figure 7 in the latest version of his document).
3. The latest version of Francis’s document repeats the same incorrect claims from previous versions. It still claims that the purpose of our article is to compare growth rates. We have already explained why this is not the case. It continues to claim that our procedure ignores country-specific inflation. We have already explained that this claim is incorrect in our response to the second version of Francis’s document. The errors noted in point 4 of our response to the second version of Francis’s document remain uncorrected in the latest version of Francis’s document.

We once more thank Joseph Francis for pointing out the mislabeling of the per-capita GDP variable in the empirical application of Abadie, Diamond, and Hainmueller (2015).

Technical appendix

Between benchmark years, PPPs extrapolate using relative inflation rates. That is, for year t attached to benchmark year b ,

$$PPP_{j,t} = PPP_{j,b} \frac{P_{j,t}/P_{j,b}}{P_{US,t}/P_{US,b}}, \quad (1)$$

where $P_{j,t}$ is the GDP deflator for country j at time t . GDP in current PPP units is obtained as

$$GDP_{j,t}^{PPP} = \frac{GDP_{j,t}^{LCU}}{PPP_{j,t}},$$

where $GDP_{j,t}^{LCU}$ is the GDP of country j at time t in local currency units.

Francis (2025) formula is

$$Y_{j,t}^{(r)} = GDP_{j,r}^{PPP} \frac{GDP_{j,t}^{real(2015)}}{GDP_{j,r}^{real(2015)}}, \quad (2)$$

where

$$GDP_{j,t}^{real(2015)} = GDP_{j,t}^{LCU} \frac{P_{j,2015}}{P_{j,t}} \frac{1}{e_{j,2015}},$$

$P_{j,t}$ is the GDP deflator for country j at time t , and $e_{j,2015}$ is the exchange rate between country j 's currency and the US dollar in 2015. As a result,

$$\begin{aligned} Y_{j,t}^{(r)} &= GDP_{j,r}^{PPP} \frac{GDP_{j,t}^{LCU}}{GDP_{j,r}^{LCU}} \frac{P_{j,r}}{P_{j,t}} \\ &= \frac{GDP_{j,t}^{LCU}}{PPP_{j,r}} \frac{P_{j,r}}{P_{j,t}}. \end{aligned} \quad (3)$$

Suppose that the period r is attached to benchmark year b . Then,

$$PPP_{j,r} = PPP_{j,b} \frac{P_{j,r}/P_{j,b}}{P_{US,r}/P_{US,b}}.$$

This gives

$$Y_{j,t}^{(r)} = \frac{GDP_{j,t}^{LCU}}{PPP_{j,b}} \frac{P_{j,b}}{P_{j,t}} \frac{P_{US,r}}{P_{US,b}}. \quad (4)$$

Our formula is

$$GDP_{j,t}^{PPP(US,r)} = GDP_{j,t}^{PPP} \frac{P_{US,r}}{P_{US,t}}. \quad (5)$$

Substituting the extrapolated $PPP_{j,t}$,

$$PPP_{j,t} = PPP_{j,b} \frac{P_{j,t}/P_{j,b}}{P_{US,t}/P_{US,b}},$$

we obtain

$$GDP_{j,t}^{PPP(US,r)} = \frac{GDP_{j,t}^{LCU}}{PPP_{j,t}} \frac{P_{US,r}}{P_{US,t}} \quad (6)$$

$$= \frac{GDP_{j,t}^{LCU}}{PPP_{j,b}} \frac{P_{j,b}}{P_{j,t}} \frac{P_{US,r}}{P_{US,b}}. \quad (7)$$

Equations (4) and (7) show that Francis’s construction is algebraically equivalent to ours within benchmark.

References

- Abadie, Alberto (2021). “Using Synthetic Controls: Feasibility, Data Requirements, and Methodological Aspects”. In: *Journal of Economic Literature* 59.2, pp. 391–425.
- Abadie, Alberto, Alexis Diamond, and Jens Hainmueller (2015). “Comparative Politics and the Synthetic Control Method”. In: *American Journal of Political Science* 59.2, pp. 495–510.
- (2025). “Erratum: Comparative Politics and the Synthetic Control Method”.
- Feenstra, Robert C., Robert Inklaar, and Marcel P. Timmer (Oct. 2015). “The Next Generation of the Penn World Table”. In: *American Economic Review* 105.10, pp. 3150–82.
- Francis, Joseph (2025). “Comparative Politics and the Synthetic Control Method Revisited”.
- Inklaar, Robert, Marcel Timmer, and Robert Feenstra (2013). *Recasting international income differences: The next generation Penn World Table*. VoxEU.org, published 7 October 2013. URL: <https://cepr.org/voxeu/columns/recasting-international-income-differences-next-generation-penn-world-table> (visited on 06/08/2024).

