# The Very Long-Run Costs of Financial Crises*

Karsten Müller† Chenzi Xu‡ Mohamed Lehbib§

September 2026

**Abstract**

We provide new empirical evidence of the aftermath of financial crises by examining output dynamics over unprecedented time horizons. To do so, we introduce the Global Macroeconomic Database, which systematically collects and harmonizes sources for 200+ economies and extends back to the origins of modern data collection. Using this extended historical coverage, we find that banking crises are associated with real GDP contractions of around 10% relative to pre-existing trends that persist for five decades following the initial shock. We also document similarly prolonged disruptions to human and physical capital accumulation, alongside weakening fiscal positions marked by declining revenues, higher spending in the aftermath of crises, and elevated debt. These results imply that crises are not just temporary disruptions of the business cycle but hold back economic development.

*This paper subsumes and replaces an earlier paper, "The Global Macro Database: A New International Macroeconomic Dataset" (Müller, Xu, Lehbib and Chen, 2025). The data, code base, and additional documentation of the Global Macro Database (Müller, Xu, Lehbib and Chen, 2025) is available at www.globalmacrodata.com. We would like to thank Sebnem Kalemli-Özcan for her comments. Astrid Tonnet provided excellent research assistance.

†SAFE, Goethe University, National University of Singapore (on leave), & CEPR, mueller@safe-frankfurt.de

‡University of California, Berkeley & NBER & CEPR; chenzixu@berkeley.edu

§National University of Singapore, NUS Business School, lehbib@u.nus.edu

# 1 Introduction

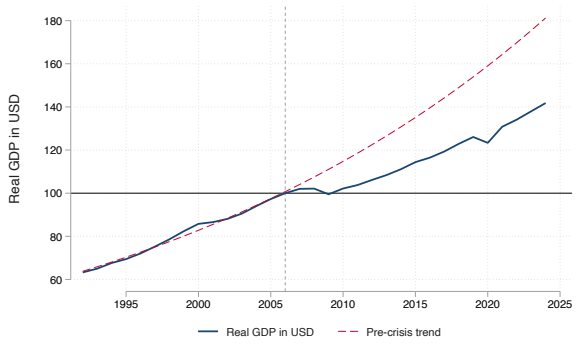
The global financial crisis of 2008 is widely viewed as a defining historical event. In many economies around the world, output fell sharply during the crisis, and the subsequent recovery was slow (Ball, 2014). Almost two decades later, economic activity remains below the trajectory implied by pre-crisis growth trends.

Figure 1 illustrates this pattern for two widely discussed cases, the United States and the United Kingdom. In each figure, we plot the realized path of real GDP (solid blue line) along with the projected path based on the pre-crisis growth trend, calculated as average growth between years $t-15$ and $t-1$ (dashed red line). In both countries, output fell sharply during the crisis and recovered only gradually. In addition, the recovery in both cases has been incomplete in the sense that real GDP remains substantially below the path implied by pre-crisis growth. Appendix Figure C.1 shows several other prominent examples.

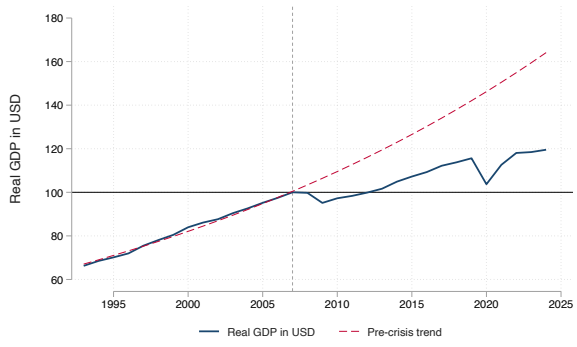
These patterns speak to the debate on whether financial crises generate long-run economic damage (e.g., Cerra and Saxena, 2008; Xu, 2022) or whether economies eventually return to levels implied by their previous growth paths (e.g., Davis and Weinstein, 2002). Specifically, how persistent are the macroeconomic consequences of financial crises? Answering this question is challenging to answer in the data due to the rare nature of crises. The ideal testing ground would be the combination of two features unavailable in existing datasets: broad cross-country coverage and very long-run time series. Only these would allow tracing economic dynamics over horizons measured in decades rather than years.

Figure 1: Output Dynamics Following the Global Financial Crisis

(a) United States




(b) United Kingdom




*Notes:* This figure illustrates the evolution of real GDP around the 2007-2008 Global Financial crisis for the United States and United Kingdom, respectively. Each panel plots real GDP (solid line) normalized to 100 in the year preceding the crisis, together with the counterfactual trajectory implied by the pre-crisis trend (dashed line). The pre-crisis trend is estimated using average growth between years $t-15$ and $t-1$, where $t$ denotes the crisis year. Vertical dashed lines indicate the onset of the banking crisis. Crisis dates follow the banking-crisis onset indicator described in Section 2.

This paper provides new evidence on this question using a newly constructed Global Macro

Database (GMD) that links modern statistics to historical time series for the full set of countries for which macroeconomic data exist. The database covers 46 across 239 countries and extends back to the origins of modern statistical measurement, combining the cross-country breadth of modern macroeconomic datasets with the temporal depth of historical reconstructions. The resulting sample is substantially larger than those of prior work (Appendix Figure A.4 provides a detailed comparison).

Studying the impact of crises over this extended sample also requires banking-crisis dates with comparable breadth. The principal chronologies in the literature—Reinhart and Rogoff (2009*b*), Laeven and Valencia (2013), Jordà, Schularick and Taylor (2017*a*), and Baron, Verner and Xiong (2021)—differ in definitions, country coverage, time coverage, and dating conventions. We consolidate these sources into a single onset indicator using a source-support rule that groups records referring to the same underlying event and assigns the onset to the year supported by more chronologies, with ties assigned to the earlier date. For sources that record every year of an ongoing crisis, we retain the first year of the episode.

We define a banking crisis as an episode of distress in the banking sector that results in losses, liquidations, or failures of one or more financial institutions. The source-support rule addresses the most common differences across chronologies: exact dating, whether the same event in adjacent years constitutes one episode or several, and variation in source coverage across countries and periods.

We estimate the dynamic effects of banking crises using local projections, which allow us to trace the evolution of output over long horizons without imposing strong parametric restrictions on adjustment dynamics. Each episode enters our estimation as a one-year impulse: the indicator equals one in the crisis onset year and zero otherwise. In addition, the crisis indicator is coded as zero only for country-years that at least one source explicitly covers as non-crisis observations; otherwise it is coded as missing, which restricts the estimation sample to periods covered by the underlying chronologies.

Because the horizon-$h$ coefficients capture the cumulative output response at every projection horizon, modeling each crisis as a one-year impulse places no restriction on the persistence the framework can identify. A central feature of our design is to include flexible controls of pre-crisis growth dynamics over short, medium, and long horizons, which capture the economic build-ups that often precede crises, and which ensure that post-crisis estimates are not conflated with the unwinding of pre-existing trends. Implementing this design with the required lag structure while still obtaining precise estimates at horizons of up to fifty years demands large samples, which the breadth and length of the GMD is uniquely positioned to provide.

Our central finding is that banking crises generate extremely persistent output losses. Real GDP declines sharply following a crisis and remains significantly below its pre-crisis trajectory for

decades. Even fifty years after the initial shock, average output remains substantially depressed relative to the counterfactual path implied by pre-crisis growth dynamics. These long-run losses are considerably larger than those estimated in previous studies that examine shorter horizons or rely on more limited samples.

The results are robust to a wide range of specification checks. First, the effects are stable across lag structures ranging from five to twenty-five years of pre-crisis growth history. Second, the results are similar when we constrain the data to a balanced panel that requires a country to be present in all the years of a given estimation. Third, conditioning on contemporaneous wars, currency crises, and sovereign debt crises leaves the impulse responses essentially unchanged, as does the inclusion of a broad set of macroeconomic controls including inflation, interest rates, credit aggregates, investment, and government debt. Fourth, excluding crises associated with the two World Wars and the Great Depression does not attenuate the findings. Finally, the results are robust to using the underlying crisis chronologies individually, as well as alternative union-based indicators.

We also subject the results to a range of alternative econometric approaches. A placebo exercise that preserves the overall frequency of crises but randomizes their occurrence generates impulse responses centered near zero at all horizons, indicating that the estimated losses are not an artifact of underlying output persistence or slow-moving macroeconomic trends. The results are also robust to estimating local projections in levels rather than cumulative growth rates, and to estimating the dynamic response using the Pesaran and Smith (1995) mean group estimator, which allows crisis effects to vary freely across countries.

Having established the magnitude and robustness of these long-run output losses, in the last section of the paper, we turn to the mechanisms underlying their persistence. Our approach is not to isolate a single structural channel, but rather to document a set of empirical patterns that collectively point to the role of financial constraints that interrupt capital accumulation and increase debt burdens in generating long-run economic scarring.

We begin by examining the recurrence of crises themselves. We find that financial crises are followed by elevated probabilities of subsequent crises, consistent with the view that crises can weaken financial systems and increase future vulnerability. However, these dynamics do not account for the persistence of output losses: even conditioning on the full path of crises, the long-run effects remain large. This suggests that persistence reflects more than repeated crisis realizations, and instead points to lasting damage from an initial shock.

We next turn to physical and human capital accumulation. The share of output devoted to investment falls sharply following a crisis and remains depressed for decades, while measures of human capital exhibit similarly persistent declines. This sustained contraction in the rate of capital formation, paired with the erosion of human capital, points to a prolonged disruption of the processes

that underpin long-run growth, consistent with binding financial constraints that limit the ability of firms and households to undertake forward-looking investments.

These costs are evident on government balance sheets as well. The revenue-to-GDP ratio falls persistently following a crisis, indicating that fiscal capacity erodes faster than the economy contracts. Combined with a short-term rise in government spending, this generates persistent deficits and a sustained build-up in debt. While revenue shares remain persistently lower, expenditures and debt burdens both eventually decline, suggesting that governments effectively shrink following these disruptions.

Taken together, these patterns indicate that the persistence of output losses coincides with prolonged shortfalls in investment and human capital and a lasting erosion of fiscal capacity, consistent with financial constraints that slow the rebuilding of productive capacity after a crisis. The long-run costs of banking crises are therefore substantially larger than previous analyses using shorter horizons suggest, underscoring the value of preventing crises and repairing balance sheets thereafter.

*Related Literature.*

This paper relates to a large literature documenting the macroeconomic consequences of financial crises. A central finding in this literature is that banking crises are associated with substantial declines in economic activity. Using long-run historical data, Reinhart and Rogoff (2009*b*), Barro (2006), and Barro and Ursúa (2008*a*) document that financial crises and other rare disasters coincide with sharp contractions in output and consumption. More recent work shows that credit booms often precede financial crises and that these crises are followed by prolonged recessions and slow recoveries (Schularick and Taylor, 2012; Jordà, Schularick and Taylor, 2013). Similarly, Cerra and Saxena (2008) show that output losses following financial crises are highly persistent relative to other shocks such as wars.

A related literature examines the broader consequences of financial crises across a range of economic and political outcomes. Crises have been shown to generate persistent effects on potential output and demand dynamics (Blanchard, Cerutti and Summers, 2015; Cerra, Fatás and Saxena, 2023), as well as on international trade patterns (Xu, 2022). Beyond real economic activity, financial crises also have important political consequences: Funke, Schularick and Trebesch (2016) document that they are followed by sharp increases in political extremism, while Guiso, Herrera, Morelli and Sonno (2017) and Quincy and Xu (2026) show that financial and economic distress contributes to rising support for populist movements. While these contributions establish that financial crises have wide-ranging effects, with the exception of Xu (2022), the empirical evidence has largely focused on horizons of up to a decade and feature little discussion of the mechanisms for persistence.

This paper also relates to a large body of work emphasizing the role of financial frictions in amplifying macroeconomic fluctuations. Models of the financial accelerator and collateral constraints

show how disruptions in credit markets can lead to large declines in real activity (Bernanke and Gertler, 1989; Kiyotaki and Moore, 1997). Consistent with these mechanisms, empirical evidence shows that financial conditions, measured for instance with credit spreads, strongly amplify business cycle fluctuations, particularly during downturns (Krishnamurthy and Muir, 2025). Other work highlights the importance of the composition of credit in shaping macroeconomic outcomes. For example, expansions in household credit and lending to financially constrained sectors are associated with more severe downturns and a higher likelihood of financial crises (Mian, Sufi and Verner, 2017; Müller and Verner, 2024). While this literature provides a compelling account of how financial disruptions generate large short- and medium-run effects, it offers more limited guidance on why the effects of financial crises remain so persistent over multi-decade horizons.

Finally, our analysis relates to a broader literature on recovery from large macroeconomic shocks. A growing body of work emphasizes that post-crisis dynamics depend critically on the ability of economies to stabilize and rebuild after the initial contraction (e.g., Romer and Romer, 2019). Our findings complement this perspective by showing that the persistence is systematically related to the disruption of investment and the erosion of fiscal capacity.

## 2 The Global Macro Database: An Introduction

Estimating the consequences of macroeconomic shocks such as financial crises over decades requires a broad cross-section of countries and long-run historical time series. This section introduces the Global Macro Database (GMD), a new harmonized, comprehensive dataset covering 46 key macroeconomic indicators, as well as additional derived values. By integrating 33 continuously updated sources with 127 historical datasets, the GMD achieves near-universal coverage of the world economy, spanning 239 economies from 1086 to 2025, including forecasts until 2030.

The construction of the GMD is motivated by a fundamental limitation of existing macroeconomic datasets. Most widely-used sources either offer broad cross-country coverage over relatively short post-Bretton Woods samples, or long historical series for a narrow set of economies. This limitation is compounded by the practical challenges involved in combining historical and modern data sources, which differ in units of account, territorial definitions, statistical methodologies, and the presence of errors, making the construction of consistent long-run series a nontrivial exercise. As a result, empirical work on the long-run consequences of macroeconomic disruptions has typically relied on small samples, short horizons, or highly selective country coverage.

The GMD is a departure from previous data efforts in several dimensions. Three deserve particular mention. First, we conduct an extensive harmonization and cleaning of raw data sources. Users can thus access 160 sources in one panel format with the same units, currency, country identifiers, and definitions (as possible). Second, we provide harmonized splice series that are ready-to-use,

based on a splicing pipeline which orders sources by their overall level of reliability. Third, we continuously update the data. The time series thus not only span centuries of data, but are also informative about recent events.

We provide extensive documentation of our approach and all sources in a technical appendix accompanying every quarterly release; the current version is included in Appendix A. In addition to this appendix document, we also provide a detailed description of the adjustments made to the raw data and time series plots for each variable and country, spanning 7,000+ pages, which shows how we arrive at our final estimates. This extensive documentation is available through www.globalmacrodata.com/documentation. We are continuously refining and extending the database, and users are encouraged to report issues on our GitHub page.

## 2.1 Coverage and Scope

The Global Macro Database (GMD) provides annual macroeconomic data for 239 countries and territories, with coverage extending as far back as the medieval period for some variables and countries, and continuing through the present. It is updated quarterly. The focus of the database are 46 macroeconomic indicators that allow us to understand the trajectory of output after crises, but also various potential mechanisms.

Appendix Table A.1 provides an overview of the 80 macroeconomic variables included in the GMD and their coverage in our dataset. These variables encompass the key macroeconomic indicators most widely used in applied research: Gross Domestic Product (GDP, nominal and real) and other national accounts variables, such as investment, consumption, government expenditure, exports, and imports; inflation, measured by both the consumer price index (CPI) and the GDP deflator; population; the current account; exchange rate measures, including the US dollar exchange rate and the real effective exchange rate; government finances, including government revenues, tax revenues, government deficits, and government debt; interest rates, including short-term and long-term interest rates and the central bank policy rate; measures of the money supply (M0, M1, M2, M3, and M4); unemployment; and house prices.

## 2.2 Sources

We draw on a total of 160 sources. A full list is in Appendix Table A.5. Appendix Table A.4 shows that these sources can be divided along two dimensions. The first dimension is the number of covered countries. We use data from 76 "aggregators", such as the World Bank, IMF, or OECD, or PWT, which provide time series for more than one country. However, not all of these aggregators are international organizations. Jordá, Schularick and Taylor (2017*b*) or Barro and Ursúa (2008*b*,*c*) are examples of academic work that compile long-run macroeconomic time series for a panel of

countries. We also draw on "country-specific" sources, such as the Statistical Bureau of Taiwan or the historical time series on the United Kingdom compiled by Thomas, Hills and Dimsdale (2010).

The second dimension is whether the data is "current", meaning continuously being updated, or "historical", meaning it is not. International organizations almost always "current." In contrast, "historical" data sources are only cleaned and incorporated into the dataset without any expectation they will be updated. Most work by economic historians falls into this category.

Overall, around two thirds of our 160 sources are of the country-specific, historical type. It is exactly these data that have been neglected in applied work, given the effort required to hunt them down and combine them with more recent time series. As such, we view one of our contributions as bringing the value provided by economic historians back into the realm of broader economics research.

Different from existing work, we do not only use these sources as an input, but also retain the original raw data (with harmonized units, currency, and so on). Every single adjustment (e.g., identified errors or unit issues) are stored as a comment on the observation-level. The starting point of our data construction is thus a code base that cleans each of the 160 underlying sources. As we outline in the next section, this process is far from trivial, given many issues even in the most widely-used databases.

### 2.3 Data Construction

**Reporting conventions.** We report all time series in the national currency currently used in each country. Historical sources may report values in currencies that are no longer in use, so currency reforms sometimes require us to convert earlier values. For example, for Eurozone countries, we report all values in euros and convert historical currencies using the irrevocable exchange rates set when the euro was introduced. Ratios and rates are reported in percent, such that a government debt to GDP ratio of "5" means 5 percent.

Real GDP in USD deserves particular attention, because it can be constructed in a number of ways. We express real GDP per capita in constant 2015 U.S. dollars following the methodology used by the World Bank (2026). We first convert the 2015 level of GDP per capita into U.S. dollars using the 2015 market exchange rate. We then extrapolate this benchmark backward and forward using growth rates of real GDP per capita measured in local currency. The resulting series can be interpreted as GDP per capita valued at the country's 2015 market exchange rate, with changes over time reflecting real growth rather than subsequent exchange rate movements.

**Combining data sources.** There are two principal benefits to combining data series. First, it allows us to generate a dataset with much larger coverage than existing work. In fact, the coverage

of even the most widely-used variables differs considerably across datasets. The World Bank's World Development Indicators, for example, only start in 1960, an arbitrary cut-off year given that much of the underlying macroeconomic data they compile is from the United Nations or International Monetary Fund, both of which report data from earlier years.

Second, combining data sources allows us to benchmark and compare different estimates. As we discuss below, this has the critical advantage of assessing whether a particular data point is plausible or not, and whether there are potential systematic issues with the reporting in a source. In many cases, such issues would not be identifiable without systematically looking at several sources at once.

After cleaning the raw data, we combine the time series from different sources in a way that preserves the growth dynamics of each source while anchoring levels to the most reliable contemporary benchmarks. This approach preserves growth rate variation over absolute historical levels and avoids artificial breaks that would arise from simply concatenating disparate sources.

We implement this principle through a systematic ratio-splicing procedure. For each country and variable, we establish a benchmark series from the highest-quality modern source, typically the IMF World Economic Outlook. We then extend this series backwards by sequentially linking earlier historical sources, which preserves their growth rates while adjusting their levels to ensure continuity at splice points.

Formally, consider a benchmark series $y_t^{(a)}$ and an earlier historical series $y_t^{(b)}$ that overlap in year $\tau$. We construct the spliced series as:

$$\tilde{y}_t^{(b)} = y_t^{(b)} \times \frac{y_\tau^{(a)}}{y_\tau^{(b)}} \quad \text{for } t < \tau$$

The final series uses $\tilde{y}_t^{(b)}$ for $t < \tau$ and $y_t^{(a)}$ for $t \geq \tau$. This procedure maintains the growth rates from the historical source—which typically reflect careful work by economic historians—while eliminating level discontinuities that would otherwise arise from differences in base years or statistical methodology.

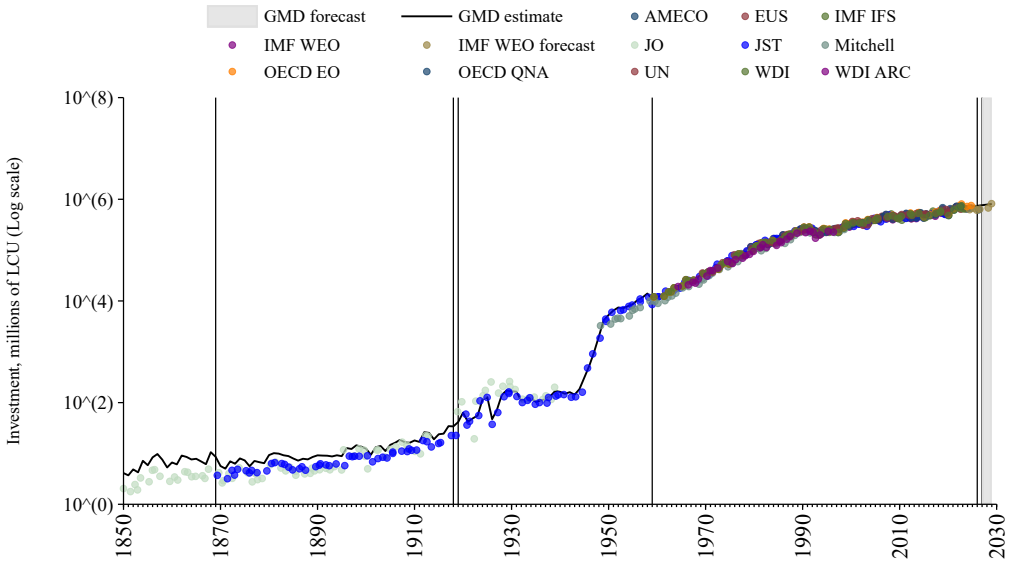
When multiple historical sources are available for a given period, we apply this procedure sequentially, working backwards from the benchmark source and prioritizing sources with broader temporal coverage and more comprehensive documentation. In the small number of cases in which adjacent sources do not overlap, we link them using the median growth rate for the three years before and after a source change. These instances are explicitly flagged in our documentation.

We do not apply the above splicing procedure to variables expressed as rates, such as exchange rates or interest rates. For these indicators, we simply combine the underlying raw data without any adjustments.

To illustrate our approach, Figure 2 plots data for investment in France for the period 1850 to

2030, including forecasts to 2029. We have data from a total of eleven sources, but five of them are sufficient to construct the full estimated series. While some sources differ in their reported "level" of investment, the trends are highly similar. We start with the IMF's WEO data and then splice the series backwards using the trends in other historical sources.

Figure 2: Example of Source Harmonization: Investment in France




To ensure full transparency and replicability, we provide a comprehensive documentation with exactly this type of exhibit for every single country and variable. These plots can also be explored on the GMD website. All code and the underlying raw data used to construct the database is available in our public GitHub repository. This structure allows researchers to trace each series back to its original sources and reproduce or modify the database for their own purposes. The GMD is therefore designed not only as a dataset for the present analysis, but as a flexible and extensible research infrastructure for future empirical work on long-run macroeconomic dynamics.

**Fixing common issues in widely-used datasets.** Integrating and harmonizing many sources allows us to spot many errors in the raw data. A non-exhaustive list of types of errors include:

- **Wrong or inconsistent units.** Even major data provider such as the IMF publish data with undocumented unit errors and inconsistencies. For example, billions may be mislabelled as millions, or values may be reported in different units across countries, which could lead to erroneous conclusions.
- **Incorrect or missing values.** Raw data may contain incorrect values, including missing observations recorded as zero or typographical errors introduced during digitization. We

correct obvious errors and set clearly erroneous values to missing when no reliable correction can be identified.

- **Inconsistent currencies** Historical series may contain undocumented breaks when currency reforms are not consistently incorporated, causing artificial jumps in the data. For example, the IMF's International Financial Statistics adjust Brazil's series for several currency changes but omit the 1993 Plano Real stabilization, leaving a large discontinuity from incorrectly converted currency units.

- **Missing data across versions.** Current releases of major databases such as the World Bank's WDI and the IMF's IFS may omit observations that appeared in earlier vintages. In some cases, values missing from the current online database remain available only in archival releases.

- **Incomplete digitization.** Some major publications, such as the OECD Key Economic Indicators (KEI), have only been incompletely digitized. In the case of the KEI in particular, we obtained historical archival versions digitized them to fill gaps in the data available online. We are in the process of digitizing the full archives of several other major publications, including the IMF's International Financial Statistics.

- **Changing base years.** Historical real GDP and price index series sometimes change base years without reporting overlapping observations, creating breaks for which the true growth rate cannot be directly calculated. This issue arises, for example, in some series reported by Mitchell (2013), although its practical importance is limited because these data receive the lowest priority in our source hierarchy.

A major contribution of the GMD is that it fixes such errors and inconsistencies, which frees up researchers' time to focus on their analysis.

**Internal consistency.** Our dataset contains both derived variables, which are those that can be constructed or calculated from their raw components, as well as the raw variables themselves. For example, real GDP per capita is a derived variable that is calculated from data on real GDP and population. Our sources do not always contain all the underlying variables for each derived variable, and sometimes the derived data within a given source differs from the values one would get from a manual computation. These inconsistencies prompt the following questions. Do we want a dataset that is internally consistent, such that real GDP per capita is equal to real GDP divided by population, even if it comes at the cost of our data at times disagreeing with the estimates of real GDP per capita from existing sources? Or do we care less about internal consistency and more

about time series for real GDP, population, and real GDP per capita that are as closely aligned as possible with existing work?

We take a hybrid approach. Whenever possible, we provide internally consistent derived measures for variables such as real GDP per capita and government debt-to-GDP ratios. When these differ from more credible estimates reported by existing sources, we additionally provide those series even if they are not internally consistent with the underlying components.

A specific version of this issue is that, in principle, we report time series for both GDP and all of its components by expenditure, such that $Y = C+I+G+(X-M)$. However, it is often the case that one of several components of GDP is missing, or that the components come from different sources for the same time period. For example, historical data on nominal GDP almost never includes data on components, and estimates of government spending ($G$) and net exports ($X - M$) usually come from entirely separate sources. To avoid imputing implausible values, the current version does not enforce the constraint that derived measures are internally consistent, although it is something we plan to improve on in future releases.

### 2.4 Access, Updates, and Revisions

The GMD can be accessed directly on www.globalmacrodata.com/data or, alternatively, using packages for Stata, R, and Python, which are also described on the website.[1] The full data and documentation are available on www.globalmacrodata.com/documentation. Most users will be interested in our ready-to-use series, which are continuously updated as new data arrive.

Some users will also want to look at the underlying raw data and additional documentation. For that purpose, we provide an extended version of the dataset that not only includes our estimates of each variable but also all underlying variables and meta data, which allow users to assess the raw data and adjustments we made.

Unlike seminal contributions such as the Macro History Database (Jordá, Schularick and Taylor, 2017*b*) or the Penn World Tables (Feenstra, Inklaar and Timmer, 2015), our data are continuously updated on a quarterly basis. The current, 9$^{\text{th}}$ release is 2026-06, and future releases are pre-announced on the GMD website. All versions of the database are archived and publicly accessible, allowing researchers to replicate earlier results or assess the sensitivity of findings to data revisions.

The frequent updating require us to pay special attention to data revisions when splicing time series. Take the case of a new GDP release in the IMF's World Economic Outlook that happens to be much higher than previous estimates. If we naively spliced together historical values back to 1850, such a revision would not only change recent values, but also a data points in the 19$^{\text{th}}$ century.

[1] For example, the user-written Stata command `gmd`, introduced in Lehbib and Müller (2025), can be installed with `ssc install gmd`. The command `gmd rGDP` downloads all time series for real GDP in national currency, while `gmd rGDP, country(FRA)` restricts the download to France.

We address this challenge by freezing historical values while allowing contemporary data to incorporate ongoing revisions. Specifically, we establish 2018 as a reference year: values prior to 2018 remain fixed unless a clear error is discovered or demonstrably superior historical evidence becomes available, while values from 2018 onward update with each new release from contemporary sources. This approach ensures that the historical portion of our series remain stable, while recent data incorporate the latest available information. Routine updates thus affect only the most recent years and do not affect earlier periods.

### 2.5 Comparison with Existing Sources

The database includes a broad set of 46 macroeconomic variables spanning real activity, public finances, money and interest rates, labor markets, and trade. In addition to real GDP, key variables used in our analysis include government revenues and expenditures, trade, public debt, investment, and unemployment. The breadth of coverage varies across variables, but the GMD always substantially extends both the temporal depth and cross-country coverage available in existing datasets.

To illustrate this, Appendix Figure A.4 compares the coverage of the GMD to the next-most comprehensive alternative, showing that the database provides a larger number of country-year observations and significantly longer time series for all variables. A complete list of variables, sources, and coverage is provided in the Technical Appendix.

### 2.6 Banking Crisis Onset Indicator

Our analysis examines the long-run effects of banking crises using an onset indicator that consolidates the principal crisis chronologies in the literature: Reinhart and Rogoff (2009*b*), Laeven and Valencia (2013), Jordà, Schularick and Taylor (2017*a*), and Baron, Verner and Xiong (2021). These sources differ in definitions, empirical criteria, and coverage, but collectively provide the most comprehensive record of historical banking crises.[2]

The construction proceeds in two steps. We first identify candidate crisis episodes from the source datasets and group records that refer to the same underlying event. We then assign each episode an onset year using a source-support rule: when sources report nearby dates for the same episode, the indicator favors the year supported by more chronologies, with ties assigned to the earlier date. For sources that code every year of an ongoing crisis, we retain the first year of the episode. This procedure converts the source chronologies into a single event-level series while preserving the timing information contained in the underlying datasets.

The resulting indicator, $BankingCrisis_{i,t}$, equals one in the onset year of each crisis episode

[2]There are important predecessors of these efforts, such as Kaminsky and Reinhart (1999) or Bordo, Eichengreen, Klingebiel and Martinez-Peria (2001), but their coverage is superseded by the four sources we consider.

and zero otherwise. A country-year is coded as zero only when at least one underlying chronology explicitly covers it as a non-crisis observation, so periods outside source coverage remain missing. In the final sample, approximately 1.8% of country-year observations contain a banking crisis. Across countries, a new crisis occurs approximately every 35 years. The indicator contains 362 episodes across 127 countries; the construction is described in greater detail in Appendix A.8.1, and the robustness exercises below compare the baseline indicator with alternative source chronologies.

Table 1: Summary Statistics

| **Panel A: Continuous Variables** | | | | | | |
|---|---|---|---|---|---|---|
| | Mean | SD | p25 | Median | p75 | N |
| Real GDP (billions of 2015 USD) | 176.15 | 924.11 | 2.67 | 12.43 | 68.63 | 17,950 |
| Government Debt (% of GDP) | 51.89 | 50.78 | 20.85 | 39.86 | 67.20 | 13,702 |
| Government Deficit (% of GDP) | -2.65 | 11.80 | -4.40 | -1.80 | 0.07 | 12,757 |
| Government Revenues (% of GDP) | 22.14 | 15.74 | 11.18 | 18.61 | 29.77 | 13,491 |
| Government Expenditure (% of GDP) | 24.62 | 18.22 | 13.03 | 21.41 | 32.73 | 13,611 |
| Investment (% of GDP) | 22.96 | 10.72 | 16.10 | 21.97 | 27.73 | 15,005 |
| Unemployment Rate (%) | 7.92 | 6.51 | 3.52 | 6.20 | 10.14 | 7,199 |
| **Panel B: Crises Variables** | | | | | | |
| | Mean | SD | p25 | Median | p75 | N |
| Banking Crisis (country × year) | 0.0201 | 0.1402 | 0.0000 | 0.0000 | 0.0000 | 19,933 |
| mean by country | 0.0189 | | | | | 163 |
| mean by year | 0.0181 | | | | | 221 |
| Sovereign Debt Crisis (country × year) | 0.0146 | 0.1199 | 0.0000 | 0.0000 | 0.0000 | 19,069 |
| mean by country | 0.0139 | | | | | 160 |
| mean by year | 0.0138 | | | | | 221 |
| Currency Crisis (country × year) | 0.0279 | 0.1647 | 0.0000 | 0.0000 | 0.0000 | 19,596 |
| mean by country | 0.0280 | | | | | 160 |
| mean by year | 0.0265 | | | | | 220 |
| **Panel C: Coverage** | | | | | | |
| | Mean | p25 | Median | p75 | Min | Max |
| Years per country | 285 | 225 | 225 | 225 | 6 | 939 |
| Countries per year | 195 | 210 | 213 | 229 | 1 | 235 |

*Notes:* This table reports summary statistics for the full sample. Panel A presents continuous variables, including real GDP (in billions of USD at constant 2015 prices), fiscal variables (as % of GDP), and the unemployment rate (in percent). Panel B reports indicator variables in percent. Indicator variables equal one in years in which the corresponding event occurs. "Mean by country" denotes the within-country frequency of the event, while "mean by year" denotes the cross-sectional share of countries experiencing the event in a given year.

Table 1 reports summary statistics for the full sample. Panel A summarizes the main macroeconomic variables used in the analysis. The data span a broad set of countries and time periods, generating substantial variation in output levels, fiscal outcomes, and labor market conditions. This dispersion reflects the wide cross-country coverage and long historical horizon of the dataset.

Panel B reports summary statistics for crisis indicators. The mean of each indicator corresponds to the share of country-year observations in which a crisis occurs. The rows labeled "mean by country" report the average frequency with which a given country experiences the event over time,

while the rows labeled "mean by year" report the cross-sectional share of countries experiencing the event in a given year. These statistics highlight that crises are infrequent in any given country-year but recur over time within countries and, at times, affect multiple countries simultaneously.

### 2.7 Additional Data

**Education and human capital.** We measure human capital using educational attainment data from Barro and Lee (2013), which provides internationally comparable measures of schooling for the population aged 15 and above at five-year intervals beginning in 1950. We combine these with the data from Lee and Lee (2016), who extend them backwards to 1820.

**Conflict and geopolitical shocks.** We incorporate data on interstate, extrastate and civil conflicts from the Correlates of War (COW) project (Sarkees and Wayman, 2010).

## 3 The Very Long-Run Output Costs of Financial Crises

### 3.1 Methodology

We estimate the dynamics of real economic activity following financial crises using impulse response functions (IRFs). These IRFs can be estimated using Vector Auto Regressions (Sims, 1980, VARs,) or Local Projections (Jordà, 2005, LPs). Different from existing work, we are interested in forecast horizons of up to $h = 50$ years, which raises a number of econometric issues.

The traditional view is that there is a bias-variance trade-off when estimating IRFs using VARs or LPs. As the forecast horizon $h$ increases, VARs produce increasingly biased estimates if they use insufficient lags (see, e.g., Li, Plagborg-Møller and Wolf, 2024; Montiel Olea, Plagborg-Møller, Qian and Wolf, 2025). The resulting recommendation for estimating long-run IRFs with VARs is thus to choose a lag length $\ell \geq h$ (e.g., Baumeister, 2025). For example, Antolin-Diaz and Surico (2025) effectively set $p = \ell$ to estimate the long-run effects of government spending.

However, recent work by Ludwig (2024) shows that the IRFs of VARs and LPs are in fact the same, in sample, in that LPs are VAR forecasts of increasing lag order; also see Plagborg-Møller and Wolf (2021). For a given lag length $L$, LPs estimate a more highly parameterized model, which explains why their IRFs exhibit greater variance as $h$ increases. Given that the main concern for the estimation of long-run effects is bias from a too parsimonious VAR model, rather than variance, our baseline specification follows Jordà (2005) and estimates LPs.

Concretely, we use panel LPs with the following specification:

$$\Delta \log Y_{i,t+h} = \beta_h \, \text{BankingCrisis}_{i,t} + \alpha_i^{(h)} + \sum_{\ell=1}^{L} \phi_\ell^{(h)} \Delta \log Y_{i,t-\ell} + \varepsilon_{i,t}^{(h)}, \tag{1}$$

in which $i$ indexes countries and $t$ indexes years. The crisis indicator, $\text{BankingCrisis}_{i,t}$, equals one in the first year of a crisis episode, as defined in Section 2. The dependent variable $\Delta \log Y_{i,t+h} \equiv \log Y_{i,t+h} - \log Y_{i,t-1}$ measures the cumulative change in real GDP from the year preceding the crisis through horizon $h$, so that the coefficients $\beta_h$ can be interpreted as impulse responses relative to the pre-crisis trajectory. Country fixed effects $\alpha_i^{(h)}$ absorb time-invariant differences across countries, such as long-run income levels. Standard errors are heteroskedasticity-robust and clustered at the country level.

We view our baseline specification as highly conservative for three reasons. First, because we estimate IRFs using LPs, bias from an insufficient number of lags is somewhat less of a concern than when using VARs. Second, we include a large number of lags and present the results for $L = 5, 10, 25$ whenever the sample size allows. As such, we take the recommendations of the VAR literature seriously that, whenever possible, $\ell \geq h$. Third, we estimate LPs in long-differences rather than levels, which requires longer time series but removes bias (Jorda and Taylor, 2025; Piger and Stockwell, 2025). The dependent variable, mainly log real GDP, is thus always defined as the difference between its realization at time $t + h$ relative to $t - 1$.

The combination of long forecast horizons, LPs rather than VARs, a large number of lags, and estimation in differences rather than levels places considerable demands on the data. The Global Macro Database (GMD) is uniquely suited for meeting these demands because it provides both broad country coverage coupled and long historical time series, and thus sufficient statistical power.

**Interpretation.** The estimated coefficients $\{\beta_h\}_{h=0}^{50}$ trace out the average path of real GDP following a banking crisis relative to the counterfactual of no crisis, conditional on the long pre-crisis output history. Interpreting these coefficients as causal effects requires the identifying assumption that, after conditioning on long lags of output growth and country-specific intercepts, the timing of crisis onset is orthogonal to unobserved determinants of future output at horizons $h = 1, \ldots, 50$. We show below that output dynamics in the twenty years prior to the onset of a crisis display no systematic pre-trends once we condition on lagged output growth, which supports the validity of this assumption. However, we do not claim that crises are exogenous events, even with these extensive controls. Rather, we interpret the estimates as capturing the persistence of output losses associated with financial crises.

### 3.2 Baseline

The baseline estimation of equation (1) uses the full panel of available data to maximize statistical power. Because the panel is unbalanced, the set of crisis observations varies across horizons as observations far into the future become unavailable for more recent crises. As a result, estimates at

longer horizons place greater weight on earlier historical episodes.

To assess whether this feature affects the estimated long-run responses, we also report results from a balanced panel that restricts the sample to crisis episodes with at least fifty years of post-crisis data, in addition to the years required to construct the lagged control variables. This restriction ensures that the composition of crisis episodes remains constant across horizons, allowing for a direct comparison of short- and long-run dynamics within the same sample. Appendix Figure B.2 illustrates the number of crises by decade in each sample; by construction, crises occurring after the 1970s are dropped from the balanced panel.

Figure 3 presents our baseline estimates of the impulse responses of cumulative real GDP following a banking crisis. The horizontal axis plots years since the onset of a crisis, while the vertical axis reports cumulative percent deviations of real GDP relative to its pre-crisis level. The dark solid line corresponds to the baseline specification with 25 lags of pre-crisis growth, and the shaded gray region denotes the associated 90 percent confidence interval. The lighter dashed lines report alternative specifications with shorter lag lengths ($L = 10$ and $L = 5$), with confidence intervals shown selectively to preserve visual clarity.

Conditioning on longer histories of pre-crisis growth modestly increases the magnitude of estimated losses at all horizons. One interpretation is that fewer lags are insufficient to account for pre-crisis trends in economic growth, and thus incorrectly understate the magnitude of the subsequent downturn.

Figure 3b reports results from the balanced panel. The estimated losses are larger in magnitude, with trough declines of approximately 15 percent relative to the pre-crisis trajectory in the two to three decades following the onset of the crisis, but the overall shape of the impulse response is very similar to that in the full panel. Conditioning on episodes with at least fifty years of post-crisis data places greater weight on earlier historical crises and excludes recent episodes whose long-run outcomes are not yet observed. The qualitative result—large and highly persistent output losses following banking crises—is common to both samples, indicating that it is not an artifact of horizon-specific changes in sample composition.

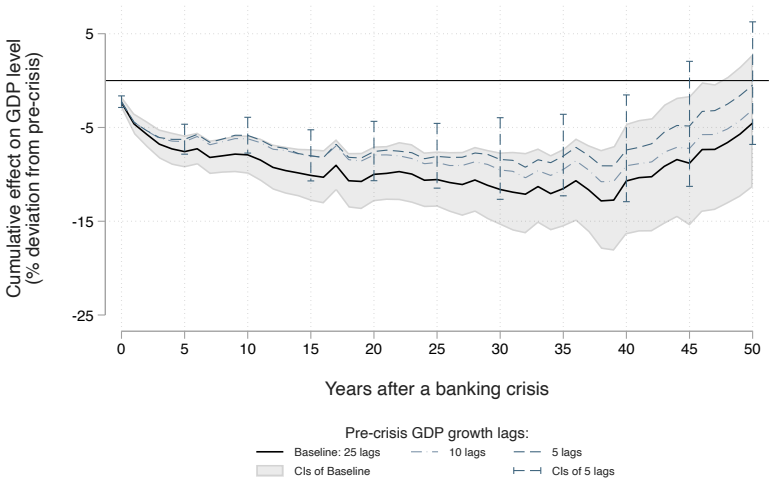
### 3.3 Robustness

We provide several robustness checks of the baseline results. Across all exercises, the central finding of large and persistent output losses following financial crises remains unchanged.

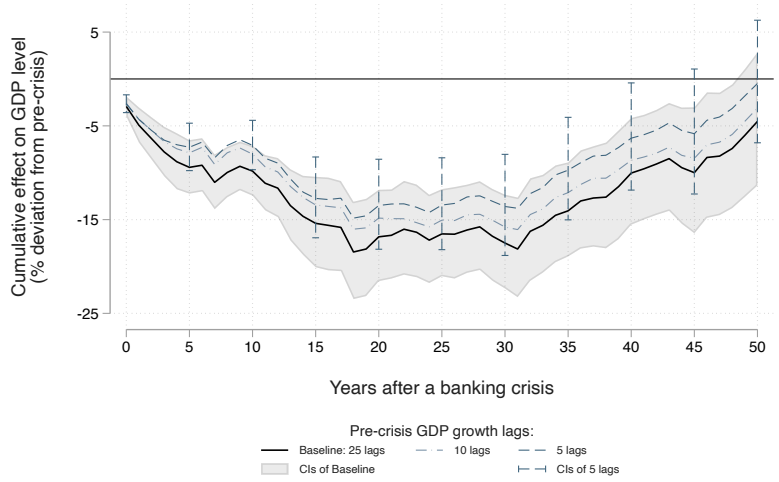
In all figures, we present the baseline specification using the full panel and conditioning on 25 lags of pre-crisis output growth. We show these baseline results as a dark line, with shaded gray bands denoting the corresponding confidence intervals. Alternative specifications are reported as thinner lines without confidence intervals, and in selected cases we report confidence intervals at

Figure 3: Long-Run Impact of Banking Crises

(a) Full Panel




(b) Balanced Panel




*Notes:* This figure reports local projection estimates of the cumulative response of real GDP to a banking crisis, relative to the year prior to the crisis. The horizontal axis denotes years since the crisis onset ($h = 0$), and the vertical axis plots cumulative percent changes in real GDP relative to the pre-crisis level. Estimates are obtained from equation (1). Panel (a) uses the full panel of available data, allowing the set of crisis observations to vary across horizons. Panel (b) restricts attention to a balanced panel of crises for which the full set of leads and lags is available. Each line corresponds to a specification that conditions on $L$ lags of pre-crisis GDP growth, with $L \in 5, 10, 25$. Shaded areas denote 90 percent confidence intervals for $L = 25$ while vertical CIs are shown every 5 years for $L = 5$. Standard errors are clustered by country. See Appendix Figures B.1 for an alternative lag structure ($L = 50$).

five-year intervals for specifications that include the full set of controls. Appendix B reports the full set of results across alternative lag structures and sample restrictions.

**Controlling for contemporaneous shocks and macroeconomic conditions.** Financial crises often coincide with other major macroeconomic disruptions, raising the concern that the estimated output losses attributed to banking crises may partly reflect the effects of concurrent shocks (e.g.,

Kaminsky and Reinhart, 1999; Laeven and Valencia, 2013). In particular, wars and other forms of macro-financial distress such as currency crises and sovereign debt crises are frequently clustered in time and may themselves have large and persistent effects on real economic activity. If such shocks are correlated with crisis timing, failing to account for them could bias the estimated dynamic response to banking crises.

To address this concern, we augment the baseline local projections specification by controlling for additional contemporaneous shocks at the crisis onset year. Specifically, we first include an indicator for wars in year $t$, and then expand the conditioning set to include indicators for currency crises and sovereign debt crises, as defined in Section 2. These controls enter contemporaneously alongside the banking crisis indicator.

Figure 4a reports the resulting impulse responses from controlling for contemporaneous wars and other types of crises while Figure 4b instead examines robustness to conditioning on a broader set of macroeconomic variables, including inflation, interest rates, broad money, investment, the current account balance, government debt, and the real effective exchange rate. In each panel, the dark line corresponds to the baseline specification, while thinner lines report estimates from specifications that add each control individually. Confidence intervals for the specification including all controls are reported at five-year intervals. Across these specifications, the estimated output dynamics following banking crises remain highly stable.

**Excluding major global disruptions.** While the preceding results address the role of contemporaneous shocks and macroeconomic conditions, a related concern is that long-run estimates may be driven by a small number of extreme global events. In particular, the two World Wars and the Great Depression represent tail events of economic disruption that affected a large share of countries simultaneously.

To address this concern, we re-estimate the baseline specification after excluding crisis episodes associated with these major global disruptions. We conduct three exercises. First, we exclude crises occurring during World War I (1914–1919) and World War II (1939–1944). Second, we exclude crises corresponding to the Great Depression period (1929–1939). Third, we exclude all of these episodes simultaneously. In each case, the sample retains the full panel of countries, but the crisis indicator during these periods is set to zero.

Figure 4c reports the corresponding impulse responses. Across all exercises, the estimated dynamics closely resemble the baseline results, indicating that the average effects we find do not depend on these global shocks. If anything, the estimated output losses are more persistent when these episodes are excluded, likely because output growth after the wars and the Great Depression was fairly strong.

**Alternative crisis chronologies.** A final concern relates to the dating of banking crises. Our baseline crisis indicator consolidates information across the canonical banking crisis datasets used in the literature and resolves differences in timing and episode definitions through the source-support rule described in Section 2. It is natural to ask whether the results are sensitive to this consolidation procedure.

To assess this, we re-estimate equation (1) using banking crisis indicators taken directly from selected underlying source datasets, replacing $BankingCrisis_{i,t}$ with the source-specific crisis onset indicator while holding fixed the outcome definition, control set, lag structure, and sample restrictions. This exercise isolates the sensitivity of the estimated impulse responses to alternative crisis dating conventions and episode definitions.

Figure 4d shows that the estimated output response is highly robust across crisis chronologies. Regardless of whether crises are dated using Reinhart and Rogoff (2009*b*), Jordà, Schularick and Taylor (2017*a*), Baron, Verner and Xiong (2021), or their union, banking crises are followed by large and persistent output losses that remain substantial for multiple decades.

Taken together, these robustness exercises show that the estimated long-run output losses following banking crises are not driven by contemporaneous macroeconomic shocks, extreme global events, or alternative crisis dating conventions. Across a wide range of alternative specifications, samples, and definitions, the impulse responses remain remarkably stable.

### 3.4 Placebo Analysis and Alternative Estimators

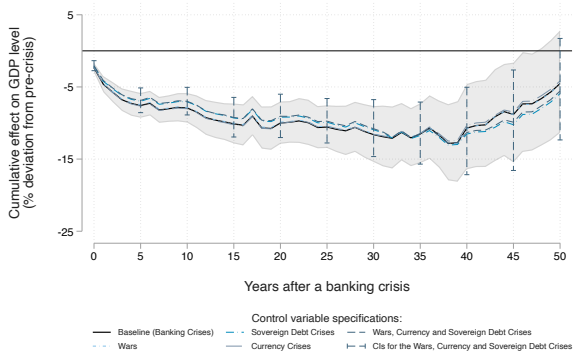
To further assess whether our results depend on the precise specification of the baseline local projections, we conduct several additional sets of exercises. First, we implement a placebo test that randomizes the timing of crises while preserving their overall frequency in the data. Second, we compare our baseline estimates to several alternative estimators and outcome definitions, including local projections in levels and long differences, and the mean group estimator.

**Placebo exercise.** A potential concern with long-horizon local projections is that persistent output dynamics may mechanically generate spurious treatment effects when shocks occur at particular points along slow-moving macroeconomic trends. While our long lag structure is designed to absorb such dynamics, we also conduct a placebo exercise based on simulated crisis timing.

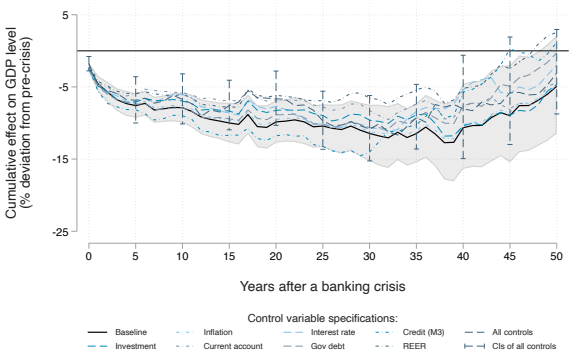
Specifically, we generate placebo crisis indicators that preserve the overall frequency of crises observed in the data but randomize their occurrence across country-year observations. For each simulation, the onset of a crisis is drawn independently so that the unconditional probability of a crisis matches the sample mean of the banking crisis indicator. We then re-estimate the baseline local projections specification in equation (1) using each simulated crisis series. Repeating this procedure

Figure 4: Robustness

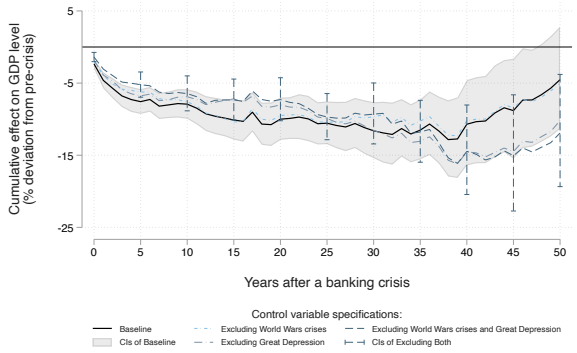
(a) Wars, Conflicts, and Other Crises



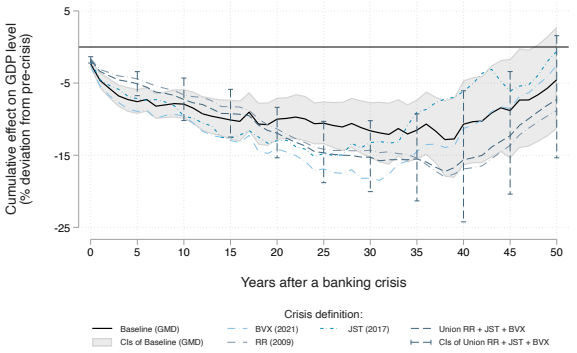

(b) Additional Macroeconomic Conditions




(c) Excluding World War & Great Depression




(d) Alternative Crisis Chronologies




*Notes:* This figure reports local projection estimates of the cumulative response of real GDP to a banking crisis, relative to the year prior to the crisis. The horizontal axis denotes years since crisis onset ($h = 0$), and the vertical axis plots cumulative percent deviations of real GDP from the pre-crisis level. Estimates are obtained from equation (1) using the full panel and conditioning on 25 lags of pre-crisis GDP growth. In all panels, the dark solid line corresponds to the baseline specification, with the shaded region denoting 90 percent confidence intervals based on standard errors clustered at the country level. Confidence intervals for specifications including the full set of controls are shown at five-year intervals. Panel (a) augments the baseline specification with controls for wars and conflicts (including "interstate", "intrastate", and 'extrastate" conflicts from Sarkees and Wayman (2010)) as well as sovereign debt and currency crises. Panel (b) instead augments the specification with macroeconomic controls measured at the time of the crisis, including inflation, interest rates, credit (M3), investment, the current account balance, government debt, and the real effective exchange rate. Panel (c) excludes crisis episodes occurring during World War I, World War II, and the Great Depression. Panel (d) reports estimates using alternative banking crisis chronologies. In each panel, thinner lines correspond to specifications that modify the baseline by adding one control or using an alternative definition. See Appendix Figures B.3, B.4, B.5, and B.6 for the full set of results across alternative lag structures (5, 10, 25) and balanced panels.

one hundred times yields the empirical distribution of impulse responses that would arise for a random shock occuring with the frequency of crises. If the estimated output losses we document were driven purely by underlying output persistence or other low-frequency trends, randomly assigning "crises" with the same overall frequency should generate similar impulse responses.