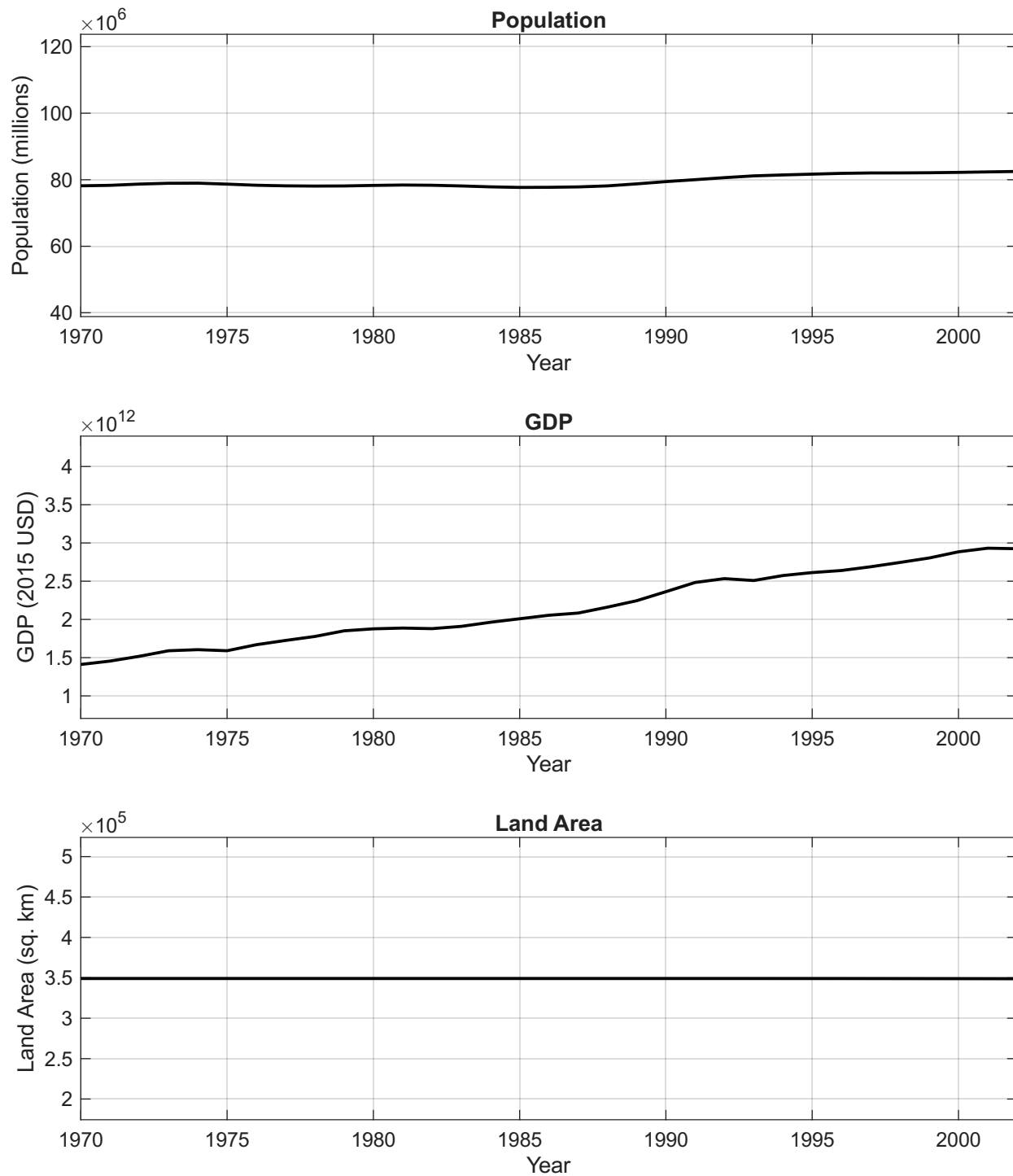


Figure 1: World Development Indicators: Germany 1970-2002