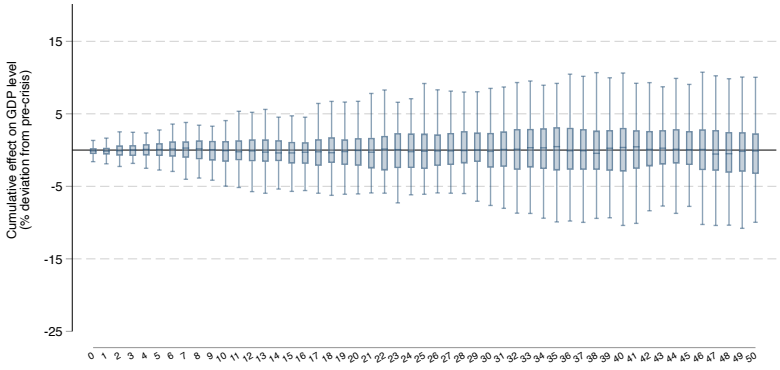
Figure 5a shows that the impulse responses generated by randomly assigned crises are centered near zero at all horizons and display no persistent decline in output. This pattern indicates that the baseline results are unlikely to arise from spurious correlations between crisis timing and slow-moving macroeconomic dynamics.

**Alternative LP specifications.** Our baseline specification measures the outcome as the cumulative change in log real GDP relative to the year prior to the crisis. To assess whether the results depend on this normalization, we estimate alternative local projection specifications using (i) the level of log real GDP and log real GDP per capita, and (ii) long differences in log real GDP per capita.

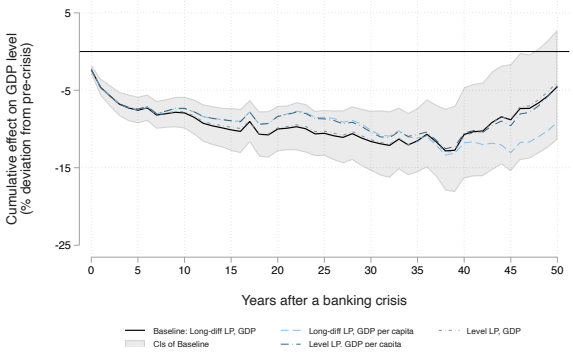
These specifications provide complementary views of the same underlying dynamics. As outlined by Jorda and Taylor (2025), long-difference specifications are less likely to be biased, but also require longer time series. Figure 5b shows that, in our setting, all alternative specifications closely track the baseline impulse response. Overall, the central pattern of long-run scarring is robust to these alternative outcome definitions.

Figure 5: Placebo Exercise and Alternative Estimators

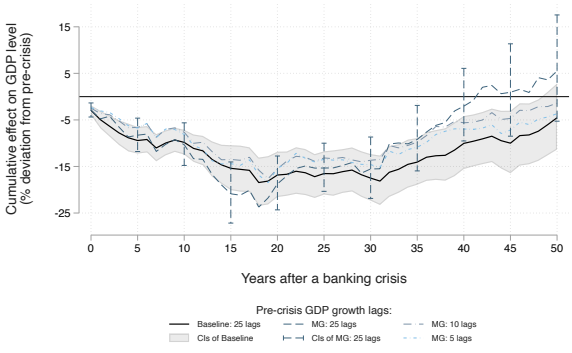
(a) Placebo Exercise




(b) Alternative LP Specifications




(c) Mean Group Estimator




*Notes:* This figure reports additional robustness exercises for the estimated long-run effects of banking crises on real GDP. The horizontal axis denotes years since the crisis onset ($h = 0$), and the vertical axis plots cumulative percent deviations of real GDP relative to the pre-crisis level. Panel (a) reports the distribution of impulse responses obtained from placebo simulations in which crisis timing is randomized while preserving the overall frequency of crises. In panel (b), the dark line denotes the baseline local-projection estimate using the full panel and conditioning on 25 lags of pre-crisis GDP growth, and the shaded region denotes the corresponding 90 percent confidence interval (standard errors clustered at the country level). Alternative specifications are shown as thinner lines without confidence intervals. Panel (b) compares the baseline estimates to alternative local-projection specifications using levels and long differences of log real GDP per capita. Panel (c) reports estimates based on the mean group estimator of Pesaran and Smith (1995) using the balanced panel. See Appendix Figures B.7, B.8, and B.9 for the full set of results across alternative lag structures and balanced panels.

**Mean group estimator.** We next re-estimate the dynamic effects of banking crises using the mean group estimator of Pesaran and Smith (1995). Unlike pooled local projections, which impose common slope coefficients across countries, the mean group estimator allows the dynamic response to vary flexibly at the country level and then averages the resulting country-specific coefficients. For each horizon $h$, local projections are estimated separately for each country, yielding a country-specific coefficient $\hat{\beta}_h^i$. The reported impulse response is the cross-country average $\hat{\beta}_h = (1/N)\sum_i \hat{\beta}_h^i$. This estimator is informative in our setting because it assigns equal weight to each country's estimated response and therefore allows for the possibility that the output losses of crises are highly heterogeneous across countries. We implement the mean group estimation using the same controls, lag structure, and dependent variable transformation as in the baseline analysis.

To isolate the role of cross-country heterogeneity, we implement this estimator on the balanced panel. This ensures that differences relative to the baseline estimates reflect the relaxation of the homogeneity assumption rather than changes in sample composition across horizons. Applying the mean group estimator to the unbalanced panel would introduce additional variation from the changing set of countries contributing to the estimates at longer horizons.

Figure 5c shows that the mean group estimates continue to display large and persistent output losses following banking crises. The qualitative dynamics closely mirror those in the pooled baseline, with output remaining substantially below its pre-crisis trajectory for several decades. However, the estimated effects are somewhat smaller in magnitude than in the pooled specification.

This pattern indicates that the pooled estimates place relatively greater weight on episodes with more severe and persistent downturns. By assigning equal weight to each country, the mean group estimator reduces the influence of such episodes and yields a somewhat more moderate estimate of the average response at very long horizons. We examine potential sources of such heterogeneity in Section 4.

Taken together, these exercises indicate that the persistent output losses following banking crises are not driven by spurious low-frequency trends or by a particular econometric specification. The placebo results show that the decline in output is not the result of random shocks occuring with the same frequency as crises. We therefore view the long-run output losses documented above as a robust feature of the data, and next turn to the mechanisms that may account for this persistence.

# 4 Mechanisms

The preceding analysis shows that financial crises are followed by exceptionally large and persistent output losses. This section investigates the mechanisms underlying this persistence. Our tests are guided by the view that crises operate through disruptions to financial intermediation, which tighten balance sheet constraints and impair the allocation of capital. Rather than testing a single model,

we examine several empirically distinct manifestations of this mechanism.

We focus on three complementary channels. First, we examine whether financial crises increase vulnerability to subsequent crises. In other words, we ask whether a crisis begets other crises, which could compound output losses over time. Second, we study the accumulation of both physical and human capital to assess whether tighter financial constraints translate into sustained reductions in productive capacity. Third, we analyze the response of government balance sheets, including revenues, expenditures, and public debt. Our overarching takeaway is that financial crises are associated with long-run scarring because they undermine the economy's productivity capacity, makes future shocks more likely, and limits governments' ability to cushion such shocks.

### 4.1 Crisis Recurrence and Financial Fragility

A natural explanation for our findings of deep output losses is that crises, like other misfortunes, seldom come alone. It is well known that crises often come at the heel of credit booms (Gourinchas and Obstfeld, 2012; Schularick and Taylor, 2012), which in turn are fueled by financial liberalization (Demirgüç-Kunt and Detragiache, 1998). If the factors giving rise to crises are persistent, or if crises themselves increase vulnerability, the result may be that crises cluster over time. The predictable result would then be a compounding of output losses, which could contribute to slow and incomplete recoveries at the scale of decades following the initial crisis.

To assess whether financial crises increase the risk of subsequent crises, we estimate the dynamic response of the cumulative probability of experiencing an additional crisis using a local projections framework analogous to that for output. Specifically, at horizon $h$, the outcome is an indicator equal to one if a country experiences at least one additional banking crisis *between* years $t+1$ and $t+h$, and zero otherwise. We then project this cumulative indicator on the onset of a banking crisis at time $t$, conditioning on country fixed effects and the same long history of pre-crisis output dynamics as in the baseline specification.[3]

The values of the dependent variable—the cumulative probability of a crisis—increase mechanically with the time horizon. However, the $\beta_h$ estimates are still informative, because they capture conditional differences between countries experiencing a crisis at baseline relative to those that do not.

Figure 6a reports the estimated relation between a banking crisis and the probability of experiencing at least one additional crisis within the subsequent $h$ years. In the short run, the estimated

[3]Formally, for each horizon $h$ we estimate

$$Y_{i,t,h} = \beta_h \,\text{BankingCrisis}_{i,t} + \alpha_i^{(h)} + \sum_{\ell=1}^{L} \phi_\ell^{(h)} \Delta \log Y_{i,t-\ell} + \varepsilon_{i,t}^{(h)},$$

where $Y_{i,t,h}$ is an indicator equal to one if country $i$ experiences at least one additional banking crisis between years $t+1$ and $t+h$, and zero otherwise.

coefficients are small or negative. In other words, it is less likely for a distinct new crisis episode to occur while the initial crisis is still unfolding. At longer horizons, however, the estimated coefficients rise sharply. Countries experiencing a banking crisis face a substantially higher probability of subsequent crises over the following decades, before the difference gradually narrows as cumulative crisis probabilities rise for both treated and untreated observations.

Taken together, these results provide evidence that crises beget future crises. While the likelihood of additional crises eventually diminishes, the elevated risk over two or three decades suggests a clustering of crisis episodes that can compound output losses over time. We next assess whether this channel can account for the long-run output losses documented in the baseline analysis.

#### 4.1.1 Accounting for Subsequent Crises

Are the long-run output losses driven by the recurrence of crises? To address this question, we re-estimate the baseline local projections while explicitly accounting for the occurrence of crises between the initial event and each projection horizon. We implement two complementary approaches.

In the first approach, we augment the baseline specification with a control for the proximity of the most recent financial crisis prior to horizon $h$. Specifically, at each horizon we include a variable measuring the number of "years since" the most recent financial crisis as of time $t+h$. In essence, we control for the time relative to the most recent crisis episodes.[4] Because this measure is undefined for countries that never experience a crisis, the specification is estimated on the subsample of countries with at least one crisis episode.

In the second approach, we directly control for crisis episodes occurring anywhere within the window $[t-L,\, t+h-1]$ by including crisis indicators for both past and future periods in the local projection specification. This absorbs the effects of any intervening crises, allowing the coefficient on the initial crisis to isolate its direct impact while conditioning on the full sequence of crisis events.[5]

Figure 6b reports the resulting impulse responses. Across both approaches, explicitly accounting for intervening crises does not attenuate the estimated long-run output losses following a banking crisis. Output remains substantially below its pre-crisis trajectory for several decades, with little evidence of convergence. If anything, the estimated effects remain comparable or somewhat larger in magnitude than in the baseline specification. Thus, the recurrence of crises alone cannot account for

[4] This specification augments equation (1) with the additional regressor $\text{YearsSince}_{i,t+h}$, which measures the number of years elapsed since the most recent financial crisis as of time $t+h$ and resets to zero at each new crisis.

[5] Leads of the crisis indicator are constructed using missing-dummy methods following Teulings and Zubanov (2014), so that the effective sample remains comparable across horizons. Formally, for each horizon $h$ we estimate

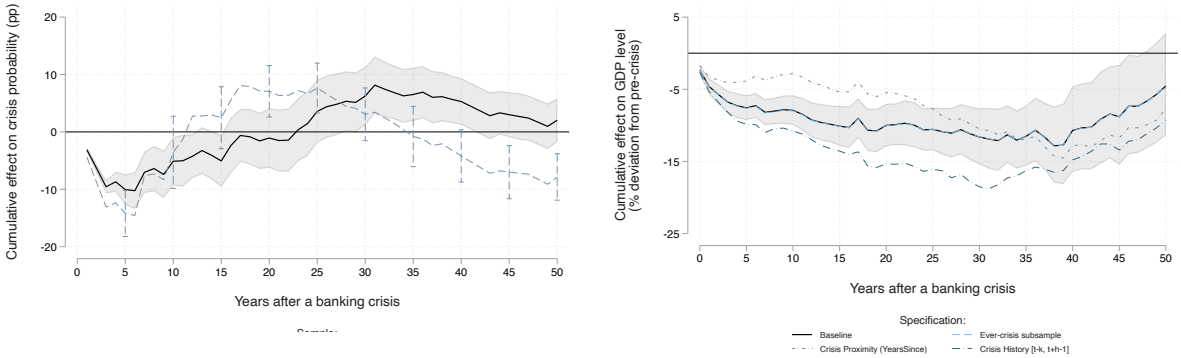
$$\Delta \log Y_{i,t+h} = \beta_h \,\text{BankingCrisis}_{i,t} + \alpha_i^{(h)} + \sum_{\ell=1}^{L} \phi_\ell^{(h)} \Delta \log Y_{i,t-\ell} + \sum_{s=-L}^{h-1} \kappa_s^{(h)} \,\text{BankingCrisis}_{i,t+s} + \varepsilon_{i,t}^{(h)},$$

where the additional terms include $L$ lags and $h-1$ leads of the crisis indicator, thereby controlling for crisis episodes occurring anywhere within the window $[t-L,\, t+h-1]$.

the magnitude and persistence of the long-run output losses. Instead, the results point to additional channels through which financial crises generate economic scarring.

Figure 6: Subsequent Financial Crises

(a) Impact on Subsequent Financial Crises (b) Accounting for Subsequent Crises




*Notes:* Panel (a) reports local projection estimates obtained from equation (1) of the effect of a banking crisis on the cumulative probability of experiencing a subsequent financial crisis. The dependent variable at horizon $h$ is an indicator equal to one if a country experiences at least one additional banking crisis between years $t+1$ and $t+h$, and zero otherwise. The horizontal axis denotes years since the initial crisis onset ($h=0$), and the vertical axis plots the estimated change in the cumulative probability of a future crisis relative to countries that do not experience a crisis at time $t$. Panel (b) reports specifications that account for intervening crisis episodes using two alternative approaches: controlling for the time since the most recent crisis and directly conditioning on crisis indicators within the window $[t-L,\, t+h-1]$. All specifications include country fixed effects and the same long history of pre-crisis output growth as in the baseline output regressions. The shaded region denotes 90 percent confidence intervals for the baseline specification, based on standard errors clustered at the country level. Thinner lines correspond to alternative specifications. See Appendix Figure B.10 for the full set of results across alternative lag structures and balanced panels.

## 4.2 Capital Accumulation and Labor Markets

When banking systems are impaired, credit intermediation becomes less effective, limiting the ability of firms and households to finance investment and sustain employment (e.g., Bernanke and Gertler, 1989; Kiyotaki and Moore, 1997). As a result, financial crises may not only reduce short-run output but also depress the accumulation of physical and human capital, generating persistent effects on productive capacity.

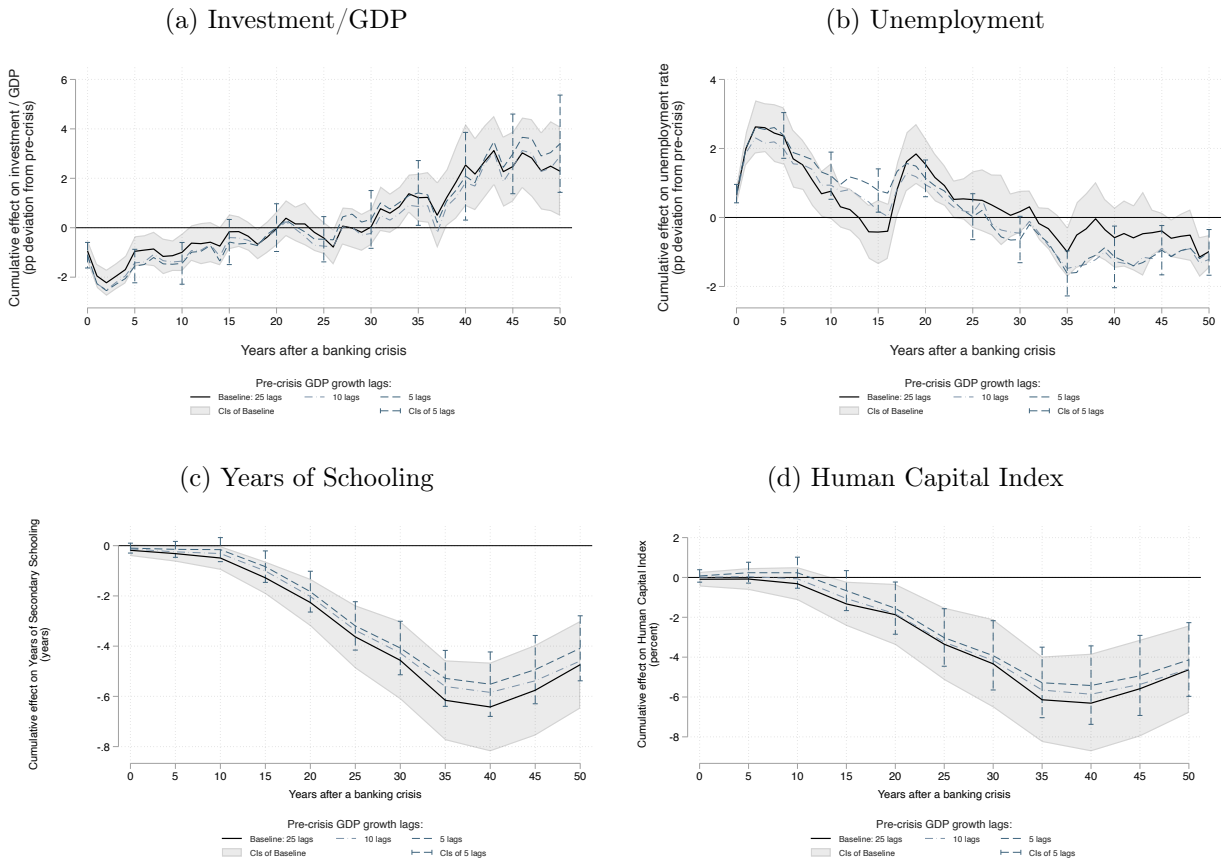
### 4.2.1 Investment

We begin by examining the evolution of investment following banking crises. Both are particularly sensitive to financial conditions because many capital expenditures require external financing. When banks experience balance sheet stress, the reduced lending capacity and heightened uncertainty that accompany crises may lead firms to delay, scale back, or cancel planned investment projects.[6]

[6]For the estimation of investment outcomes, we focus on the balanced panel, because long-horizon cumulative responses are particularly sensitive to horizon-specific sample composition in the unbalanced panel. Holding the set of crisis episodes fixed across horizons yields more stable and interpretable estimates.

Figure 7a plots the response of investment as a share of GDP following banking crises. The investment-to-GDP ratio declines sharply in the years immediately following the crisis, indicating that investment falls by more than output, and remains below its pre-crisis level for a period of around 15 to 20 years. At the horizon of approximately four decades after the crisis, investment actually exceeds its pre-crisis level by a magnitude comparable to the initial trough decline, suggesting eventual recovery in the economy's capital stock.

Figure 7: Capital Accumulation after Banking Crises




*Notes:* This figure reports local projection estimates of the response of capital accumulation, labor market outcomes, and human capital to a banking crisis. The horizontal axis denotes years since crisis onset ($h = 0$), and the vertical axis plots the cumulative response of each outcome relative to the year prior to the crisis. Estimates are obtained from equation (1) using the full panel and conditioning on 25 lags of pre-crisis output growth. Panel (a) reports the response of investment as a share of GDP, and Panel (b) reports the response of the unemployment rate. Panels (c) and (d) report the response of human capital measured in terms of years of schooling and an overall index, for both male and female from Barro and Lee (1996) and Lee and Lee (2016). These data are only available every five years, and therefore estimates are reported at those intervals. In all panels, the dark solid line denotes the baseline specification, and the shaded region denotes 90 percent confidence intervals based on standard errors clustered at the country level. See Appendix Figures B.11 for the full set of results across balanced panels.

### 4.2.2 Labor Markets and Human Capital

Financial crises may also affect output persistently through their impact on labor markets and human capital accumulation. When firms face tighter financing constraints, they may reduce hiring,

cut hours, or lay off workers in response to declining demand and limited access to credit (e.g., Mian, Sufi and Verner, 2017). Disruptions to firm balance sheets may also lead to layoffs and shutdowns.

Figure 7b presents the response of unemployment following banking crises using the same local projection framework. We find that unemployment rises significantly in the aftermath of crises and recovers only gradually over time. In the baseline specification, it takes unemployment around 15 years to get back to pre-crisis level, which is then followed by another spike. Our estimates thus suggest that unemployment only fully recovers perhaps 25-30 years after a crisis.

In addition to affecting employment, financial crises may disrupt investments in human capital. Households facing income shocks or tighter credit conditions may reduce their spending on education or delay entry into skill-intensive occupations. At the same time, governments experiencing fiscal stress may cut education spending or reduce support for training programs.[7]

Figures 7c and 7d present impulse responses for years of schooling and a human capital index using data from Barro and Lee (1996) and Lee and Lee (2016). Human capital accumulation shows severe and permanent reductions following a crisis. In fact, even after five decades, we observe only limited evidence of a recovery to previous levels of human capital. Alternative samples, and the breakdown by male and female respectively, are reported in Appendix B and yield similar patterns.

Taken together, these findings suggest that financial crises have long-lasting consequences for investment, labor markets, and human capital. Sluggish investments in physical and human capital are thus a likely contributor to the sustained output losses of crises.

### 4.3 Fiscal Impact

A long literature has shown that banking crises do not only impair the balance sheets of the private sector, but also have consequences for government finances. Different from existing work, we focus on the very long-run trajectory of fiscal matters, again up to five decades after a crisis hits.

Figure 8 presents impulse responses for government revenues, government spending, fiscal balances, and public debt, all as a share of GDP, using the same empirical framework as in the baseline. Government revenues decline sharply and persistently, even as a share of GDP. Crises are thus followed by a permanent reduction in government income in a way that is not fully explained by lower output. Government expenditure rises in the immediate aftermath of the crisis, consistent with a countercyclical response. After around five years, however, expenditures revert and then decline as a share of GDP at longer horizons. The fiscal balance correspondingly deteriorates strongly in the first five years following a crisis before returning close to its pre-crisis level.

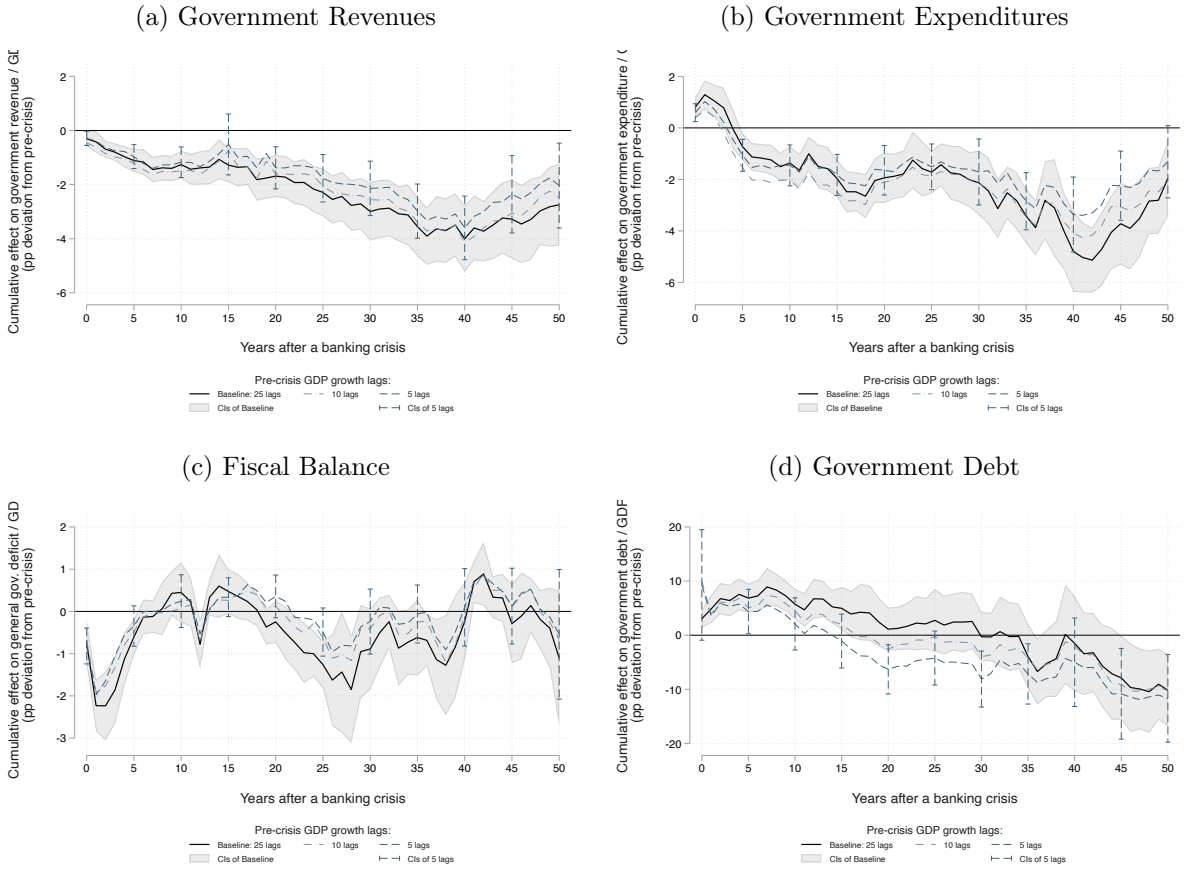
Faced with lower revenues and higher expenditures, governments incur considerable debt in the aftermath of financial crises. At a horizon of ten years, the ratio of government debt to GDP stands

[7] We will return to government balance sheets Section 4.3.

at around 10 percentage points higher than if no crisis had occured. It is only after around 30 years that elevated debt levels eventually get back to pre-crisis levels.

Our findings on the short-run dynamics align with existing evidence that banking crises trigger a countercyclical fiscal expansion and modest debt accumulation (e.g., Reinhart and Rogoff, 2009*a*,*b*; Laeven and Valencia, 2013). Over the multi-decade horizon enabled by our data, however, the initial countercyclical expansion is reversed, debt eventually declines below pre-crisis levels, and the long-run fiscal position is characterized by persistent revenue erosion alongside a corresponding contraction in the size of the public sector. These patterns are consistent with binding constraints on sustained deficits, with countries ultimately bringing expenditure into line with the reduced revenue base. The initial debt accumulation is gradually offset by the combination of this expenditure restraint and the eventual recovery of output.

Figure 8: Fiscal Dynamics Following Banking Crises




*Notes:* This figure reports local projection estimates of the response of fiscal outcomes to a banking crisis. The horizontal axis denotes years since crisis onset ($h = 0$), and the vertical axis plots cumulative changes in each fiscal variable relative to the year prior to the crisis, expressed as a share of GDP. Estimates are obtained from equation (1) using the full panel and conditioning on 25 lags of pre-crisis GDP growth. The dark solid line denotes the baseline specification, and the shaded region denotes 90 percent confidence intervals based on standard errors clustered at the country level. Thinner lines correspond to alternative specifications. See Appendix Figures B.12 for the full set of results across balanced panels.

# 5 Conclusion

Banking crises are often framed as a kind of hiccup in business cycle dynamics. This paper documents that we may better want to think about such crises as a potential threat to long-run economic development. As we show using a new macroeconomic dataset encompassing a broad set of countries and several centuries of data, real GDP remains substantially below its pre-crisis trend for many decades, with little evidence of full recovery. These patterns are robust across specifications, samples, and alternative crisis definitions, and are not explained by the occurrence of subsequent crises alone.

The evidence in this paper points to a broader interpretation of financial crises as episodes that shift the economy onto a persistently lower trajectory by tightening financial constraints. The sustained declines in investment, employment, and human capital suggest that financial frictions are not merely cyclical propagation mechanisms but a central determinant of long-run economic dynamics.

Our findings speak to a broader policy debate on whether the occasional crisis is a "price worth paying." Rancière, Tornell and Westermann (2008), for example, show that countries with more negatively skewed credit growth have higher real GDP growth. They argue that systemic risk is a byproduct of looser borrowing constraints and ultimately desirable. The evidence presented here challenges this view by showing that crises are sufficiently disruptive to hold back economic development for decades relative to a plausible counterfactual.

Finally, the Global Macro Database (GMD) presented here has already proven to be useful in many other domains, such as macroeconomics, development economic, international trade, or comparative political science. It is also widely used for teaching economics. Going forward, we plan to continuously extend and update the data with the goal of turning it into a trusted one-stop resource for economists worldwide.

# References


**Antolin-Diaz, Juan, and Paolo Surico.** 2025. "The Long-Run Effects of Government Spending." *American Economic Review*, 115(7): 2376–2413.

**Ball, Laurence.** 2014. "Long-Term Damage from the Great Recession in OECD Countries." *European Journal of Economics and Economic Policies: Intervention*, 11(2): 149–160.

**Baron, Matthew, Emil Verner, and Wei Xiong.** 2021. "Banking crises without panics." *The Quarterly Journal of Economics*, 136(1): 51–113.

**Barro, Robert J.** 2006. "Rare Disasters and Asset Markets in the Twentieth Century." *Quarterly Journal of Economics*, 121(3): 823–866.

**Barro, Robert J, and Jong Wha Lee.** 1996. "International measures of schooling years and schooling quality." *The American Economic Review*, 86(2): 218–223.

**Barro, Robert J., and Jong Wha Lee.** 2013. "A new data set of educational attainment in the world, 1950–2010." *Journal of Development Economics*, 104(C): 184–198.

**Barro, Robert J., and José F. Ursúa.** 2008*a*. "Macroeconomic Crises since 1870." *Brookings Papers on Economic Activity*, 2008(1): 255–350.

**Barro, Robert J., and José F. Ursúa.** 2008*b*. "Macroeconomic Crises Since 1870." *Brookings Papers on Economic Activity*, 2008: 255–335. Accessed: 2024-11-22.

**Barro, Robert J, and José F Ursúa.** 2008*c*. "Macroeconomic Crises Since 1870." National Bureau of Economic Research.

**Baumeister, Christiane.** 2025. "Comment on "Local Projections or VARs? A Primer for Macroeconomists""." *NBER Macroeconomics Annual 2025, volume 40*, 153–171. University of Chicago Press.

**Bernanke, Ben S., and Mark Gertler.** 1989. "Agency Costs, Net Worth, and Business Fluctuations." *American Economic Review*, 79(1): 14–31.

**Blanchard, Olivier, Eugenio Cerutti, and Lawrence Summers.** 2015. "Inflation and Activity: Two Explorations and Their Monetary Policy Implications." *IMF Working Paper*, , (15/230).

**Bordo, Michael, Barry Eichengreen, Daniela Klingebiel, and Maria Soledad Martinez-Peria.** 2001. "Is the Crisis Problem Growing More Severe?" *Economic Policy*, 16(32): 52–82.

**Cerra, Valerie, and Sweta Chaman Saxena.** 2008. "Growth dynamics: the myth of economic recovery." *American Economic Review*, 98(1): 439–457.

**Cerra, Valerie, Antonio Fatás, and Sweta Chaman Saxena.** 2023. "Hysteresis and Business Cycles." *Journal of Economic Literature*, 61(1): 181–225.

**Davis, Donald R, and David E Weinstein.** 2002. "Bones, bombs, and break points: the geography of economic activity." *American economic review*, 92(5): 1269–1289.

**Demirgüç-Kunt, Asli, and Enrica Detragiache.** 1998. "The Determinants of Banking Crises in Developing and Developed Countries." *IMF Staff Papers*, 45(1): 81–109.

**Feenstra, Robert C., Robert Inklaar, and Marcel P. Timmer.** 2015. "The Next Generation of the Penn World Table." *American Economic Review*, 105(10): 3150–3182.

**Funke, Manuel, Moritz Schularick, and Christoph Trebesch.** 2016. "Going to Extremes: Politics after Financial Crises, 1870–2014." *European Economic Review*, 88: 227–260.

**Gourinchas, Pierre-Olivier, and Maurice Obstfeld.** 2012. "Stories of the Twentieth Century for the Twenty-First." *American Economic Journal: Macroeconomics*, 4(1): 226–265.

**Guiso, Luigi, Helios Herrera, Massimo Morelli, and Tommaso Sonno.** 2017. "Demand and Supply of Populism." *CEPR Discussion Paper*, , (11871).

**Jordà, Òscar.** 2005. "Estimation and inference of impulse responses by local projections." *American economic review*, 95(1): 161–182.

**Jorda, Oscar, and Alan M. Taylor.** 2025. "Local Projections." *Journal of Economic Literature*, 63(1): 59–110.

**Jordâ, Oscar, Moritz Schularick, and Alan M. Taylor.** 2013. "When Credit Bites Back." *Journal of Money, Credit and Banking*, 45(s2): 3–28.

**Jordà, Òscar, Moritz Schularick, and Alan M Taylor.** 2017*a*. "Macrofinancial history and the new business cycle facts." *NBER macroeconomics annual*, 31(1): 213–263.

**Jordá, Óscar, Moritz Schularick, and Alan M. Taylor.** 2017*b*. "Macrofinancial History and the New Business Cycle Facts." *NBER Macroeconomics Annual*, 31(1): 213–263.

**Kaminsky, Graciela L., and Carmen M. Reinhart.** 1999. "The Twin Crises: The Causes of Banking and Balance-of-Payments Problems." *American Economic Review*, 89(3): 473–500.

**Kiyotaki, Nobuhiro, and John Moore.** 1997. "Credit Cycles." *Journal of Political Economy*, 105(2): 211–248.

**Krishnamurthy, Arvind, and Tyler Muir.** 2025. "How Credit Cycles Across a Financial Crisis." *Journal of Finance*, 80(3): 1339–1378.

**Laeven, Luc, and Fabian Valencia.** 2013. "Systemic banking crises database." *IMF Economic Review*, 61(2): 225–270.

**Lee, Jong-Wha, and Hanol Lee.** 2016. "Human capital in the long run." *Journal of Development Economics*, 122(C): 147–169.

**Lehbib, Mohamed, and Karsten Müller.** 2025. "GMD: The Easy Way to Access the World's Most Comprehensive Macroeconomic Database." Working Paper. Accessed: 2025-12-29.

**Li, Dake, Mikkel Plagborg-Møller, and Christian K. Wolf.** 2024. "Local projections vs. VARs: Lessons from thousands of DGPs." *Journal of Econometrics*, 244(2): None.

**Ludwig, Julian F.** 2024. "Local projections are VAR predictions of increasing order." *Available at SSRN 4882149*.

**Mian, Atif, Amir Sufi, and Emil Verner.** 2017. "Household Debt and Business Cycles Worldwide." *Quarterly Journal of Economics*, 132(4): 1755–1817.

**Mitchell, Brian.** 2013. *International Historical Statistics. International Historical Statistics*. 1 ed., Palgrave Macmillan London.

**Montiel Olea, José Luis, Mikkel Plagborg-Møller, Eric Qian, and Christian K Wolf.** 2025. "Local Projections or VARs? A Primer for Macroeconomists." National Bureau of Economic Research Working Paper 33871.

**Müller, Karsten, and Emil Verner.** 2024. "Credit Allocation and Macroeconomic Fluctuations." *Review of Economic Studies*, 91(6): 3645–3676.

**Müller, Karsten, Chenzi Xu, Mohamed Lehbib, and Ziliang Chen.** 2025. "The Global Macro Database: A New International Macroeconomic Dataset." National Bureau of Economic Research Working Paper 33714.

**Pesaran, M Hashem, and Ron Smith.** 1995. "Estimating long-run relationships from dynamic heterogeneous panels." *Journal of econometrics*, 68(1): 79–113.

**Piger, Jeremy, and Thomas Stockwell.** 2025. "Differences From Differencing: Should Local Projections With Observed Shocks Be Estimated in Levels or Differences?" *Journal of Applied Econometrics*, 40(7): 759–787.

**Plagborg-Møller, Mikkel, and Christian K. Wolf.** 2021. "Local Projections and VARs Estimate the Same Impulse Responses." *Econometrica*, 89(2): 955–980.

**Quincy, Sarah, and Chenzi Xu.** 2026. "Regulation in Crisis: Bank Failures and Entry Barriers."

**Rancière, Romain, Aaron Tornell, and Frank Westermann.** 2008. "Systemic Crises and Growth." *The Quarterly Journal of Economics*, 123(1): 359–406.

**Reinhart, Carmen M, and Kenneth S Rogoff.** 2009*a*. "The aftermath of financial crises." *American Economic Review*, 99(2): 466–472.

**Reinhart, Carmen M, and Kenneth S Rogoff.** 2009*b*. *This time is different: Eight centuries of financial folly*. princeton university press.

**Romer, Christina D., and David H. Romer.** 2019. "Fiscal Space and the Aftermath of Financial Crises: How It Matters and Why." *Brookings Papers on Economic Activity*, 2019(1): 239–313.

**Sarkees, Meredith Reid, and Frank Wayman.** 2010. "COW War Data, 1816–2007 (Version 4.0)." *Correlates of War Project Data Set*, Available at https://correlatesofwar.org/data-sets/cow-war/.

**Schularick, Moritz, and Alan M. Taylor.** 2012. "Credit Booms Gone Bust: Monetary Policy, Leverage Cycles, and Financial Crises, 1870–2008." *American Economic Review*, 102(2): 1029–1061.

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

**Teulings, Coen N, and Nikolay Zubanov.** 2014. "Is economic recovery a myth? Robust estimation of impulse responses." *Journal of Applied Econometrics*, 29(3): 497–514.

**Thomas, Ryland, Sally Hills, and Nicholas Dimsdale.** 2010. "The UK Recession in Context — What Do Three Centuries of Data Tell Us?" *Bank of England Quarterly Bulletin*, 50(4): 277–291.

**World Bank.** 2026. "How Do You Derive Constant Price Series in USD?" *World Bank Data Help Desk*, World Development Indicators methodological documentation.

**Xu, Chenzi.** 2022. "Reshaping global trade: the immediate and long-run effects of bank failures." *The Quarterly Journal of Economics*, 137(4): 2107–2161.

# A Data Appendix

## A.1 Introduction

The Global Macro Database (GMD), introduced by Müller, Xu, Lehbib and Chen (2025), is the world's most comprehensive open-source repository of cross-country macroeconomic statistics. Updated on a quarterly basis, the 2026–06 version of the GMD contains 46 variables, spanning 239 countries and territories over the period 1086 to 2025 with projections up to 2030. The data and further information can be found at www.globalmacrodata.com.

This appendix is specific to the 2026–06 release of the GMD.

## A.2 Terms of Use

**Version 1.0.** By downloading or using the Global Macro Database (GMD), you accept these Research Use Terms.

**About these Terms** The GMD is built for academic and non-profit researchers. These Research Use Terms explain how you can use the GMD. They follow the spirit of the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 (CC BY-NC-SA 4.0) license. If anything here differs from it, these Terms apply.

Not sure your use is allowed? Treat it as commercial and contact us before downloading at kmueller@globalmacrodata.com.

**Permitted users** You can use the GMD if you are:

- A university or college, or its staff and students.
- A non-profit research organization.
- An individual researcher, for your own scholarly work meant for publication or peer review.
- A journalist or author, for reporting, books, or commentary, with citation.
- A teacher or instructor, for the classroom.
- A government body, central bank, regulator, finance ministry, or other public-sector or international organization, for non-commercial research, analysis, and policy work.

If you are not on this list, please get in touch before using the GMD.

**Permitted uses** If you are one of the above, you are permitted to use the GMD to:

- Conduct research for papers, working papers, book chapters, theses, and other open scholarship.
- Teach and supervise student research.
- Report on it or write commentary, with citation.
- Engage in non-commercial research, analysis, and policy work inside a public-sector institution.

For these uses, you can copy, analyze, and build on the GMD, as long as you follow the citation and share-alike rules below.

**Do not re-host or redistribute it.** Without explicit permission, do not republish the GMD, in whole or in part, on another website, API, platform, product, or service, or under another name, and do not present it as your own. Send people to globalmacrodata.com so they get the latest data and cite it. One exception: you can include the specific data used in a paper in that paper's replication package if you clearly label the data as coming from the GMD and point people to globalmacrodata.com.

**Non-permitted uses** The GMD is not for commercial use or for use inside for-profit companies. In particular, you are not permitted to:

- **Use it inside a for-profit company.** That means any for-profit organization, including but not limited to banks, asset managers, hedge funds, family offices, consultancies, advisory firms, data vendors, fintechs, tech firms, insurers, rating agencies, and any company's research, strategy, treasury, or risk team, using it for the company's work, including business, investment, advisory, analytics, model-building, operations, or client services, even purely internally, and whether or not you ever share the data.
- **Build it into something you sell.** Do not put the GMD, whole, partial, or in derived form, into any product, platform, service, database, dashboard, API, model, index, benchmark, signal, paid report, or client deliverable that is sold, licensed, paywalled, or otherwise monetized.
- **Train commercial AI with it.** Do not train, fine-tune, or test machine-learning models that go into a commercial product or service, in whole, in part, or in derived form, or into paid client work.
- **Use it in marketing.** No ads, sponsored content, or other promotional material.

If you are not sure, treat your use as commercial and contact us first.

**Citation and attribution** Please credit the GMD by citing it as specified in Section A.3. If the documentation names a specific source for data you rely on heavily, please cite that source too. Do not imply that the GMD or its authors endorse you or your work.

**Share-alike** If you build on the GMD in any way, and want to share the result, do so under these same Terms. In case you want to distribute data, in whole or in part or derived form, building on the GMD, contact us first.

**Name and endorsement** Do not suggest that the GMD, its authors, its contributors, or its sources endorse you or your work. Do not use the "Global Macro Database" or "GMD" name or logo in a way that implies endorsement or a connection, without our written permission.

**No warranty** The GMD comes "as is," with no warranties of any kind. The authors and maintainers are not liable for any claims or damages that come from using it.

**Code license** The GMD's reusable helper functions are licensed under the MIT License unless a specific file says otherwise. Other repositories, packages, website code, and release infrastructure are governed by the license file or terms in the relevant repository. That covers code only, not the GMD data; the data is governed by these Terms.

**Contact** Questions about these Terms, or a use they do not cover? Email `kmueller@globalmacrodata.com`.

### A.3 Citation

If you use the GMD, it must be cited as follows. Because the GMD is updated quarterly, you may want to refer to the exact version you used:

```
@techreport{MuellerXuLehbib2026,
  title = {The Very Long-Run Costs of Financial Crises},
  author = {M{\"u}ller, Karsten and Xu, Chenzi and Lehbib, Mohamed},
  year = {2026},
  note = {Working paper}
}
```

## A.4 Version History

This section summarizes major updates to the Global Macro Database (GMD). For full release notes and version details, please visit www.globalmacrodata.com/data.

### Version 2026_06 – *current*

**Overview** The 2026-06 quarterly update expands coverage with thirty nine new aggregator and country-level historical sources, improves our continuous automated error-monitoring system to improve data quality, resolves dozens of data-quality bugs spanning units, currency redenominations, and series splicing, completes a comprehensive audit of the Mitchell historical statistics, and improves the Stata, R, and Python packages. The entire release is produced by the GMD pipeline running unfiltered, end-to-end, on cloud infrastructure.

**New Sources** This release adds thirty nine new sources, bringing the total in the database to 160 (up from 121 at the previous release). The additions are dominated by long-run country-level historical sources, complemented by several new cross-country aggregators. Highlights include:

- World Bank (2026): the World Bank Global Economic Monitor, providing global macroeconomic indicators.
- Consejo Monetario Centroamericano (2017): historical macroeconomic statistics from the Central American Monetary Council (SECMCA) for Costa Rica, the Dominican Republic, Guatemala, Honduras, Nicaragua, and El Salvador (1960--2017).
- Ritschl (2002): German interwar and pre-WWII macroeconomic series compiled from primary historical sources.
- Martín Moreno (2026): the Archivo de Historia Econ'omica de M'exico --- long-run Mexican national accounts, prices, money, and public finances (1876--2025).
- Abdulai (2022): historical monetary, inflation, and output series for Ghana (1980--2017).

- Foster (1996): Foster's Australian Economic Statistics (1949/50--1994/95).
- Eichholtz (1997): the Herengracht house-price index (1628--1973), one of the longest continuous real-estate price series available.
- Min. of Economy and Finance, Korea (2026): Open Fiscal Data from the Republic of Korea's Ministry of Economy and Finance.
- Xu et al. (2015): Chinese national income estimates, 1661--1933.
- Feinstein (1972): Feinstein's National Income, Expenditure and Output of the United Kingdom (1855--1965).
- Sefton and Weale (1995): the Sefton--Weale balanced estimates of UK national income (1920--1990).

Beyond these highlights, the release adds many further long-run country-level historical sources --- including new slots for Germany (six new historical series), the United Kingdom, India, New Zealand, the Netherlands, Australia, Spain, Sweden, Turkey, Japan, Russia, Israel, Bulgaria, and Venezuela --- alongside additional cross-country aggregators.

In addition, the CEPII trade-and-macro source (Fouquin and Hugot (2016), two centuries of bilateral trade and gravity data, 1827--2014) has been refreshed and renamed from `CEPII_TRADHIST` to `CEPII`.

**Extended Error Monitoring** This release extends our continuous error-monitoring system that automatically scans the database for ratio discrepancies, unit breaks, and other anomalies, and routes each flagged case through an auto-triage workflow for review.

**Data Quality and Bug Fixes** Driven in large part by the improved monitoring system and by reports from our community of contributors, this release resolves dozens of bugs. The most impactful corrections fall into two groups:

- **Units, scaling, and redenominations.** We corrected unit-scale and currency-redenomination errors across many sources, including the IMF IFS, IMF GFS, Andersson, the Bordo monetary aggregates, Flora, the African Development Bank source, HFS, and the United Nations.
- **Splicing and ratio consistency.** We improved the handling of gaps between sources, and corrected the US long-term interest-rate series.

**Mitchell Historical Audit** We completed a comprehensive audit of the Mitchell International Historical Statistics, fixing issues and correcting the euro-cutoff handling for European government tax series.

**Easier Download of Underlying Raw Data** Previously, users could only access the raw data underlying the spliced GMD series by either downloading the Excel file or using our Stata, R, or Python packages. From this version onwards, users can also directly include all underlying raw data when downloading the csv or Stata versions.

**Package Updates** The Stata, R, and Python packages have all been improved in this release. The three distributions remain at full feature parity, so every user has the same access to the combined database, the underlying source data, and the documentation regardless of which language they work in.

**Documentation and License** We substantially expanded the Technical Appendix. It now includes a clarified license and a comprehensive Frequently Asked Questions section. A redesigned set of per-variable and per-country PDF documentation make it easier to trace where the data are taken from.

**Acknowledgements** Thanks to everyone who reported issues and suggested improvements. Many of the fixes in this release came directly from user reports.

### Version 2026_03

**Overview** This release adds eleven new data sources, introduces two new variables, improves the methodology for splicing government finance ratios, harmonizes all ratio variables, and updates the Python and R packages to full feature parity with the Stata package.

**New Sources** We added eleven new sources to the database:

- Banushi et al. (2024): The wiiw COMECON Dataset, covering economic time series for the command economies of Eastern Europe (1944--1994), including GDP, consumption, trade, government finance, monetary, and price data for nine countries.
- Cogneau, Dupraz and Mesplé-Somps (2021): Colonial fiscal, GDP, trade, and population data for French colonies (1833--1962), covering Algeria, Tunisia, Morocco, Madagascar, Cameroon, and Togo.

- Kehoe and Nicolini (2022): The Monetary and Fiscal History of Latin America project, covering GDP, inflation, fiscal balances, government debt, exchange rates, and monetary base for eleven Latin American countries (1960--2017).
- Andersson (2026): Central government revenue data from Per F. Andersson, covering revenue and GDP for multiple countries from the 1800s onwards.
- Herranz-Loncan and Peres-Cajias (2016): Historical real GDP for Bolivia from Herranz-Loncan & Peres-Cajias (2016), extending coverage back to the mid-nineteenth century (1846, 1890--2010).
- Peres-Cajias (2014): Bolivian public finance data from Peres-Cajias (2014), covering central and general government revenue, expenditure, and tax ratios (1882--2010).
- Banco Central de Reserva, Perú (2022): Historical macroeconomic data from the Banco Central de Reserva del Per'u, covering national accounts, prices, trade, monetary aggregates, and fiscal data for Peru (1922--2021).
- Banco de la República, Colombia (2024): Historical series from Colombia's central bank, covering government finances, trade, current account, and monetary aggregates.
- Hong Kong Institute (2024): Historical data for Hong Kong including monetary aggregates, government finance, trade, GDP, exchange rates, and prices (1843--2002).
- Hubmann, Jobst and Maier (2020): Long-run Austrian CPI series from Hubmann, Jobst & Maier (2020), covering 1800--2018.
- HM Treasury (2026): UK historical public finances from HM Treasury, covering government revenue and expenditure.

**New Variables** We introduced two new consumption variables: household consumption (`hcons`) and government consumption (`gcons`). These complement the existing total consumption (`cons`) variable and provide a finer decomposition of the expenditure side of GDP.

**Improved Government Finance Ratio Splicing** We introduced a new methodology for combining government finance ratios (revenue, expenditure, tax, debt, and deficit as % of GDP). Previously, we spliced the underlying level series and then derived the ratios. We now splice the ratios directly, which avoids compounding errors that arise when the numerator and denominator are spliced separately with different chainlinking adjustments.

**Ratio Harmonization** All ratio variables (e.g., `govdebt_GDP`, `exports_GDP`, `CA_GDP`) are now consistently expressed in percent, so that a value of 50 means 50% of GDP.

**Package Updates** The Python and R packages have been updated to match the full functionality of the Stata package, including access to underlying source data and documentation features.

**Data Quality** We incorporated feedback from GMD users and improved data quality across multiple variables and sources.

### Version 2026_01

**Overview** This release introduces significant enhancements to data accuracy and infrastructure. Key updates include a comprehensive revision of the real GDP series and the deployment of a fully automated, cloud-based data processing pipeline to ensure timely future updates.

**Real GDP Improvement** We have conducted a major review of the real GDP series. The data has been rigorously corrected and is now consistently rebased to the year 2015, ensuring greater comparability and accuracy across the dataset.

**Automated Pipeline** We have implemented a new automated pipeline. This system autonomously handles downloading, processing, and merging data from all sources in the cloud, ensuring the maintenance process and allowing for more frequent and reliable database updates.

**Stata Package & Documentation** We have launched a dedicated repository for the official Stata package [1]. Moreover, we have released a comprehensive companion paper[2], which serves as a detailed guide to using the package effectively.

### Version 2025_12

**Overview** The *2025_12* version includes updated data as of December 2025 and introduces various important patches and improvements. We also rewrote the Stata package from scratch: get the new version by typing `ssc install gmd`. Lehbib and Müller (2025) provides more details.

**Improved Government Finance Statistics** We further improved the construction of combined government finance statistics. Relative to before, the combined time series are now mostly based on

[1] To access the package source-code, please visit: github.com/KMueller-Lab/Global-Macro-Database-Stata.

[2] To access the paper, please visit: github.com/KMueller-Lab/Global-Macro-Database-Stata/blob/main/Global_Macro_Database_Stata.pdf.

chain-linking ratios, with some exceptions, and we more commonly use a country-specific priority ordering of sources.

**Extended Technical Appendix** We considerably improved the technical appendix to enhance clarity and readability. Going forward, we will provide a dedicated technical appendix with each release.

**Major Update to Stata Package** We rewrote the Stata package from scratch to make it faster and added various new functionalities, including the ability to easily access all the (cleaned) data underlying the GMD. A new companion paper (Lehbib and Müller, 2025) now describes the package in detail.

**Bug Fixes** Thanks to the support of many GMD users, we were able to identify and fix many bugs. Noteworthy examples include real GDP per capita for Venezuela and the inflation rates of a few countries.

**New Variable** The GMD now includes the World Bank's income classification.

### Version 2025_09

**Overview** This quarterly update introduces improved government finance statistics, streamlined source handling, a new outlier detection process, and numerous fixes and small improvements.

**Improved Government Finance Statistics** The database now explicitly distinguishes between data referring to the central or general government. The GMD now contains separate series for central and government variables, as well as consolidated aggregates (which maximize data availability).

**Improved Download Infrastructure** The download infrastructure has been updated to pull data directly from primary institutions like the IMF, Eurostat, OECD, and the UN, rather than relying on dbnomics. Additionally, all IMF data collection now uses its newly released API (3.0) for better reliability.

**Pipeline Improvements** The underlying GMD processing pipeline has undergone a comprehensive overhaul to optimize performance and data handling. As a result of these structural changes, the total runtime for updating the database has improved by approximately 10x.

**New and Improved Sources** To streamline data management, various disparate datasets from the IMF and OECD are now treated as single, unified "sources" where appropriate. We have also expanded our historical coverage by adding monetary statistics for France and unemployment series from Eurostat.

**Automated Error Checking** Our automated quality control system has been expanded to detect multiple distinct error types across the entire dataset. Any values flagged as suspicious by these new automated checks are subsequently reviewed and confirmed manually by our team.

**Bug Fixes** Thanks to reports from our community of contributors, we have identified and resolved a variety of small bugs in this release. Notably, this update corrects systematic errors in the World Bank's WDI and the IMF's FPP data series.

### Version 2025_08

**Technical Documentation Updates** We have improved the technical documentation to provide a better user experience. This includes enhancing the clarity of both data source descriptions and our methodological explanations.

**Long-Term Interest Rate Updates** We updated the long-term interest rate data using the latest available information. These updates also improved the overall data quality and consistency for interest rate variables.

### Version 2025_06

The data was updated as of June 2025. In this release, we patched issues in inflation data caused by breaks in CPI series and revised the priority list accordingly. We added six new sources: Reinhart and Rogoff (2010), Internl Inst. Social History (2024), UN Statistics Division (2025), Fortin-Gagnon, Leroux, Stevanovic and Surprenant (2022), Cha, Kim, Park and Park (2022), and National Bank of Belgium (2024). Furthermore, we updated the source priority order for `nGDP`, `cons`, `inv`, `finv`, `imports`, and `exports`.

Methodologically, government finance variables (`govdebt`, `govdef`, `govexp`, `govrev`, `govtax`) are now derived from GDP ratios, after which the estimated levels series are chain-linked with other series. Finally, we improved the data cleaning pipeline to resolve minor source inconsistencies and fixed specific issues with the government deficit and current account balance data.

### Version 2025_03

Data was updated as of March 2025, and we introduced official support for Python, R, and Stata packages. Beyond a general update across all indicators, we extended historical coverage using data from Ferreres et al. (2005), Stuart (2018), and Taiwan Statistical Office (2025). We also refined the treatment of the IMF WEO within our source hierarchy and fixed inconsistencies in several long-run series, such as Australian GDP.

### Version 2025_01 - *Initial Release*

First release with date up-to-date as of January 2025. The coverage spans 46 variables, 243 countries, and annual data spanning from the 1800s to 2030 (including forecasts). The GMD integrates data from 32 major contemporary sources and 78 historical datasets, providing a unified structure for macroeconomic indicators across countries and time.

## A.5 Acknowledgments

The development of the Global Macro Database would not have been possible without the generous funding provided by the Singapore Ministry of Education (MOE) through the PYP grants (WBS A-0003319-01-00 and A-0003319-02-00), a Tier 1 grant (A-8001749-00-00), and the NUS Risk Management Institute (A-8002360-00-00). This financial support laid the foundation for the initial completion of this extensive project.

We would also like to thank the many people who have supported this project. For helpful feedback on data issues, we would like to thank Andrea Presbitero, Matteo Iacoviello, Xander Uyttenbroek, Flint O'Neil, Luis Libonatti, Jeroen Bouma, and Tom Hamburger. Shixuan Yuan provided excellent research assistance.

## A.6 Accessing the Data

The Global Macro Database (GMD) is available through multiple access methods:

### A.6.1 Web Download

Visit www.globalmacrodata.com/data to download the dataset in CSV, Excel, or Stata format.

### A.6.2 Stata Package

Install the official SSC package and retrieve data using the `gmd` command:

```
ssc install gmd, replace
gmd
```

```
gmd rGDP, country(FRA)
```

To download the package from github:

```
net install gmd, from("https://raw.githubusercontent.com/KMueller-Lab/Global-Macro-
    Database-Stata/main/stata") replace
```

Please make sure you use the most recently updated version. Type `help gmd` in Stata for full documentation. We provide more details on the extensive functionalities of the package in a dedicated companion paper (Lehbib and Müller, 2025).

One useful feature of the `gmd` function is that it also allows you to easily download and compare the raw data underlying the harmonized GMD estimates, and also includes the option to generate BibTeX citations for the GMD and all of its underlying sources.

### A.6.3 Python Package

Install the package via pip and use `gmd()` to access the data with filtering options:

```
pip install global_macro_data
from global_macro_data import gmd

df = gmd()
df = gmd(version="2025_01", country=["USA", "CHN"], variables=["rGDP", "CPI"])
```

### A.6.4 R Package

Install the R package from GitHub and access the data in the following manner:

```
install.packages("devtools")
devtools::install_github("KMueller-Lab/Global-Macro-Database-R")

library(globalmacrodata)
df <- gmd(version = "2025_01", country = c("USA", "CHN"), variables = c("rGDP", "CPI"))
```

For full setup instructions and usage examples, see www.globalmacrodata.com/data.

## A.7 Frequently Asked Questions

This section summarizes frequently asked questions about the Global Macro Database. The answers refer to the 2026–06 release unless otherwise noted.

### A.7.1 Getting started

**What is the Global Macro Database?** The Global Macro Database (GMD) is a dataset that brings together macroeconomic statistics for 239 countries and territories in a single, consistent format. It integrates 160 historical and contemporary sources, is updated quarterly, and is introduced in Müller, Xu, Lehbib and Chen (2025).

**What does the GMD cover?** The GMD covers 46 variables across national accounts, consumption and investment, the external sector, government finances, money and interest rates, and prices and labor markets. The dataset runs from 1086 to 2025, with forecasts up to 2030. Because not every country has every variable for the full period, the GMD is an unbalanced panel and coverage depends on the underlying sources.

**Is the GMD free? Can I use it commercially?** The GMD is free for academic and non-profit research, provided users attribute and cite it. Commercial use is not permitted without a separate license. If you are unsure whether your intended use is commercial, treat it as commercial and contact us at `kmueller@globalmacrodata.com`. See the Terms of Use section above for the full terms.

**How do I cite the GMD?** Please cite Müller, Xu, Lehbib and Chen (2025). See section A.3.

### A.7.2 Accessing the data

**How do I download the data?** The data can be downloaded from www.globalmacrodata.com/data in CSV, Excel, and Stata formats. The same page also links to package-based access through Stata, Python, and R.

**How do I access the GMD in Stata, Python, or R?** The Stata, Python, and R packages support filtering by version, country, and variable. The Stata package is available through SSC via `ssc install gmd`. The Python package is available via `pip install global_macro_data`. The R package can be installed from GitHub. Usage examples are provided in the Accessing the Data section above and on the GMD website.

**Can I get the raw underlying source data?** Yes. The download page and access packages provide options to return the harmonized raw data from each underlying source. These data are harmonized to make units, currencies, and identifiers comparable, but otherwise preserve the information reported by the original sources.

**How often is the GMD updated?** The GMD is updated quarterly. Each release is versioned, for example as `2026_06`. Values can change between releases because of new source data, revised source data, source-priority changes, corrections, or methodological improvements. Release notes are published on the Data page of the website.

### A.7.3 Understanding the data

**Why does the GMD sometimes differ from the IMF or World Bank?** The GMD combines many sources into a single long-run dataset. It prioritizes sources based on reliability, update frequency, and country coverage, and then harmonizes and splices them into continuous series. The resulting GMD series can therefore differ from any individual source, including the IMF or World Bank, because it reflects our effort to combine the available information consistently across countries and time.

**Why do GDP components not add up exactly to GDP?** Many GMD series are constructed by splicing multiple sources together, often using ratio chain-linking. As a result, national-accounts identities such as $Y = C + I + G + NX$ need not hold exactly. The current GMD does not force components to reconcile mechanically, so that each series remains close to the growth rates reported in the underlying sources. We plan to offer reconciled versions in future releases.

**In which units are variables expressed?** Units depend on the variable type. Level variables such as GDP, consumption, investment, trade, government revenue, government expenditure, government debt, and monetary aggregates are generally in millions of local currency units. US-dollar variants are in millions of US dollars. Ratio variables ending in `_GDP` are in percent of GDP. Rates such as inflation, unemployment, and interest rates are in percent. Index variables such as the CPI, GDP deflator, REER, and house-price index are rebased to a common reference year. Population is measured in thousands of persons.

**What base year are real and index variables expressed in?** Real and index variables are generally rebased to 2015, using the IMF World Economic Outlook as the reference where applicable. For sources without data in 2015, we use the most recently available year. This makes real and index series more comparable across countries and over time.

**Why do recent values change between versions, but historical values do not?** The GMD uses an anchor-year approach. Historical data before the anchor year are chain-linked backward from the anchor and remain fixed across releases unless an error is found. More recent data are chain-linked forward and incorporate new revisions and forecasts. This keeps long-run history stable while

allowing recent years to update.

**How are currency changes and redenominations handled?** Long-run series can span major currency reforms and redenominations. The GMD converts observations to a single consistent unit using official conversion rates where available. Examples include euro conversion rates and historical redenominations such as Turkey's pre-2005 lira. Conversions are documented in country notes where relevant.

**How are ratio variables expressed?** Ratio variables such as `govdebt_GDP`, `exports_GDP`, and `CA_GDP` are expressed in percent. A value of 50 therefore means 50% of GDP.

**Which variables are constructed directly from sources, and which are derived?** Some variables are constructed directly, meaning that the GMD harmonizes and splices observations reported by the underlying sources. Others are derived from those series using accounting identities or transformations. Directly constructed variables include many core series such as nominal and real GDP, consumption, investment, exports and imports, prices, population, monetary aggregates, interest rates, the exchange rate, and several government-finance ratios. Derived variables include many US-dollar versions, per-capita series, the GDP deflator, national-accounts ratios, consolidated government aggregates, and government-finance levels that are back-computed from share-of-GDP series. The exact status of each variable is recorded in the documentation.

**How is real GDP in US dollars calculated?** Real GDP in US dollars is not computed as a year-by-year conversion of local real GDP at current exchange rates. Instead, the GMD follows a method that preserves each country's real growth. Nominal GDP in the base year is converted to US dollars at the base-year market exchange rate, and the resulting level is extrapolated backward and forward using real GDP growth rates. This differs from nominal GDP in US dollars, which is converted year by year using the market exchange rate.

**What is the difference between `cons`, `hcons`, and `gcons`?** `cons` is total consumption and remains the headline consumption series, because historical sources often report total consumption. `hcons` is household consumption, and `gcons` is government consumption. These provide a finer breakdown where the data allow.

**What is the difference between central and general government?** Central-government variables refer to the central state. General-government variables are broader and include state and local government and social security where reported. The GMD also includes consolidated

government aggregates that maximize data availability. Users should choose the concept that best matches their research question.

### A.7.4 Data quality and methodology

**How does the GMD ensure data quality?** Every series is visually inspected. For each variable and country, the available sources are plotted together, allowing us to review level shifts, disagreements across sources, outliers, unit or currency errors, and splicing problems. Issues are documented and corrected where possible, while remaining uncertainties are flagged.

**What is splicing or chain-linking?** Splicing or chain-linking is the process of building one long series from multiple sources. Where sources overlap, the GMD often matches their growth rates rather than simply pasting raw levels together. This avoids artificial jumps in the combined series. Where sources do not overlap, a representative growth rate is used. Breaks and methods are recorded in the country notes.

**How are outliers and impossible values handled?** The GMD flags doubtful observations, including implausible outliers and values that are impossible by definition. Potential issues are checked against other sources and known economic events. Confirmed errors are corrected or removed, while genuine economic events are retained.

**Where can I see how a specific country or variable was constructed?** The documentation page links to per-country and per-variable PDFs that show the raw underlying sources, the final GMD series, splice points, and notes on adjustments for each country-variable combination. The same information can also be explored interactively on the Explore page.

### A.7.5 Contributing and contact

**I found an error or have data to contribute. What should I do?** Please email kmueller@globalmacrodata.
or open an issue on the GitHub repository. We actively fix reported issues and welcome suggestions for new data sources.

**How do I get notified about updates?** The email provided when downloading the data is used to notify users about critical updates. Release notes for every version are published on the Data page of the website.

## A.8 Variables

Table A.1 shows the key variables included in the Global Macro Database (GMD). The table provides detailed information on each variable, including the variable name, the abbreviation used in the database, units of measurement, temporal coverage, forecast horizons, and country coverage. The variables are grouped into six categories: national accounts, consumption and investments, external sectors, government finances, money and interest rates, and prices and labor market. For detailed definitions and descriptions of each variable, please refer to Section A.11.

Table A.1: Variable Overview

| Variable | Abbreviation | Derivations | Range | Forecasts | Countries |
|---|---|---|---|---|---|
| **A. National accounts** | | | | | |
| Nominal GDP | `nGDP` | USD | 1086-2030 | 5 | 229 |
| Real GDP | `rGDP` | USD, pc | 1270-2030 | 5 | 228 |
| Real GDP in USD | `rGDP_USD` | pc | 1791-2030 | 5 | 218 |
| GDP deflator | `deflator` | | 1270-2030 | 5 | 228 |
| Total consumption | `cons` | USD, GDP | 1800-2027 | 2 | 220 |
| Household consumption | `hcons` | USD, GDP | 1800-2027 | 2 | 221 |
| Government consumption | `gcons` | USD, GDP | 1800-2027 | 2 | 218 |
| Gross capital formation | `inv` | USD, GDP | 1807-2030 | 5 | 219 |
| Gross fixed capital formation | `finv` | USD, GDP | 1800-2027 | 2 | 217 |
| **B. External sectors** | | | | | |
| Current account , % GDP | `CA_GDP` | LC, USD | 1772-2030 | 5 | 206 |
| Exports | `exports` | USD, GDP | 1280-2030 | 5 | 227 |
| Imports | `imports` | USD, GDP | 1560-2030 | 5 | 227 |
| Real effective exchange rate | `REER` | | 1870-2026 | 1 | 180 |
| USD exchange rate | `USDfx` | | 1791-2030 | 5 | 235 |
| **C. Consolidated government finances** | | | | | |
| Gov. debt , % GDP | `govdebt_GDP` | LC | 1670-2030 | 5 | 198 |
| Gov. deficit , % GDP | `govdef_GDP` | LC | 1689-2030 | 5 | 199 |
| Gov. expenditure , % GDP | `govexp_GDP` | LC | 1689-2030 | 5 | 200 |
| Gov. revenue , % GDP | `govrev_GDP` | LC | 1689-2030 | 5 | 199 |
| Gov. tax revenue , % GDP | `govtax_GDP` | LC | 1789-2027 | 2 | 173 |
| **D. General government finances** | | | | | |

*Continued on next page*

*Continued from previous page*

| **Variable** | **Abbreviation** | Derivations | **Range** | **Forecasts** | **Countries** |
|---|---|---|---|---|---|
| Gov. debt , % GDP | `gen_govdebt_GDP` | LC | 1700-2030 | 5 | 196 |
| Gov. deficit , % GDP | `gen_govdef_GDP` | LC | 1700-2030 | 5 | 198 |
| Gov. expenditure , % GDP | `gen_govexp_GDP` | LC | 1800-2030 | 5 | 199 |
| Gov. revenue , % GDP | `gen_govrev_GDP` | LC | 1800-2030 | 5 | 196 |
| Gov. tax revenue , % GDP | `gen_govtax_GDP` | LC | 1798-2027 | 2 | 148 |
| **E. Central government finances** | | | | | |
| Gov. debt , % GDP | `cgovdebt_GDP` | LC | 1670-2026 | 1 | 186 |
| Gov. deficit , % GDP | `cgovdef_GDP` | LC | 1689-2026 | 1 | 168 |
| Gov. expenditure , % GDP | `cgovexp_GDP` | LC | 1689-2026 | 1 | 174 |
| Gov. revenue , % GDP | `cgovrev_GDP` | LC | 1689-2026 | 1 | 172 |
| Gov. tax revenue , % GDP | `cgovtax_GDP` | LC | 1789-2026 | 1 | 148 |
| **F. Money and interests** | | | | | |
| M0 | `M0` | | 1619-2026 | 1 | 175 |
| M1 | `M1` | | 1841-2026 | 1 | 166 |
| M2 | `M2` | | 1841-2026 | 1 | 155 |
| M3 | `M3` | | 1819-2026 | 1 | 165 |
| M4 | `M4` | | 1870-2020 | — | 4 |
| Central bank policy rate | `cbrate` | | 1694-2026 | 1 | 167 |
| Short-term interest rate | `strate` | | 1695-2026 | 1 | 136 |
| Long-term interest rate | `ltrate` | | 1727-2026 | 1 | 83 |
| **G. Prices, labor market, and population** | | | | | |
| Consumer price index | `CPI` | | 1209-2030 | 5 | 217 |
| House price index | `HPI` | | 1628-2026 | 1 | 58 |
| Real house price index | `rHPI` | | 1628-2026 | 1 | 58 |
| Inflation | `infl` | | 1210-2030 | 5 | 218 |
| Unemployment rate | `unemp` | | 1760-2030 | 5 | 220 |
| Population | `pop` | | 1277-2030 | 5 | 235 |
| **H. Financial crisis dummy** | | | | | |
| Banking crisis | `BankingCrisis` | | 1800-2020 | — | 163 |
| Sovereign debt crisis | `SovDebtCrisis` | | 1800-2020 | — | 160 |

| Currency crisis | `CurrencyCrisis` | 1800-2019 — | 160 |
|---|---|---|---|

**Note:** This table presents the variables included in the Global Macro Database (GMD), along with their abbreviations/variable names, available derivations, temporal coverage, forecast horizons, and country coverage. Except for variables explicitly in USD, all variables in levels are recorded in millions of local currency units. Indices (CPI, HPI) are recorded such that 2015 = 100. Interest rates, inflation, and unemployment are recorded in percent (e.g., 5 means 5%). Derivations refer to other versions of the variables available in the data, where "LC" stands for local currency (instead of % of GDP), "USD" for US dollars, and "GDP" for a fraction of GDP.

### A.8.1 Banking crisis onset indicator

The empirical analysis uses a country-year banking crisis onset indicator constructed from the principal crisis chronologies in the literature: Reinhart and Rogoff (2009*b*), Laeven and Valencia (2013), Jordà, Schularick and Taylor (2017*a*), and Baron, Verner and Xiong (2021). The goal is to convert these source datasets into a single event-level series that can be used in local projections over a long historical panel. The construction preserves two pieces of information from the underlying sources: whether a country-year is covered by at least one chronology, and whether a covered country-year belongs to the onset of a banking crisis episode.

We first merge the source indicators by country and year. For each country-year, the construction records crisis observations, explicit non-crisis observations, and years that fall outside the coverage of the source chronologies. This distinction determines the estimation sample. The baseline indicator is coded as zero only when at least one source explicitly covers the country-year as a non-crisis observation. Country-years outside source coverage remain missing.

The onset indicator is constructed at the episode level. The implementation begins with candidate crisis years reported by Laeven and Valencia (2013) or Baron, Verner and Xiong (2021). It then uses the multi-year crisis spells in Reinhart and Rogoff (2009*b*) to group candidate dates that fall within the same underlying episode. When a Reinhart-Rogoff spell contains multiple candidate years, the first candidate year is retained. When a Reinhart-Rogoff spell contains no Laeven-Valencia or Baron-Verner-Xiong candidate year, the first year of the Reinhart-Rogoff spell is used as the onset. This converts multi-year crisis records into one-year onset observations.

The Jordà-Schularick-Taylor chronology enters as an additional source of coverage and cross-validation: JST crisis dates are used to assess whether the baseline onset indicator misses episodes identified in the long-run macro-financial literature, and source-specific JST indicators are retained for the robustness exercises below.

The dating rule also addresses adjacent candidate years outside Reinhart-Rogoff spells. Candidate onsets that are one year apart and are not part of a Reinhart-Rogoff episode are treated as

a single event, with the earlier year retained. For adjacent candidate years in which one year falls within a Reinhart-Rogoff spell and the other does not, the construction counts support from the Laeven-Valencia, Baron-Verner-Xiong, and Reinhart-Rogoff indicators. The onset is assigned to the year with greater source support, and ties are assigned to the earlier year.

The construction also produces alternative indicators used in the robustness exercises. These include source-specific crisis onset indicators, union-based indicators that count a crisis when any major source identifies one, an indicator for bank-run episodes that do not coincide with banking-crisis observations in the other sources, and a union of the Reinhart-Rogoff, Jordà-Schularick-Taylor, and Baron-Verner-Xiong chronologies. The results in Section 3 compare the baseline onset indicator with these alternatives.

### A.8.2 Data collection

The GMD is based on two broad categories of sources:

1. *Modern sources*: Datasets from major international organizations (e.g., Eurostat, OECD) that are continuously updated. The GMD retrieves these via an automated code pipeline and APIs to ensure continuous updates.

2. *Historical sources*: Academic work or datasets from national authorities (e.g., Mitchell (2013) or Inklaar, de Jong, Bolt and van Zanden (2018*a*)) that are updated infrequently or cannot be accessed through our automated pipeline. Integrating these sources often requires manual digitization and extensive harmonization of data from archival records.

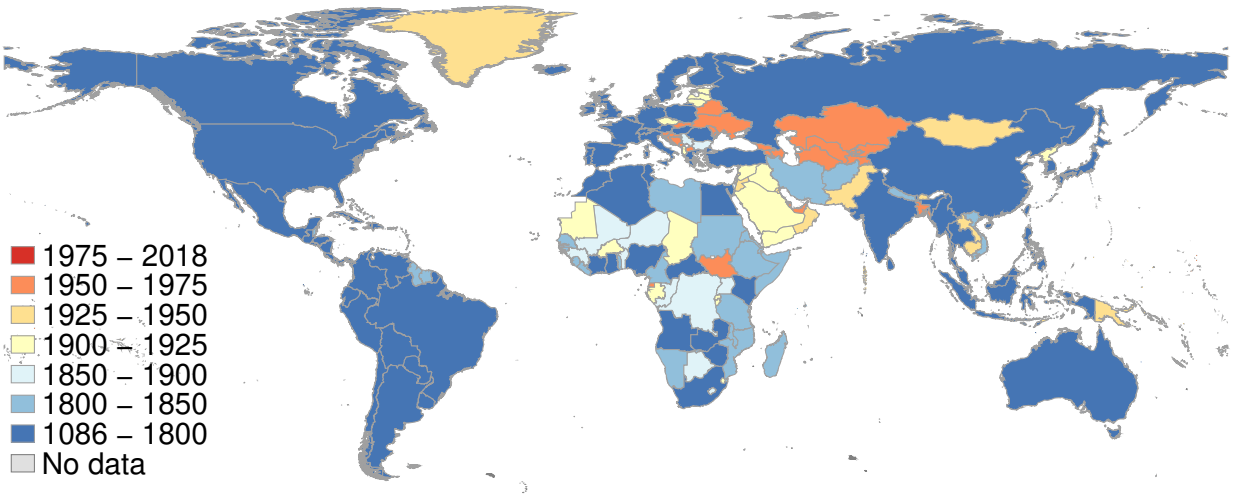
## A.9 Data Coverage

The Global Macro Database comprises a collection of 46 macroeconomic variables, spanning 239 countries and territories. The majority of countries have time series data extending to pre-1950 periods, with many reaching back to the early $20^{\text{th}}$ century or even the $19^{\text{th}}$ century, including several developing nations. Figure A.1 presents a world map indicating the earliest year for which any macroeconomic statistic is available for each country (excluding population data).

### A.9.1 Dataset overview

Figure A.2 illustrates the number of countries covered by the Global Macro Database (GMD) for each variable in the years 1900, 1980, and 2008. A significant proportion of the variables have coverage for over 20 countries since 1900, over 100 countries since 1980, and nearly 200 countries since 2008, reflecting the increasing scope and comprehensiveness of the database over time. Figure A.3 depicts the share of GDP for which data is available on key variables in the GMD.

Figure A.1: First Year in Dataset, By Country



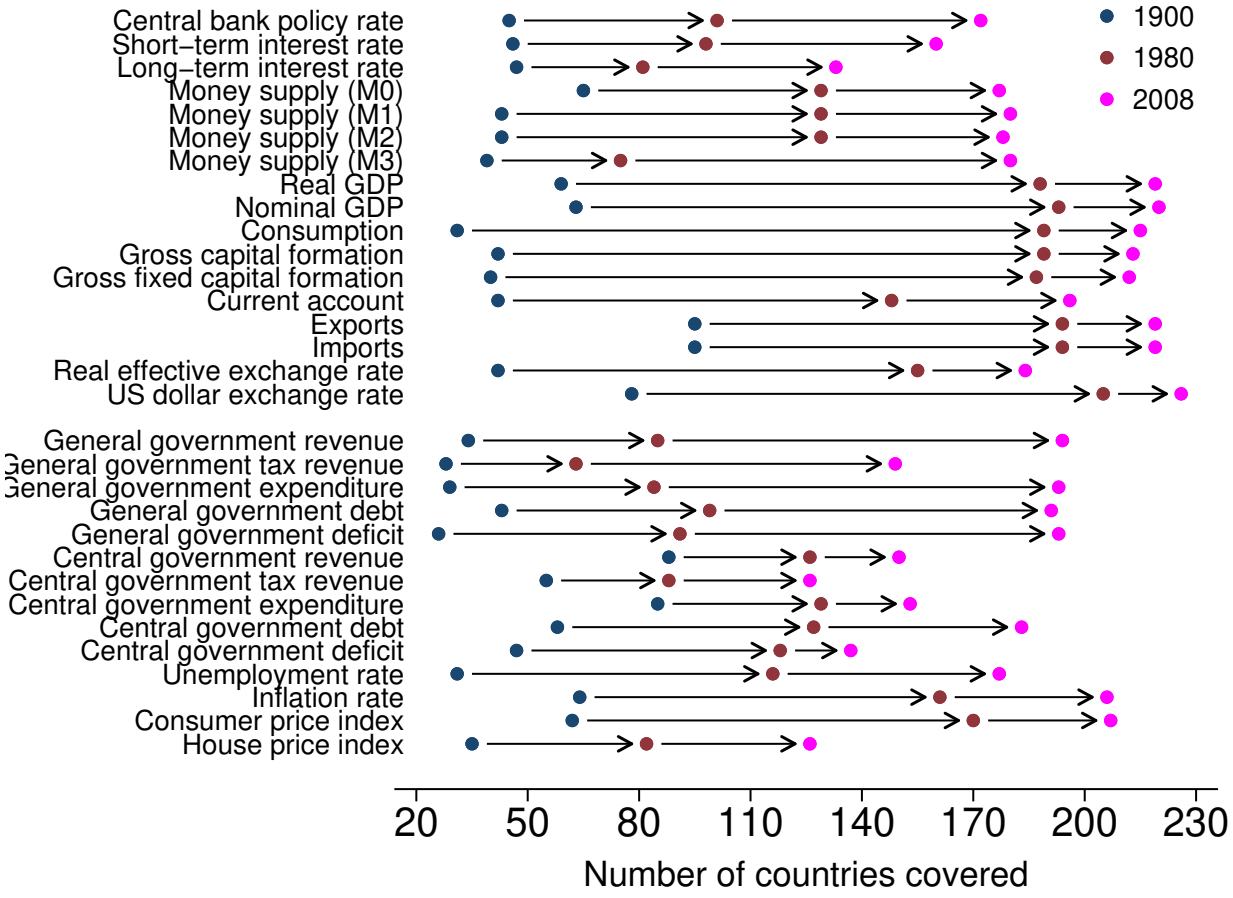

**Note:** This figure visualizes, for each country, the first year for which we have data on any macroeconomic variable other than population in the Global Macro Database (GMD). The map demonstrates that for the vast majority of countries, we have time series before 1950, with many series extending back to the 1800s for both advanced and developing economies. For some countries, such as the United Kingdom, data coverage begins as early as 1086.

### A.9.2 Comparison with other sources

By design, the coverage of our dataset surpasses that of all existing publicly available sources we are aware of, as these sources have been integrated into the Global Macro Database (GMD). Table A.2 compares the coverage of key variables in the GMD with those offered by other widely-used data providers. Table A.3 presents the number of country-year observations included in the GMD and the fraction covered by the major providers in percentage. Figure A.4 compares the coverage of the Global Macro Database (GMD) with that of the next best source for each key variable.

In addition to publicly available sources, Table A.2 also includes a comparison with the dataset provided by the commercial data provider Global Financial Data (GFD). Beyond surpassing the coverage of many commercial products, the Global Macro Database (GMD) offers more comprehensive and transparent documentation, supported by an open-source code base. For further details, please refer to our GitHub repository.

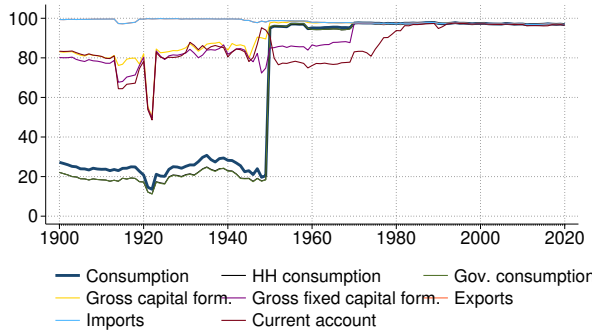
Figure A.2: Number of Countries Covered, By Variable



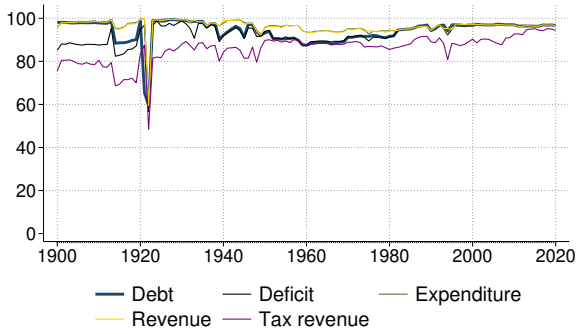

**Note:** This figure plots the number of countries with available data for all macroeconomic variables contained in the Global Macro Database (GMD) at three points in time: 1900, 1980, and 2008. Each line represents a variable, with dots indicating the coverage at these three dates. The graph reveals substantial variation in coverage across variables and time.

Figure A.3: Share of GDP Covered Over Time, By Variable

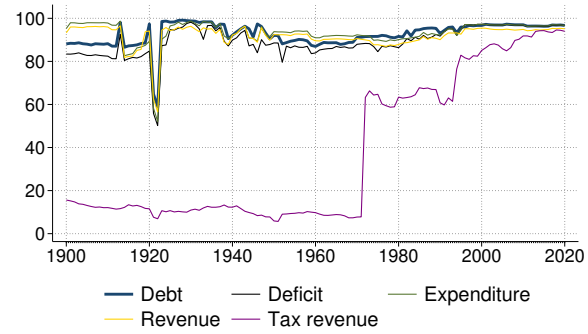
(a) National Accounts



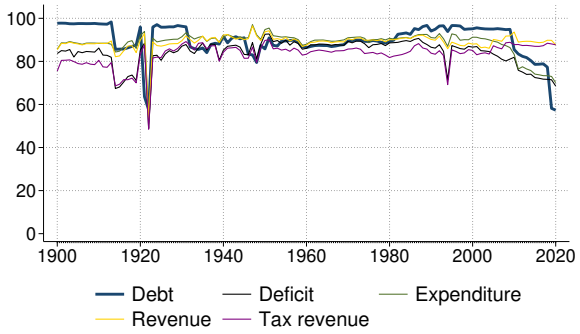

(b) Combined Government Finances



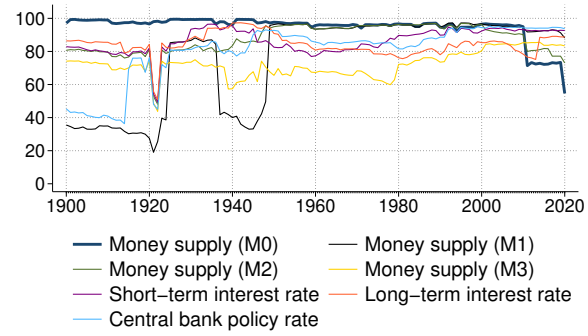

(c) General Government Finances



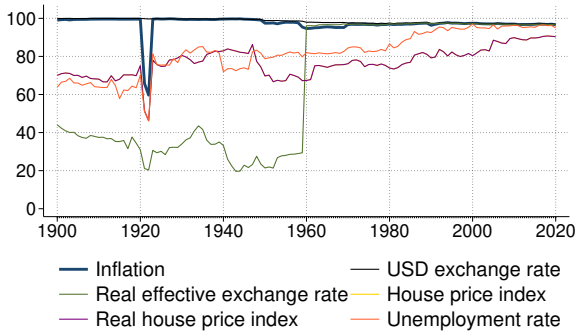

(d) Central Government Finances




(e) Money and Interest Rates




(f) Prices and Labor



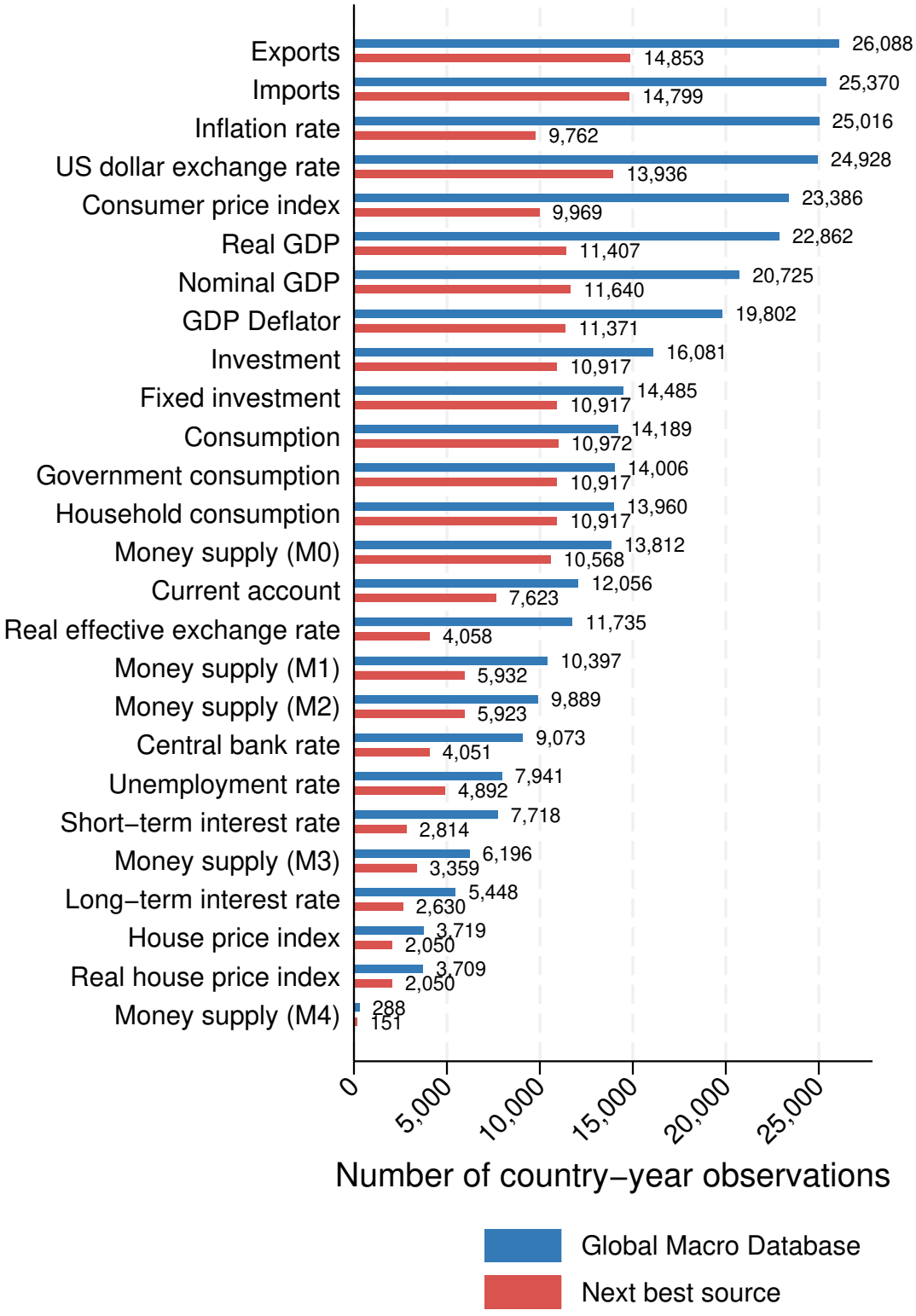

**Note:** This figure plots the share of GDP covered by each variable in the Global Macro Database (GMD) between 1900 and 2020.

Table A.2: Comparing the Coverage of Key Macroeconomic Variables

| Source | Start Year | | Latest | | Countries | | Variables |
|---|---|---|---|---|---|---|---|
| | First | Median | Actual | Forecast | Number | Year-Obs. | |
| GMD | 1086 | 1800 | 2025 | 2030 | 239 | 56,030 | 80 |
| GFD | 1000 | 1820 | 2024 | — | 236 | 39,345 | 37 |
| IFS | 1914 | 1950 | 2025 | 2025 | 223 | 14,813 | 27 |
| WEO | 1960 | 1970 | 2025 | 2030 | 209 | 11,417 | 26 |
| JST | 1870 | 1870 | 2020 | — | 18 | 2,666 | 31 |
| Maddison | 1253 | 1876 | 2022 | — | 169 | 20,590 | 2 |
| OECD EO | 1960 | 1960 | 2025 | 2026 | 49 | 2,837 | 37 |
| PWT | 1950 | 1960 | 2023 | — | 212 | 12,521 | 16 |
| UN | 1900 | 1970 | 2024 | — | 215 | 11,931 | 22 |
| WDI | 1960 | 1960 | 2024 | — | 220 | 13,792 | 53 |

**Note:** This table compares the coverage of the key macroeconomic variables included in our dataset across a selected sample of widely-used data providers. See Table A.1 for the list of variables, which also includes derived measures such as GDP-scaled variables. The number of variables refers to the subset of macroeconomic indicators we cover that are available in a given source, not the total number of variables available. The acronyms in the source column refer to the following datasets: GFD = Global Financial Data, IFS = IMF International Financial Statistics, WEO = IMF World Economic Outlook, JST = Jordà-Schularick-Taylor Macro History Database, MAD = Maddison Historical Statistics, OECD EO = OECD Economic Outlook, PWT = Penn World Table, UN = United Nations Statistics, WDI = World Bank World Development Indicators. "Year-Obs." under the "Countries" header refers to the total number of country-year observations with non-missing information on any of the variables. Note that, for the Maddison dataset, we keep only continuous observations which started in 1253.

Figure A.4: Comparing Dataset Coverage by Variable




**Note:** This figure compares, for each variable, the coverage of the Global Macro Database (GMD) with that of the next most comprehensive source. Blue bars show the number of country-year observations in the Global Macro Database, and the red bars show the number of country-year observations for the next best source.

Table A.3: Coverage of GMD Variables in Selected Sources

| | | Fraction of observations in the GMD covered in... (in %) | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| **Variable** | **GMD** | IFS* | WEO | OECD EO | WDI† | UN | JST | Mitchell | GFD‡ |
| Central bank policy rate | 9,073 | 45 | — | 18 | — | — | — | — | 71 |
| Short-term interest rate | 7,718 | 36 | — | 28 | — | — | 33 | — | 54 |
| Long-term interest rate | 5,448 | 36 | — | 36 | — | — | 48 | — | 133 |
| Money supply (M0) | 13,812 | 14 | — | — | — | — | 3 | 77 | 97 |
| Money supply (M1) | 10,397 | — | — | 13 | — | — | 19 | 57 | 44 |
| Money supply (M2) | 9,889 | — | — | — | — | — | 15 | 60 | — |
| Money supply (M3) | 6,196 | 54 | — | 20 | — | — | 13 | — | — |
| Real GDP | 22,862 | 23 | 35 | 11 | 50 | 47 | 12 | 28 | 86 |
| Nominal GDP | 20,725 | 28 | 38 | 12 | 56 | 53 | 13 | 36 | 79 |
| Consumption | 14,189 | 23 | — | 16 | 61 | 77 | — | — | — |
| Household consumption | 13,960 | — | — | 17 | 61 | 78 | — | — | — |
| Government consumption | 14,006 | 34 | — | 17 | 61 | 78 | — | — | — |
| Gross capital formation | 16,081 | 27 | 42 | 14 | 54 | 68 | 15 | 33 | 27 |
| Gross fixed capital formation | 14,485 | 31 | — | 16 | 56 | 75 | — | 31 | 27 |
| Current account | 12,056 | 37 | 63 | 17 | 63 | — | 21 | — | 23 |
| Exports | 26,088 | 18 | 34 | 9 | 35 | 42 | 10 | 57 | 70 |
| Imports | 25,370 | 18 | 36 | 9 | 36 | 43 | 10 | 58 | 72 |
| Real effective exchange rate | 11,735 | 35 | — | 8 | 35 | — | — | — | 38 |
| US dollar exchange rate | 24,928 | 56 | 31 | 11 | 49 | 44 | 11 | — | 90 |
| Combined government debt | 14,858 | — | — | — | — | — | — | — | 12 |
| Combined government deficit | 13,659 | — | — | — | — | — | — | — | 93 |
| General government debt | 13,464 | 86 | 42 | 6 | — | — | 18 | — | — |
| General government deficit | 11,068 | 79 | 58 | 16 | — | — | — | — | — |
| General government expenditure | 11,932 | 81 | 54 | 14 | — | — | — | — | — |
| General government revenue | 11,642 | 83 | 55 | 14 | — | — | — | — | — |
| General government tax revenue | 3,823 | 74 | — | — | — | — | — | — | — |
| Central government debt | 12,083 | 61 | — | — | 17 | — | — | — | — |
| Central government deficit | 9,644 | 26 | — | — | — | — | — | 57 | — |
| Central government expenditure | 11,026 | 23 | — | — | — | — | 23 | 55 | — |
| Central government revenue | 10,785 | 25 | — | — | — | — | 23 | 54 | — |
| Central government tax revenue | 6,790 | 40 | — | — | — | — | — | 42 | — |
| Unemployment rate | 7,941 | 62 | 53 | 26 | 37 | — | 25 | — | 64 |
| Inflation rate | 25,016 | 39 | 31 | 8 | 36 | — | 11 | 32 | 77 |
| Consumer price index | 23,386 | 43 | 34 | 6 | 39 | — | 11 | 35 | 77 |
| House price index | 3,719 | — | — | — | — | — | 55 | — | 30 |
| Real house price index | 3,709 | — | — | — | — | — | 55 | — | — |

**Note:** This table shows the number of country-year observations in the Global Macro Database (GMD) and the fraction covered in major providers of macroeconomic data (in percent). Our dataset begins in the year 1086.
* IFS includes data from the IMF GFS datasets. † Unemployment data from the World Bank is added from International Labour Organization (ILO). ‡ The GFD coverage is based on the subset of data currently available to us and may not represent the entirety of the GFD dataset.

## A.10 Sources and Data Access

### A.10.1 Types of data sources

Our dataset combines information from 160 distinct sources that can be categorized along two primary dimensions. The first dimension is update frequency. Current sources (33) are continuously updated with regular release calendars, primarily from major international organizations and statistical agencies. Historical sources (127) are either never or infrequently updated without a regular schedule.

The second dimension is coverage. The 76 aggregators we include provide data for multiple countries (e.g., IMF, World Bank). The 84 country-specific sources focus on individual countries, and are national statistical offices or academic research.

Table A.5 provides a comprehensive overview of all sources currently incorporated in the Global Macro Database (GMD), along with detailed information about each source. This includes the source reference, the abbreviation used in the GMD, the latest update date, an indication of whether the source is newly digitized, temporal coverage, variable coverage, country coverage, and whether it is classified as a historical source (i.e., a dataset that is not updated or is only infrequently updated).

Table A.4: Types of Sources in the Global Macro Database

| | Current | Historical | **Total** |
|---|---|---|---|
| Aggregators | 21 | 55 | **76** |
| *Examples* | *WEO* | *JST* | |
| Country-specific | 12 | 72 | **84** |
| *Examples* | *FRED* | *Thomas, Hills and Dimsdale (2010)* | |
| **Total** | **33** | **127** | **160** |

**Note:** This table plots information on the number of sources used in the Global Macro Database. We differentiate sources along two dimensions: (1) whether they are continuously updated (*current*) or contain only historical data (*historical*), and (2) whether they report information on several countries (*aggregators*) or only a single country (*country-specific*). Note that we count as historical sources those that have been updated on an ad-hoc basis but do not have a clear release calendar, such as the Jordá, Schularick and Taylor (2017*b*) Macro History Database.

Table A.5: Dataset Overview

| # | Source | Abbreviation | Updated | Digitized | From | To | Forecasts | Variables | Countries | Historical |
|---|---|---|---|---|---|---|---|---|---|---|
| **Panel A: Aggregator Sources** | | | | | | | | | | |
| 1 | Arroyo Abad, Lindert et al. (2015) | AAL | | No | 1870 | 2013 | — | 5 | 5 | Yes |
| 2 | Asian Development Bank (2024) | ADB | | No | 2000 | 2023 | — | 28 | 49 | Yes |
| 3 | African Development Bank (2018) | AFDB | 2026-07-01 | No | 1980 | 2019 | — | 10 | 53 | No |
| 4 | African Union (2024) | AFRISTAT | | No | 1990 | 2023 | — | 11 | 22 | Yes |
| 5 | IEC, Hitotsubashi University (2020) | AHSTAT | | No | 1860 | 2013 | — | 30 | 5 | Yes |
| 6 | European Commission (2024) | AMECO | 2026-07-01 | No | 1960 | 2027 | — | 30 | 27 | No |
| 7 | Arab Monetary Fund (2024) | AMF | | No | 1971 | 2021 | — | 26 | 22 | Yes |
| 8 | Andersson (2026) | ANDERSSON | | No | 1800 | 2024 | — | 5 | 29 | Yes |
| 9 | Barro and Ursúa (2008*c*) | BARRO | | No | 1790 | 2009 | — | 1 | 42 | Yes |
| 10 | BCEAO (2024) | BCEAO | 2026-07-01 | No | 1960 | 2024 | — | 29 | 8 | No |
| 11 | Broadberry and Gardner (2022) | BG | | No | 1885 | 2008 | — | 2 | 8 | Yes |
| 12 | BIS (2025*c*) | BIS CPI | 2026-07-01 | No | 1661 | 2025 | — | 0 | 62 | No |
| 13 | BIS (2025*e*) | BIS HPI | 2026-07-01 | No | 1927 | 2026 | — | 0 | 57 | No |
| 14 | BIS (2025*d*) | BIS REER | 2026-07-01 | No | 1994 | 2026 | — | 0 | 63 | No |
| 15 | BIS (2025*a*) | BIS USDfx | 2026-07-01 | No | 1791 | 2025 | — | 0 | 189 | No |
| 16 | BIS (2025*b*) | BIS cbrate | 2026-07-01 | No | 1945 | 2026 | — | 0 | 48 | No |
| 17 | Banca d'Italia (2024) | BIT | | No | 1955 | 2024 | — | 1 | 9 | Yes |

*Continued on next page*

*Continued from previous page*

| # | Source | Abbreviation | Updated | Digitized | From | To | Forecasts | Variables | Countries | Historical |
|---|---|---|---|---|---|---|---|---|---|---|
| 18 | Bordo, Eichengreen, Klingebiel and Martinez-Peria (2001) | BORDO | | No | 1880 | 1997 | — | 14 | 56 | Yes |
| 19 | Darvas (2021) | BRUEGEL | | No | 1960 | 2023 | — | 1 | 178 | Yes |
| 20 | Baron, Verner and Xiong (2020) | BVX | | No | 1870 | 2016 | — | 7 | 48 | Yes |
| 21 | Fouquin and Hugot (2016) | CEPII | 2026-07-01 | No | 1827 | 2014 | — | 6 | 216 | Yes |
| 22 | Internl Inst. Social History (2024) | CLIO | | No | 1500 | 2011 | — | 3 | 175 | Yes |
| 23 | Banushi et al. (2024) | COMECON | | No | 1945 | 1993 | — | 21 | 9 | Yes |
| 24 | Cogneau, Dupraz and Mesplé-Somps (2021) | Cogneau Dupraz | | No | 1833 | 1970 | — | 15 | 6 | Yes |
| 25 | Mack and Martínez-García (2011) | DallasFED | | No | 1975 | 2023 | — | 2 | 26 | Yes |
| 26 | Univ. of California, Davis (2022) | Davis | | No | 1818 | 2012 | — | 1 | 55 | Yes |
| 27 | United Nations (2024*a*) | ECLAC | | No | 1950 | 2023 | — | 21 | 34 | Yes |
| 28 | Eurostat (2025) | EUS | 2026-07-01 | No | 1949 | 2025 | — | 39 | 44 | No |
| 29 | FAO (2024) | FAO | | No | 1970 | 2023 | — | 5 | 214 | Yes |
| 30 | Univ. of California, Davis (2024) | FLORA | | No | 1790 | 1975 | — | 7 | 13 | Yes |
| 31 | Banque de France (2024*b*) | FRANC ZONE | | No | 1991 | 2019 | — | 7 | 15 | Yes |
| 32 | Flandreau and Zumer (2004) | FZ | | No | 1880 | 1913 | — | 15 | 16 | Yes |
| 33 | Smits, Woltjer and Ma (2009) | GNA | | No | 1800 | 2005 | — | 2 | 6 | Yes |
| 34 | Gapminder (2024) | Gapminder | | No | 1800 | 2030 | 7 | 1 | 197 | Yes |
| 35 | Grimm (2024) | Grimm | | No | 1945 | 2023 | — | 1 | 166 | Yes |
| 36 | Schuler (2015) | HFS | | No | 1800 | 2008 | — | 38 | 63 | Yes |

*Continued on next page*

*Continued from previous page*

| # | Source | Abbreviation | Updated | Digitized | From | To | Forecasts | Variables | Countries | Historical |
|---|---|---|---|---|---|---|---|---|---|---|
| 37 | Homer and Sylla (1996) | Homer Sylla | | No | 1798 | 1989 | — | 3 | 26 | Yes |
| 38 | Ellison, Lee and O'Rourke (2024) | IHD | | No | 1925 | 1936 | — | 11 | 37 | Yes |
| 39 | Internl Labour Organization (2024) | ILO | | No | 2000 | 2023 | — | 1 | 215 | Yes |
| 40 | Mauro, Romeu, Binder and Zaman (2015) | IMF FPP | | No | 1800 | 2023 | — | 4 | 151 | Yes |
| 41 | Mbaye, Moreno-Badia and Chae (2018) | IMF GDD | | No | 1950 | 2018 | — | 5 | 185 | Yes |
| 42 | International Monetary Fund (2024) | IMF GFS | 2026-07-01 | No | 1972 | 2025 | — | 16 | 140 | No |
| 43 | International Monetary Fund (2010) | IMF HDD | | No | 1800 | 2015 | — | 1 | 188 | Yes |
| 44 | International Monetary Fund (2025*a*) | IFS | 2026-07-01 | No | 1914 | 2025 | — | 27 | 223 | No |
| 45 | International Monetary Fund (2025*b*) | WEO | 2026-07-01 | No | 1960 | 2030 | 5 | 26 | 209 | No |
| 46 | Albers, Jerven and Suesse (2023) | JERVEN | | No | 1900 | 2015 | — | 3 | 46 | Yes |
| 47 | Jones and Obstfeld. (1997) | JO | | No | 1850 | 1945 | — | 7 | 13 | Yes |
| 48 | Jordá, Schularick and Taylor (2017*b*) | JST | | No | 1870 | 2020 | — | 31 | 18 | Yes |
| 49 | Ljungberg (2019) | LUND | | No | 1870 | 2016 | — | 1 | 27 | Yes |
| 50 | Laeven and Valencia (2013*b*) | LV | | No | 1970 | 2017 | — | 4 | 155 | Yes |
| 51 | Kehoe and Nicolini (2022) | MAFHOLA | | No | 1960 | 2017 | — | 10 | 11 | Yes |
| 52 | Cox and Dincecco (2021) | MD | | No | 1703 | 1913 | — | 5 | 2 | Yes |
| 53 | Officer and Williamson (2024) | MW | | No | 1209 | 2024 | — | 10 | 41 | Yes |
| 54 | Inklaar, de Jong, Bolt and van Zanden (2018*b*) | Maddison | | No | 1253 | 2022 | — | 2 | 169 | Yes |
| 55 | Mitchell (2013) | Mitchell | | Yes | 1750 | 2010 | — | 27 | 142 | Yes |

*Continued on next page*

*Continued from previous page*

| # | Source | Abbreviation | Updated | Digitized | From | To | Forecasts | Variables | Countries | Historical |
|---|---|---|---|---|---|---|---|---|---|---|
| 56 | Bértola and Rey (2018) | Moxlad | | No | 1870 | 2010 | — | 6 | 20 | Yes |
| 57 | National Bank of Serbia (2024) | NBS | | No | 1833 | 1950 | — | 27 | 8 | Yes |
| 58 | OECD (2024*a*) | OECD EO | 2026-07-01 | No | 1960 | 2026 | — | 37 | 49 | No |
| 59 | OECD (2024*b*) | OECD HPI | 2026-07-01 | No | 1960 | 2025 | — | 2 | 47 | No |
| 60 | OECD (2024*c*) | OECD KEI | 2026-07-01 | No | 1914 | 2025 | — | 7 | 47 | No |
| 61 | OECD (2024*d*) | OECD MEI | 2026-07-01 | No | 1951 | 2025 | — | 5 | 37 | No |
| 62 | OECD (1986) | OECD MEI ARC | | Yes | 1955 | 1984 | — | 3 | 18 | Yes |
| 63 | Feenstra, Inklaar and Timmer (2015) | PWT | | No | 1950 | 2023 | — | 16 | 212 | Yes |
| 64 | Piketty and Zucman (2014) | Piketty Zucman | | No | 1791 | 2011 | — | 4 | 8 | Yes |
| 65 | Reinhart and Rogoff (2009*a*) | RR | | No | 1719 | 2016 | — | 9 | 72 | Yes |
| 66 | Consejo Monetario Centroamericano (2026) | SECMCA | 2026-07-01 | No | 1960 | 2025 | — | 51 | 6 | Yes |
| 67 | Consejo Monetario Centroamericano (2017) | SECMCA HIST | | Yes | 1960 | 2018 | — | 35 | 6 | Yes |
| 68 | Albers (2018) | TH ID | | No | 1925 | 1936 | — | 2 | 27 | Yes |
| 69 | Federico and Tena-Junguito (2019) | Tena | | No | 1800 | 1938 | — | 6 | 150 | Yes |
| 70 | Federico and Tena-Junguito (2025) | Tena Pop | | No | 1800 | 1938 | — | 0 | 143 | Yes |
| 71 | United Nations (2024*b*) | UN | 2026-07-01 | No | 1900 | 2024 | — | 22 | 215 | No |
| 72 | UN Statistics Division (2025) | UN Trade | | Yes | 1900 | 1960 | — | 2 | 22 | Yes |
| 73 | World Bank (2026) | WB GEM | 2026-07-01 | No | 1987 | 2026 | — | 7 | 195 | No |
| 74 | Ha, Kose and Ohnsorge (2023) | WB infl | | No | 1970 | 2023 | — | 3 | 205 | Yes |

*Continued on next page*

*Continued from previous page*

| # | Source | Abbreviation | Updated | Digitized | From | To | Forecasts | Variables | Countries | Historical |
|---|---|---|---|---|---|---|---|---|---|---|
| 75 | World Bank (2024) | WDI | 2026-07-01 | No | 1960 | 2024 | — | 53 | 221 | No |
| 76 | World Bank (1999) | WDI ARC | | No | 1960 | 1997 | — | 21 | 208 | Yes |
| **Panel B: Country Specific Sources** | | | | | | | | | | |
| 77 | Nakamura and Zarazaga (2001) | ARG 1 | | Yes | 1901 | 1935 | — | 0 | 1 | Yes |
| 78 | Ministerio de Economia (2024) | ARG 2 | | No | 1940 | 2023 | — | 0 | 1 | Yes |
| 79 | Ferreres et al. (2005) | ARG 3 | | No | 1810 | 2018 | — | 0 | 1 | Yes |
| 80 | Hutchinson and Ploeckl (2024) | AUS 1 | | No | 1789 | 2020 | — | 0 | 1 | Yes |
| 81 | Vamplew (1987) | AUS 2 | | Yes | 1788 | 1917 | — | 0 | 1 | Yes |
| 82 | Foster (1996) | AUS 3 | | No | 1950 | 1997 | — | 0 | 1 | Yes |
| 83 | Australian Bur. of Statistics (2008) | AUS 4 | 2026-07-01 | Yes | 1788 | 2007 | — | 0 | 1 | Yes |
| 84 | Schulze (2000) | AUT 1 | | No | 1870 | 1913 | — | 0 | 1 | Yes |
| 85 | Hubmann, Jobst and Maier (2020) | AUT 2 | | No | 1801 | 2018 | — | 0 | 1 | Yes |
| 86 | Ivanov (2012) | BGR 1 | | Yes | 1870 | 1945 | — | 0 | 1 | Yes |
| 87 | Herranz-Loncan and Peres-Cajias (2016) | BOL 1 | | No | 1890 | 2010 | — | 0 | 1 | Yes |
| 88 | Peres-Cajias (2014) | BOL 2 | | No | 1882 | 2010 | — | 0 | 1 | Yes |
| 89 | IPEA (2024) | BRA 1 | | No | 1872 | 2023 | — | 0 | 1 | Yes |
| 90 | Statistics Canada (2024) | CAN 1 | | No | 1867 | 1977 | — | 0 | 1 | Yes |
| 91 | Fortin-Gagnon, Leroux, Stevanovic and Surprenant (2022) | CAN 2 | | No | 1914 | 2025 | — | 0 | 1 | Yes |

*Continued on next page*

*Continued from previous page*

| # | Source | Abbreviation | Updated | Digitized | From | To | Forecasts | Variables | Countries | Historical |
|---|---|---|---|---|---|---|---|---|---|---|
| 92 | Swiss National Bank (2009) | CHE 1 | | No | 1838 | 2005 | — | 0 | 1 | Yes |
| 93 | Hist. Statistics of Switzerland (2012) | CHE 2 | | No | 1851 | 1992 | — | 0 | 1 | Yes |
| 94 | Natl Bureau of Statistics, China (2024) | CHN 1 | | No | 1949 | 2023 | — | 0 | 1 | Yes |
| 95 | Xu et al. (2015) | CHN 3 | | Yes | 1661 | 1933 | — | 0 | 1 | Yes |
| 96 | Banco de la República, Colombia (2024) | COL 1 | 2026-07-01 | No | 1950 | 2025 | — | 0 | 1 | No |
| 97 | Banco de la República, Colombia (2026) | COL 2 | | No | 1919 | 2025 | — | 0 | 1 | Yes |
| 98 | Deutsche Bundesbank (1998) | DEU 1 | | Yes | 1948 | 1997 | — | 0 | 1 | Yes |
| 99 | Deutsche Bundesbank (1976) | DEU 2 | | Yes | 1875 | 1945 | — | 0 | 1 | Yes |
| 100 | Sprenger (1982) | DEU 3 | | Yes | 1835 | 1998 | — | 0 | 1 | Yes |
| 101 | Statistisches Reichsamt (1925) | DEU 4 | | Yes | 1913 | 1923 | — | 0 | 1 | Yes |
| 102 | Holtfrerich (1980) | DEU 5 | | Yes | 1910 | 1923 | — | 0 | 1 | Yes |
| 103 | Sensch (2012) | DEU 6 | | Yes | 1881 | 1977 | — | 0 | 1 | Yes |
| 104 | Ritschl (2002) | DEU 7 | | No | 1913 | 1943 | — | 0 | 1 | Yes |
| 105 | Ritschl and Spoerer (1997) | DEU 8 | | No | 1901 | 1944 | — | 0 | 1 | Yes |
| 106 | Abildgren (2017) | DNK 1 | | No | 1487 | 2023 | — | 0 | 1 | Yes |
| 107 | Bank of Algeria (2023) | DZA 1 | | No | 1974 | 2023 | — | 0 | 1 | Yes |
| 108 | Inst. Nacional de Estadística (2024) | ESP 1 | 2026-07-01 | No | 1995 | 2023 | — | 0 | 1 | No |
| 109 | Banco de España (2024) | ESP 2 | | No | 1277 | 2014 | — | 0 | 1 | Yes |
| 110 | Martín-Aceña (2018) | ESP 3 | | Yes | 1830 | 1998 | — | 0 | 1 | Yes |

*Continued on next page*

*Continued from previous page*

| # | Source | Abbreviation | Updated | Digitized | From | To | Forecasts | Variables | Countries | Historical |
|---|---|---|---|---|---|---|---|---|---|---|
| 111 | Banque de France (2024*a*) | FRA 1 | 2026-07-01 | No | 1970 | 2026 | — | 0 | 1 | No |
| 112 | Levy-Garboua and Monnet (2016) | FRA 2 | | No | 1800 | 2015 | — | 0 | 1 | Yes |
| 113 | Ngakegni (1991) | FRA 3 | | Yes | 1876 | 1990 | — | 0 | 1 | Yes |
| 114 | Villa (1994) | FRA 4 | | No | 1815 | 1994 | — | 0 | 1 | Yes |
| 115 | Thomas, Hills and Dimsdale (2010) | GBR 1 | | No | 1086 | 2016 | — | 0 | 1 | Yes |
| 116 | HM Treasury (2026) | GBR 2 | | No | 1700 | 2022 | — | 0 | 1 | Yes |
| 117 | Feinstein (1972) | GBR 3 | | Yes | 1855 | 1965 | — | 0 | 1 | Yes |
| 118 | Sefton and Weale (1995) | GBR 4 | | Yes | 1920 | 1990 | — | 0 | 1 | Yes |
| 119 | Abdulai (2022) | GHA 1 | | No | 1961 | 2017 | — | 0 | 1 | Yes |
| 120 | Hong Kong Institute (2024) | HKG 1 | | No | 1843 | 2002 | — | 0 | 1 | Yes |
| 121 | Bank Indonesia (2023) | IDN 1 | 2026-07-01 | No | 2008 | 2018 | — | 0 | 1 | No |
| 122 | Sivasubramonian (2000) | IND 1 | | Yes | 1900 | 1946 | — | 0 | 1 | Yes |
| 123 | Government of India (1922) | IND 2 | | Yes | 1840 | 1919 | — | 0 | 1 | Yes |
| 124 | Stuart (2018) | IRL 1 | | No | 1933 | 2014 | — | 0 | 1 | Yes |
| 125 | Jónsson, Magnússon and Snorrason (1997) | ISL 1 | | Yes | 1870 | 2016 | — | 0 | 1 | Yes |
| 126 | Statistics Iceland (1997) | ISL 2 | | No | 1625 | 1990 | — | 0 | 1 | Yes |
| 127 | Bank of Israel (2001) | ISR 1 | | No | 1961 | 2000 | — | 0 | 1 | Yes |
| 128 | Baffigi (2013) | ITA 1 | | No | 1861 | 2011 | — | 0 | 1 | Yes |
| 129 | Piselli and Vercelli (2023) | ITA 2 | | No | 1861 | 2016 | — | 0 | 1 | Yes |

*Continued on next page*

*Continued from previous page*

| # | Source | Abbreviation | Updated | Digitized | From | To | Forecasts | Variables | Countries | Historical |
|---|---|---|---|---|---|---|---|---|---|---|
| 130 | Ist. Naz. di Statistica (2024) | ITA 3 | 2026-07-01 | No | 1995 | 2023 | — | 0 | 1 | No |
| 131 | Banca d'Italia (2011) | ITA 4 | | No | 1861 | 2010 | — | 0 | 1 | Yes |
| 132 | Bank of Japan (2024) | JPN 1 | | No | 1882 | 2017 | — | 0 | 1 | Yes |
| 133 | Diebolt (2003) | JPN 2 | | No | 1868 | 1940 | — | 0 | 1 | Yes |
| 134 | IER, Hitotsubashi University (2019) | KOR 1 | | No | 1911 | 2016 | — | 0 | 1 | Yes |
| 135 | Cha, Kim, Park and Park (2022) | KOR 2 | | No | 1877 | 2015 | — | 0 | 1 | Yes |
| 136 | Min. of Economy and Finance, Korea (2026) | KOR 3 | 2026-07-01 | Yes | 2017 | 2026 | 2 | 0 | 1 | No |
| 137 | Gardner (2022) | LBR 1 | | No | 1845 | 1979 | — | 0 | 1 | Yes |
| 138 | Bank Al-Maghrib (2024) | MAR 1 | | No | 1985 | 2024 | — | 0 | 1 | Yes |
| 139 | Martín Moreno (2026) | MEX 1 | | Yes | 1876 | 2025 | — | 0 | 1 | Yes |
| 140 | CBS and Univ. Groningen (2001) | NLD 2 | | No | 1800 | 1998 | — | 0 | 1 | Yes |
| 141 | Eichholtz (1997) | NLD 3 | | Yes | 1628 | 1973 | — | 0 | 1 | Yes |
| 142 | Grytten (2022) | NOR 1 | | No | 1816 | 2019 | — | 0 | 1 | Yes |
| 143 | Eitrheim, Klovland and Qvigstad (2023) | NOR 2 | | No | 1516 | 2022 | — | 0 | 1 | Yes |
| 144 | NZ Inst. Economic Research (2026) | NZL 1 | | No | 1800 | 2025 | — | 0 | 1 | Yes |
| 145 | New Zealand Treasury (2026) | NZL 2 | 2026-07-01 | No | 1972 | 2025 | — | 0 | 1 | No |
| 146 | Banco Central de Reserva, Perú (2022) | PER 1 | | No | 1922 | 2021 | — | 0 | 1 | Yes |
| 147 | Statistics Poland (2024) | POL 1 | 2026-07-01 | No | 1989 | 2020 | — | 0 | 1 | No |
| 148 | Inst. Nacional de Estatística (2001) | PRT 1 | | Yes | 1549 | 1998 | — | 0 | 1 | Yes |

*Continued on next page*

*Continued from previous page*

| # | Source | Abbreviation | Updated | Digitized | From | To | Forecasts | Variables | Countries | Historical |
|---|---|---|---|---|---|---|---|---|---|---|
| 149 | Markevich and Harrison (2011) | RUS 1 | | Yes | 1913 | 1928 | — | 0 | 1 | Yes |
| 150 | Saudi Central Bank (2024) | SAU 1 | 2026-07-01 | No | 1973 | 2020 | — | 0 | 1 | No |
| 151 | Schön and Krantz (2017) | SWE 1 | | No | 1293 | 2014 | — | 0 | 1 | Yes |
| 152 | Lobell (2025) | SWE 2 | | No | 1300 | 2022 | — | 0 | 1 | Yes |
| 153 | Central Bank of Türkiye (2024) | TUR 1 | 2026-07-01 | No | 1994 | 2024 | — | 0 | 1 | No |
| 154 | Pamuk (2004) | TUR 2 | | Yes | 1469 | 1914 | — | 0 | 1 | Yes |
| 155 | National Statistics, Taiwan (2024) | TWN 1 | | No | 1951 | 2021 | — | 0 | 1 | Yes |
| 156 | Taiwan Statistical Office (2025) | TWN 2 | | No | 1951 | 2025 | — | 0 | 1 | Yes |
| 157 | Fed. Reserve Bank of St. Louis (2024) | FRED | 2026-07-01 | No | 1915 | 2026 | — | 0 | 1 | No |
| 158 | Carter et al. (2006) | USA 2 | | No | 1774 | 2003 | — | 0 | 1 | Yes |
| 159 | Baptista (1997) | VEN 1 | | Yes | 1825 | 1995 | — | 0 | 1 | Yes |
| 160 | South African Reserve Bank (2024) | ZAF 1 | 2026-07-01 | No | 1959 | 2020 | — | 0 | 1 | No |

**Note:** This table lists all datasets currently incorporated in the Global Macro Database (GMD). "Updated" refers to the last time a particular set was updated; it is blank when the dataset is downloaded only once. "Digitized" tags whether a dataset was newly digitized (i.e., copied from scans or PDF into digital format). "From" and "To" refer to the years covered. "Forecasts" is the number of forecasts into the future. "Variables" refers to the number of variables. "Countries" refers to the number of countries covered by a dataset. "Historical" means a dataset is not or only infrequently updated.

### A.10.2 How we update the data

**Automated downloads** For current sources, we use an automated data collection process through APIs and structured web queries. Our data downloading and processing pipeline has the following key functions:

- Automatic downloading and processing of new data as it becomes available
- Data validation and harmonization
- Integration of new information into the existing dataset

**Manual collection** Historical sources require careful manual processing. This involves the one-time collection of historical datasets, digitization of printed materials where necessary, standardization of formats and units, and integration with the automated updating system for current data. This process ensures that the historical data maintains the same quality standards as our automated collections.

### A.10.3 Data storage and version control

**Raw data storage** A critical feature for combining many dozens of historical sources with up-to-date recent data is the maintenance of a comprehensive archive of previous versions. For every source file, we always maintain the original data files in their native formats (csv, Excel, Stata, etc.) at the time they were downloaded. We also maintain any additional source documentation or metadata, as available. If the data is newly digitized or comes from an unusual source, we store the PDF files and archive the relevant web pages. In all cases, we always record the timestamp when a dataset was downloaded and the source URLs.

**Version control** Our version control system tracks data vintages from each source. We store newly-downloaded incremental updates and always record whether data values are genuinely new information or represent revisions (i.e., updates of data points that were already previously available). As such, we maintain complete historical records of all data points, and document all modifications and harmonization steps. This systematic approach ensures full traceability of every transformation made to the original data.

### A.10.4 Data update process

**Update monitoring and integration** Our update process begins with a continuous monitoring of current sources for new releases, supported by an automated notification system for new data availability and regular checks of historical sources for potential updates. The integration phase

involves automated comparison with existing data, identification of revised values, integration of new data points, and maintenance of vintage data.

**Quality control** Each update undergoes a rigorous quality control process consistent with the initial release of the Global Macro Database. Among others, we always run (a) a set of automated data validity checks, (b) unit and format verification checks, (c) consistency checks relative to previous versions, and (d) manually review all significant changes. This process is designed to make the resulting dataset as error-free as possible.

In addition to these initial checks, we always manually check the time series plot for each country and variable for potential issues.

## A.11 Variable Definitions

### A.11.1 Statistical framework

Wherever possible, our variable definitions follow the System of National Accounts 2008 (SNA 2008), the international statistical standard for national accounts adopted by the United Nations Statistical Commission. The SNA 2008 provides a comprehensive, consistent, and flexible framework for collecting and reporting macroeconomic statistics.

The SNA framework is a reasonable starting point for our dataset because it provides internationally agreed-upon concepts, definitions, and classifications, ensures consistency and comparability across countries and time periods as much as possible, and offers standardized methodologies for data compilation.

For variables that are not reported as part of national accounts statistics, we try to adopt a consensus approach by surveying the metadata and best practices in existing work.

### A.11.2 Nominal gross domestic product

Nominal Gross Domestic Product (GDP) measures the total market value of all final goods and services produced within a country's borders in a specific time period, typically a year. It reflects the economic output valued at the prices of that same year, often referred to as "current prices."

The international standard, the System of National Accounts (SNA), outlines three distinct approaches for calculating GDP. In theory, all three methods should result in the same values, but there are various reasons why they may differ in practice.

- **Production Approach:** Calculates GDP as the sum of gross value added by all resident producer units, plus any taxes and less any subsidies on products (United Nations, 2009, Chapter 16, Section C).

- **Income Approach:** Sums the incomes generated by production. This includes compensation of employees (wages and salaries), gross operating surplus of enterprises (profits), and taxes on production and imports less subsidies.
- **Expenditure Approach:** Computes GDP by summing all final expenditures. It is calculated as the sum of final consumption ($C$), gross capital formation (investment, $I$), government spending ($G$), and net exports (exports ($X$) minus imports ($M$)). The formula is:

$$\text{nGDP} = C + I + G + (X - M)$$

Our `nGDP` series is constructed from sources that may use any of these methods. It is possible to reconstruct nominal GDP based on the expenditure approach because the GMD provides series for its main components, namely household and government final consumption, gross capital formation, and net exports. That said, as discussed in Section A.13.6, the GMD currently does not impose accounting identities to hold, which would lead to other complications. `nGDP` is measured in millions of local currency units at current market prices.

### A.11.3 Real gross domestic product

**Real GDP in Local Currency (`rGDP`)** Real GDP measures a country's total economic output adjusted for the effects of price changes (i.e., inflation or deflation). It is valued using the prices of a fixed base year, thus providing a constant-price series that reflects changes in volume. In our dataset, `rGDP` is measured in millions of constant 2015 local currency units (LCU).

**Real GDP in U.S. Dollars (`rGDP_USD`)** There are various ways of calculating real GDP in U.S. dollars, all of which can make sense depending on what one is trying to measure. The GMD follows the methodology used by the World Bank, which ensures that the real growth rates of the local economy are preserved accurately.[3]

The calculation proceeds as follows:

1. First, the nominal GDP for the base year (2015) is converted to U.S. dollars using the 2015 average market exchange rate ($E_{c,2015}$). This single value serves as the anchor for the entire series. The anchor value is calculated as:

$$\text{rGDP}_{\text{USD},c,2015} = \frac{\text{nGDP}_{c,2015}}{E_{c,2015}} \tag{2}$$

   For the base year 2015, the values in current and constant prices are the same.

[3]The methodology is detailed by the World Bank here: https://datahelpdesk.worldbank.org/knowledgebase/articles/114943-how-do-you-derive-constant-price-series-in-usd.

2. Second, the annual real growth rates ($g_{c,t}$) are calculated from the constant local currency series (`rGDP`):

$$g_{c,t} = \frac{\text{rGDP}_{c,t}}{\text{rGDP}_{c,t-1}} - 1 \tag{3}$$

3. Finally, these growth rates are applied forwards and backwards from the 2015 constant U.S. dollar value to generate the complete time series. The series is constructed recursively as follows:

$$\text{rGDP}_{\text{USD},c,t} = \begin{cases} \text{rGDP}_{\text{USD},c,t-1} \times (1 + g_{c,t}) & \text{if } t > 2015 \\ \frac{\text{rGDP}_{\text{USD},c,t+1}}{1+g_{c,t+1}} & \text{if } t < 2015 \end{cases} \tag{4}$$

This method ensures that the resulting series reflects the true volume growth of the local economy, free from the distorting effects of both domestic inflation and exchange rate volatility.

**GDP deflator (`deflator`)** The GDP deflator is defined as a price index derived by dividing nominal GDP by real GDP (United Nations, 2009, Chapter 15, Section C). `deflator` measures the overall price level of all domestically produced final goods and services. The deflator is set to 100 in the base year (2015) and expressed as an index.

### A.11.4 Consumption and investment

**Total consumption (`cons`)** Total consumption is the total expenditure on the consumption of goods and services (United Nations, 2009, Chapter 9, Section D). `cons` is measured in millions of local currency units. `cons_GDP` is in percent of nominal GDP, so that 100 means 100%.

**Household consumption (`hcons`)** Household final consumption expenditure consists of the expenditure incurred by resident households on consumption goods and services (United Nations, 2009, Chapter 9, Section D). It includes expenditure on durable goods, semi-durable goods, non-durable goods, and services. `hcons` is measured in millions of local currency units. `hcons_GDP` is in percent of nominal GDP.

**Government consumption (`gcons`)** Government final consumption expenditure consists of expenditure incurred by general government on both individual consumption goods and services and collective consumption services (United Nations, 2009, Chapter 9, Section D). It includes compensation of employees, intermediate consumption, and social transfers in kind. `gcons` is measured in millions of local currency units. `gcons_GDP` is in percent of nominal GDP.

**Gross capital formation (`inv`)** Gross capital formation is measured by the total value of the gross fixed capital formation, changes in inventories, and acquisitions less disposals of valuables.

(United Nations, 2009, Chapter 10, Section B). `inv` is measured in millions of local currency units. `inv_GDP` is in percent of nominal GDP.

**Gross fixed capital formation (`finv`)** Gross fixed capital formation is measured by the total value of a producer's acquisitions, less disposals, of fixed assets during the accounting period plus certain specified expenditure on services that adds to the value of non-produced assets. (United Nations, 2009, Chapter 10, Section B). `finv` is measured in millions of local currency units. `finv_GDP` is in percent of nominal GDP.

### A.11.5 External sector

**Exports (`exports`)** Exports are goods and services produced in one economy and sold to another economy, valued free on board (f.o.b.) at the border of the exporting country (United Nations, 2009, Chapter 15, Section B). `exports` is measured in millions of local currency units. `exports_GDP` is in percent of nominal GDP.

**Imports (`imports`)** Imports are goods and services purchased by residents from nonresidents, originally valued cost, insurance, and freight (c.i.f.) but converted to f.o.b. (United Nations, 2009, Chapter 15, Section B). `imports` is measured in millions of local currency units. `imports_GDP` is in percent of nominal GDP.

**Current account balance (`CA`)** The current account is defined as the sum of the balances on goods, services, primary income, and secondary income (United Nations, 2009, Chapter 16, Section B). It is considered a key indicator of an economy's saving-investment relationship with the rest of the world. `CA` is measured in millions of local currency units. `CA_GDP` is in percent of nominal GDP.

**Real effective exchange rate (`REER`)** The real effective exchange rate, or REER, is the trade-weighted average of bilateral exchange rates adjusted for relative price levels. `REER` is expressed as an index equal to 100 in the base year (2015) and thus measures (changes in) international competitiveness.

**USD exchange rate (`USDfx`)** The US dollar exchange rate is defined as the value of one U.S. dollar in terms of local currency units. Wherever possible, we report end-of-period rates.

### A.11.6 Government finance

Government finance statistics are among the most difficult to create harmonized time series for. Many sources are not consistent in only reporting *central* or *general* government data, where the

latter also includes various types of local governments. This is true both across sources—some may only report central government debt—but also within the *same source*. For example, the data on government debt reported by a given source may for some countries cover general government debt, but in others only central government debt.

In fact, such inconsistencies even occur within a given source. A look at the appendix material of the widely-used "Public Finances in Modern History Database" produced by IMF staff (Mauro, Romeu, Binder and Zaman, 2015) is revealing. Even when looking at the same country over time, their series are sometimes based on data covering the general government, and sometimes on the central government alone. In Australia, for example, their reported government spending doubles from 11.2% to 22.3% of GDP in 1960, which likely results from using general rather than central government expenditures starting in 1960.

Partly because of the above, there is often considerably more disagreement about government statistics than other variables. It is thus not straightforward to combine data on government statistics from several sources. As an example, splicing either data on outstanding debt or on the ratio of debt to GDP will inevitably be based on contentious and often widely-differing data points. We are thus faced with a tough choice. If we chainlink these data, the resulting values can end up being entirely implausible. If we do not, this will introduce statistical breaks (and thus implausible year-on-year *changes* in a variable) when the differences in the underlying raw data are large.

To arrive at series that are as reliable as possible, we use the following procedure. First, we carefully combine data on general and central government statistics separately. In doing so, we attempt to select data that are as consistent as possible with each other in whether they cover only the central or general government. These have variable codes starting with the prefix `c` (e.g., `cgovdebt_GDP`) or `gen_` (e.g., `gen_govdebt_GDP`) for central and general government, respectively.

Starting with the 2026-03 release, government finance ratio variables (e.g., `cgovdebt_GDP`) are spliced directly as ratios rather than being derived from separately spliced level series. This avoids compounding errors that arise when the numerator and denominator receive different chainlinking adjustments. The corresponding level variables (e.g., `cgovdebt`) are then derived from the spliced ratios using the formula level $=$ (ratio $\times$ nGDP)/100.

Second, equipped with these data, we further construct a combined measure of government debt, which combines these estimated general and central government series. Note that combining the statistics in this way is already standard practice in existing work (see e.g., Mauro, Romeu, Binder and Zaman, 2015). The resulting variables do *not* have a prefix (e.g., `govdebt_GDP`). It is in the nature of the exercise that these combined measures may have higher measurement error but, on the flipside, also larger coverage. Note that, for combined variables, we also do not have "raw data" that could be accessed through, for example, the `gmd` Stata command.

For the remainder of the variable definitions, we will denote all the government finances to start with `*_` (e.g., `*_govdebt`) to highlight that the GMD contains three versions of each (central, general, and combined).

**Government debt (`*_govdebt`)** Government debt refers to the total liabilities of the government requiring future payments of interest and/or principal. It includes loans, debt securities, and other borrowings (United Nations, 2009, Chapter 22, Section D). `govdebt` is measured in millions of local currency units. `govdebt_GDP` is in percent of nominal GDP, so that 100 means 100%.

**Government revenues (`*_govrev`)** Government revenue is the increase in net worth of the government resulting from incoming transactions; it includes taxes, social contributions, grants, and other revenue (United Nations, 2009, Chapter 22, Section C). `govrev` is measured in millions of local currency units. `govrev_GDP` is in percent of nominal GDP.

**Government tax revenue (`*_govtax`)** Government tax revenues are compulsory transfers to government units, including taxes on income, profits, goods and services, and international trade (United Nations, 2009, Chapter 22, Section C). `govtax` is measured in millions of local currency units. `govtax_GDP` is in percent of nominal GDP.

**Government expenditure (`*_govexp`)** Government expenditure is a decrease in net worth of the government resulting from transactions, including compensation of employees, use of goods and services, and transfers (United Nations, 2009, Chapter 22, Section C). `govexp` is measured in millions of local currency units. `govexp_GDP` is in percent of nominal GDP.

**Government deficit (`*_govdef`)** Government deficit refers to net lending/borrowing by the government, i.e. the difference between revenue and expenditure (United Nations, 2009, Chapter 22, Section C). `govdef` is the primary measure of a country's fiscal position and is measured in millions of local currency units. `govdef_GDP` is in percent of nominal GDP.

### A.11.7 Money and interest rates

**Money supply (`M0`)** M0 is defined as notes and coins in circulation outside depository corporations (International Monetary Fund, 2016, Chapter 6, Section IV). It is considered the most liquid monetary aggregate. `M0` is measured in millions of local currency units.

**Money supply (`M1`)** M1 is defined as currency in circulation plus transferable (demand) deposits (International Monetary Fund, 2016, Chapter 6, Section III). It is the most liquid monetary aggregate, including bank deposits. `M1` is measured in millions of local currency units.

**Money supply (`M2`)** M2 is defined as M1 plus time and savings deposits and includes less liquid monetary assets (International Monetary Fund, 2016, Chapter 6, Section III). `M2` is measured in millions of local currency units.

**Money supply (`M3`)** M3 is defined as M2 plus marketable instruments issued by depository corporations (refer to (International Monetary Fund, 2016, Chapter 6, Section III)). It represents the broadest monetary aggregate measured within the banking system. `M3` is expressed in millions of local currency units.

**Money supply (`M4`)** M4 is defined as M3 plus debt securities issued by the central government that are held by money holders (refer to (International Monetary Fund, 2016, Chapter 6, Section III)). It is the most comprehensive measure of the money supply. `M4` is also expressed in millions of local currency units.

**Central bank policy rate (`cbrate`)** The central bank policy rate is the key interest rate used to implement or signal monetary policy stance (International Monetary Fund, 2016, Chapter 4, Section II). It usually applies to short-term liquidity operations. `cbrate` is measured in percent per annum.

**Short-term interest rate (`strate`)** The short-term interest rate is the market rate on short-term government securities or interbank rates, usually referring to three-month maturity (International Monetary Fund, 2016, Chapter 4, Section II). `strate` is measured in percent per annum.

**Long-term interest rate (`ltrate`)** The long-term interest rate is the market yield on long-term government bonds, usually referring to ten-year maturity (International Monetary Fund, 2016, Chapter 4, Section II). `ltrate` is measured in percent per annum.

### A.11.8 Prices, labor market, and population

**Consumer price index (`CPI`)** The consumer price index (CPI) is a measure of the average change in prices paid by consumers, with the weights based on household consumption patterns (United Nations, 2009, Chapter 15, Section A). `CPI` is expressed as an index equal to 100 in the base year (2015).

**House price index (`HPI`)** The house price index (HPI) measures changes in residential property prices, where possible adjusted for quality (Eurostat, 2013). `HPI` is expressed as an index equal to 100 in the base year (2015).

**Inflation (`infl`)** Inflation is defined as the period-on-period percentage change in the consumer price index, which measures the rate of price level changes (United Nations, 2009, Chapter 15, Section A). `infl` is expressed in percent per annum.

**Unemployment rate (`unemp`)** The unemployment rate is defined as the number of unemployed persons as a percentage of the labor force (United Nations, 2009, Chapter 19, Section D). Wherever possible, it is based on International Labor Organization standards. `unemp` is expressed in percent.

**Population (`pop`)** Population is the total number of persons present in the economic territory, including both nationals and foreigners (United Nations, 2009, Chapter 19, Section B). `pop` is measured in millions of persons.

## A.12 Priority Ordering of Data Sources

Our extensive experience working with macroeconomic time series has led us to develop a hierarchy for how different sources should be prioritized. This ordering reflects both the reliability of the data and the practical considerations of maintaining a comprehensive database. The hierarchy follows three main tiers:

### A.12.1 Tier 1: Modern official sources

Modern official sources, particularly those from international organizations, national statistical offices, and central banks receive the highest priority. These institutions typically have the most accurate and up-to-date information for their respective countries, along with detailed documentation of methodologies and regular revision schedules.

### A.12.2 Tier 2: Country-specific historical sources

The second tier consists of country-specific historical sources, often compiled by economic historians or research institutions focusing on particular countries or regions. These sources frequently offer invaluable historical data that has been carefully reconstructed and has often been evaluated through the academic peer-review process. Notable examples include the historical statistics for Australia compiled by Vamplew (1987), Portuguese historical statistics by Inst. Nacional de Estatística (2001), and Argentinian long-run series by Nakamura and Zarazaga (2001).

### A.12.3 Tier 3: Other aggregators

International aggregators such as the IMF, World Bank, or OECD form the third tier. While these sources provide extensive cross-country coverage and standardized definitions, they sometimes

sacrifice historical depth or country-specific accuracy.

This ordering has proven robust through extensive testing and practical application. It balances the trade-offs between data accuracy, historical coverage, methodological consistency, and maintenance feasibility. The hierarchy is not rigid, however, and we document any deviations from this general ordering in our detailed source notes for each country and variable.

## A.13 Measurement Issues

### A.13.1 Data quality

A key contribution of our dataset is the systematic approach to data quality control through comprehensive visual inspection of all time series. For each of our 46 variables and 239 countries/territories, we plot the data from every available source on a single graph, allowing for detailed comparison and anomaly detection. This results in over 4,000 individual plots that we manually inspect for data quality issues.

**Visual inspection process** Our visual inspection process creates plots for each country-variable combination. These plots comprehensively display our final GMD estimates, a GMD forecast (where available), data points from all available sources, a clear indication of splice points between different sources, and notes on major adjustments or concerns.

This visualization process allows us to identify several critical types of data quality issues. We detect level shifts, which appear as unexpected jumps in the series that might indicate currency changes, definition changes, or data errors. We also identify source discrepancies, where different sources report substantially different values for the same period. The process also reveals outliers that deviate significantly from the series trend, inconsistencies in the units or currency in which a data series is recorded, and splicing problems where different data sources are not correctly "stitched together."

**Documentation and correction** After identifying a potential data quality issue, we implement a systematic correction procedure. We begin by thoroughly documenting the nature of the anomaly and cross-referencing it with other sources and historical events. This investigation helps us determine whether the issue reflects a genuine economic event, a definitional change, or a data error requiring correction. Based on this assessment, we apply the necessary corrections while maintaining a detailed documentation of all adjustments. Throughout this process, we carefully flag any remaining uncertainties in the dataset to ensure full transparency.

We maintain a comprehensive set of PDF documents containing plots for each country and variable, accessible at www.globalmacrodata.com. This documentation provides extensive details,

including complete source information, the time spans each data source is used for constructing our time series, notes on specific adjustments or concerns, and detailed explanations of any splicing procedures applied.

As an illustrative example, consider the case of government expenditure data in Argentina. We use the following sources:

- Mitchell (2013): 1864-1989
- International Monetary Fund (2024): 1990-1992
- International Monetary Fund (2025*b*): 1993-2029

These plots serve dual purposes: they function as both a quality control tool and transparent documentation of our data construction process. This comprehensive approach allows users to evaluate the reliability of specific series and understand precisely how different sources were combined to create our final estimates.

### A.13.2 Dealing with data revisions

A key challenge in maintaining a long-run macroeconomic database is the treatment of data revisions. Statistical offices and international organizations frequently revise their estimates as new information becomes available or methodologies improve. Without proper treatment, these revisions would propagate through the entire historical series when using standard chain-linking procedures, causing implausible changes to historical values.

**Anchor year approach** To address the issue of data revisions, we implement an anchor year approach that effectively separates historical data from recent data revisions. We establish 2018 as our anchor year for the current version of the dataset, with historical data (pre-2018) chain-linked backwards and contemporary data (post-2018) chain-linked forwards from this anchor year. Note that, in the rare cases where countries do not have data available as of 2018, we use the most recent available year as an anchor year.

This approach creates a clear framework for handling new data and revisions:

- Values before 2018 remain unchanged
- Values after 2018 incorporate all new information and revisions
- The anchor year serves to prevent the historical data from being continuously revised, while allowing for revisions to more recent data

Going forward, we will occasionally consider a change in the anchor year to take into account possible data revisions to 2018 values of the data.

**Example** To illustrate our approach of fixing the anchor year, consider a hypothetical case for real GDP. We begin with a historical source providing data for 1850-1989 and a contemporary source (e.g., International Monetary Fund (2025*b*)) covering 1990-2024. When a new (International Monetary Fund, 2025*b*) release revises the values for the 2022-2024 period and adds 2025-2029 forecasts, the treatment of these revisions differs significantly under different approaches.

Under traditional chain-linking starting from the most recent data, revisions to recent years would affect the entire series back to 1850, meaning historical values would change despite no new historical information becoming available. In contrast, our anchor year approach provides a more stable framework:

- 2018 serves as the fixed anchor point
- Data for 1850-2018 is constructed by chain-linking backwards from 2018
- Data for 2019-2025 is constructed by chain-linking forwards from 2018
- Revisions to 2022-2024 and the addition of 2025 only affect post-2018 values

**Implementation** For each level variable, we implement a systematic procedure that begins with fixing the level of the series in 2018 using our preferred contemporary source. For earlier years ($t < 2018$), we use growth rates from historical sources to chain-link backwards from 2018, with these values remaining fixed unless errors are discovered. For later years ($t > 2018$), we employ growth rates from contemporary sources to chain-link forwards from 2018, allowing these values to update with each new data release.

This methodology ensures that our historical estimates remain stable while still incorporating all relevant contemporary revisions and updates. We plan to periodically update the anchor year (e.g., to the most recent years) in the next major releases of the dataset. At this point, the entire series will be recomputed to incorporate any important historical revisions.

### A.13.3 Base years for index variables

Several variables in our dataset are expressed either in constant prices or as indices (e.g., Consumer Price Index, GDP deflator, real GDP). These variables require a common base year for meaningful comparison; we set the base year to 2015. We implement a systematic approach to standardize base years.

**Primary reference source** We use the IMF's World Economic Outlook (International Monetary Fund, 2025*b*) as our primary reference source for base years. This choice is motivated by several key

advantages: the WEO offers wide country coverage, maintains a regular updating schedule, employs consistent methodological treatment, and enjoys broad usage in applied research.

### A.13.4 Rebasing methodology

To ensure consistency across countries and time periods, we rebase series to a common base year, set to 2015. To illustrate the rebasing procedure, we use real GDP as an example, which uses country-specific GDP deflators. Formally, the deflator is computed as:

$$\text{Deflator}_{c,t} = \left( \frac{\text{nGDP}_{c,t}}{\text{rGDP}_{c,t}} \right) \times 100, \tag{5}$$

where $\text{nGDP}_{c,t}$ and $\text{rGDP}_{c,t}$ denote nominal and real GDP for country $c$ in year $t$, respectively.

We then loop over each country separately to anchor each series to the base year 2015 deflator value:

$$\text{rGDP}_{c,t}^{\text{rebased}} = \text{rGDP}_{c,t} \times \frac{\text{Deflator}_{c,2015}}{100}, \tag{6}$$

where $\text{Deflator}_{c,2015}$ is the GDP deflator value for country $c$ in the base year 2015. After rebasing, we update the GDP deflator to reflect the recalculated real GDP values:

$$\text{Deflator}_{c,t}^{\text{updated}} = \left( \frac{\text{nGDP}_{c,t}}{\text{rGDP}_{c,t}^{\text{rebased}}} \right) \times 100. \tag{7}$$

This rebasing method ensures temporal consistency and comparability of real GDP figures across different countries in our dataset.

### A.13.5 Changes in currency

The long-run nature of our dataset requires careful handling of currency changes and redenominations. Historical sources often report values in currencies that were in use at the time of recording, while modern sources typically adjust historical values to current currencies. We identify and address two distinct types of currency changes:

1. **Adoption of new currencies:** Complete changes in the monetary unit, such as:
    - Introduction of new national currencies post-independence
    - Currency union adoptions (e.g., Euro)
    - Post-hyperinflation currency reforms
2. **Redenomination of existing currencies:** Technical adjustments to the same basic monetary unit, typically:

- Removal of zeros after hyperinflation periods
- Technical currency reforms maintaining the same basic unit

For new currency adoptions, we apply the official conversion rate at the time of change. For example, in the case of Euro adoption, we use the irrevocable exchange rates set by the European Central Bank (e.g., 1.95583 Deutsche Mark = 1 Euro for Germany). All historical values are converted using these official rates, and we document the conversion rate and date in country-specific notes.

Currency redenominations, particularly following periods of hyperinflation, require special attention. For instance, Turkey removed six zeros from the Lira in 2005, requiring all pre-2005 values to be divided by 1,000,000 to maintain consistency. Brazil presents a more complex case, with multiple changes between 1942 and 1994:

- 1942-1967: Cruzeiro
- 1967-1970: Cruzeiro Novo, 1000:1 redenomination
- 1970-1986: Cruzeiro
- 1986-1989: Cruzado, 1000:1 redenomination
- 1989-1990: Cruzado Novo, 1000:1 redenomination
- 1990-1993: Cruzeiro
- 1993-1994: Cruzeiro Real, 1000:1 redenomination
- 1994-present: Real, 2750:1 conversion

For such complex cases, we implement a systematic procedure in four steps. First, we document the complete chain of currency changes. Second, we apply conversions sequentially using official exchange rates. Third, we cross-validate the results with multiple sources when available. Fourth, we record potential measurement uncertainties in periods of extreme inflation.

To ensure accurate currency conversions, we employ several quality control measures. These start with an automated detection of unusual jumps in series around known currency change dates and a cross-validation with multiple sources where available. Most importantly, we manually review all major currency changes.

For each country-year observation affected by currency changes, we maintain detailed documentation of the original currency and value, applied conversion rate(s), and final adjusted value. This information is available in the country-specific technical notes. In cases where sources disagree about the appropriate conversion rate, particularly during periods of high inflation, we use geometric averages of available estimates and flag these observations for higher uncertainty.

### A.13.6 Accounting identities and internal consistency

Some variables in the dataset, such as nominal GDP, have several components that should in principle add up to the total. In the case of GDP, the expenditure approach implies that $Y = C + I + G + NX$. Ideally, this accounting identity should thus always hold in the data. In practice, data providers such as Eurostat include an unexplained "statistical discrepancy" even for recent data points that would have to be incorporated for the components to add up to GDP.

Because we construct time series from many different underlying sources using ratio-splicing, accounting identities such as the one outlined above will not hold exactly. While we do not explicitly provide it, the statistical discrepancy in our data can be calculated as the difference between GDP and the sum of its underlying components.

The alternative to accepting a statistical discrepancy would be to use a hierarchical reconciliation method. These methods rescale the underlying components so that they exactly add up to the total. Because using a statistical discrepancy is the standard approach taken by almost all data providers, however, we do not use hierarchical reconciliation. To ensure the highest possible data quality, we instead manually investigate all cases where the statistical discrepancy is large to identify potential errors before each release.

### A.13.7 Adjusting for source changes and series breaks

Series breaks occur when the same variable exhibits a discontinuity due to methodological changes, redefinitions, or changes in source data. We identify a series break when either:

- There is an explicit documentation of methodological change
- There is an unexplained jump in the series that unambiguously cannot be attributed to economic events
- Other sources than the one in use report substantially different growth rates or values for the same period

We employ two distinct methodologies for addressing series breaks, depending on the availability of overlapping data.

**Method 1: Overlapping data available** When we have overlapping observations between two sources at the break point $t_0$, we adjust the historical series using the ratio splicing method:

$$\hat{y}_s = y_s^{next} \times \frac{y_{t_0}^{previous}}{y_{t_0}^{next}} \quad \text{for all } s < t_0 + 1 \tag{8}$$

where:

- $\hat{y}_s$ is the adjusted value for period $s$
- $y_s^{next}$ is the value from the next series used in the chain-linking process
- $y_t^{previous}$ is the value from the previous series at the break point
- $y_t^{next}$ is the value from the next series at the break point

**Method 2: No overlapping data** In a few cases, there is no overlapping data to "connect" the raw data from adjacent years. In these cases, we use an ad-hoc approach to estimate a plausible growth rate for the year where we have no information available. Specifically, we adjust the data using a "representative" growth rate around the year of the break:

1. **Compute growth rates from the new source after the break:**

$$g_t^{\text{previous}} = \frac{y_{t_0}^{\text{previous}} - y_{t_0-1}^{\text{previous}}}{y_{t_0-1}^{\text{previous}}}, \quad \text{for } t \in \{t_0+1, t_0+2, t_0+3\} \tag{9}$$

2. **Compute growth rates from the old source before the break:**

$$g_t^{\text{next}} = \frac{y_{t_0}^{\text{next}} - y_{t_0-1}^{\text{next}}}{y_{t_0-1}^{\text{next}}}, \quad \text{for } t \in \{t_0-3, t_0-2, t_0-1\} \tag{10}$$

3. **Create a combined growth rate series:**

$$g_t = \begin{cases} g_t^{\text{previous}} \\ g_t^{\text{next}} \end{cases} \tag{11}$$

   This series pools available growth rates from both sources to increase robustness.

4. **Compute the median growth rate:**

$$\bar{g} = \text{median}(g_t) \tag{12}$$

   The median is used for robustness against outliers and to adjust the pre-break value:

5. **Adjust the pre-break value using the median growth rate:**

$$y_{t_0+1}^{\text{break_adjusted}} = \frac{y_{t_0+1}^{previous}}{1+\bar{g}} \tag{13}$$

6. **Compute the chainlinking ratio:**

$$\theta = \frac{y_{t_0+1}^{\text{break_adjusted}}}{y_{t_0}^{\text{next}}} \tag{14}$$

7. **Apply adjustment to historical series:**

$$\hat{y}_s = y_s^{next} \times \theta \quad \text{for all } s < t_0 + 1 \tag{15}$$

**Method 3: Weighted splice estimator** Starting with the 2026-03 release, we introduce a weighted splice estimator that blends three correction operators—Replace, Chainlink, and Different—using time-varying weights. Let $T$ denote the splice point and $d_t = T - t \geq 1$ the distance from it. The blended estimate is:

$$\tilde{y}_t \;=\; \pi_{R,t}\,\hat{y}_t^{(R)} \;+\; \pi_{C,t}\,\hat{y}_t^{(C)} \;+\; \pi_{D,t}\,\hat{y}_t^{(D)}, \tag{16}$$

where $\hat{y}_t^{(R)}$ is the raw (replace) value, $\hat{y}_t^{(C)}$ is the chainlink-adjusted value, and $\hat{y}_t^{(D)}$ is the different-operator value. The weights satisfy $\pi_{k,t} \geq 0$ and $\sum_k \pi_{k,t} = 1$ for all $t < T$. These weights are estimated by minimizing an objective function $\mathcal{L} = \mathcal{L}_3 + \mathcal{L}_4 + \mathcal{L}_5$ subject to the simplex constraint. The loss uses a Huber function $h(u;\, c) = \frac{1}{2}u^2$ for $|u| \leq c$ and $c(|u| - \frac{c}{2})$ otherwise, which limits the influence of outliers.

The objective balances three goals. First, a dual time-varying prior ($\mathcal{L}_3$) encodes the belief that raw data should receive more weight far from the splice point while the chainlink correction is most trustworthy near it, with both targets decaying exponentially at rate $\rho$. Second, an adaptive reallocation term ($\mathcal{L}_4$) compares operator deviations from the observed data and redistributes weight toward whichever correction operator performs better at each point in time. Third, an adaptive smoothness penalty ($\mathcal{L}_5$) discourages abrupt changes in the weight trajectory but relaxes when any operator produces an extreme value—defined by a threshold $\tau$ set from domain knowledge—allowing the optimizer to reallocate weights rapidly when needed. This approach generalizes Methods 1 and 2 by letting the data inform how much chain-linking correction to apply at each point, rather than imposing a single ratio uniformly across the entire pre-splice period.

**Implementation notes** We strictly give preference to Method 1 whenever overlapping data is available. Method 2 is only used when no overlapping observations exist. When multiple breaks exist in a series, we apply these methods sequentially. Crucially, all series breaks and adjustment methods are documented in the country-specific notes.

### A.13.8 Doubtful data points

In compiling a comprehensive macroeconomic database, we systematically identify and address doubtful data points that may reflect measurement errors rather than genuine economic phenomena. We classify these into two main categories: outliers and definitionally impossible values.

**Outliers** We define outliers as observations that exhibit implausible changes in the level of a series. Our identification process combines economic validation and source comparison. For economic validation, we cross-reference flagged observations with other sources, verify if large changes can be explained by known economic shocks, and compare with related economic variables for consistency. The source comparison involves checking alternative data sources for the same period and evaluating the reliability of different sources.

When outliers are identified, we follow a three-step protocol. First, we replace values with another source if a more reliable source is available. Second, if no reliable alternative exists, we set the observation to missing. Third, all adjustments are meticulously recorded in the country-specific notes.

**Definitionally impossible values** Certain variables have natural constraints on their possible values based on economic definitions. We systematically check for and address two types of constraints. First, sign restrictions dictate that stock variables (e.g., money supply, population), price indices, and exchange rates must be positive, while flow variables (e.g., government deficit, current account) can be negative. Second, logical bounds require that percentages (e.g., unemployment rate) must lie between 0 and 100, and ratios (e.g., debt-to-GDP) must be non-negative.

When encountering impossible values, we implement a three-step verification process. First, we compare the values with alternative sources to identify correct values. Second, we check for potential unit or sign errors in the source data. Third, we investigate possible definitional changes or special circumstances that might give rise to unusual values.

**Quality control process** For both types of doubtful data points, we implement a systematic review process:

1. Automated flagging of potential issues
2. Manual review of flagged observations
3. Cross-validation with multiple sources
4. Documentation of all adjustments
5. Regular review of flagging thresholds

All identified doubtful data points and their resolution are recorded in our technical documentation, allowing users to assess the reliability of specific observations and understand any adjustments made to the raw data.

### A.14 Planned Improvements

We are currently working on:

- Further improvements to the government finance statistics, especially to better take into account the different reporting of measures referring to central and general government
- Further harmonizing the definition of monetary aggregates across sources
- Adding measures of total factor and labor productivity
- Extending the coverage of exchange rate data, especially for former colonies
- Adding additional new sources we already processed but that are not yet integrated
- Quantifying statistical uncertainty across sources
- Improving internal consistency of the data, e.g., for national accounts variables

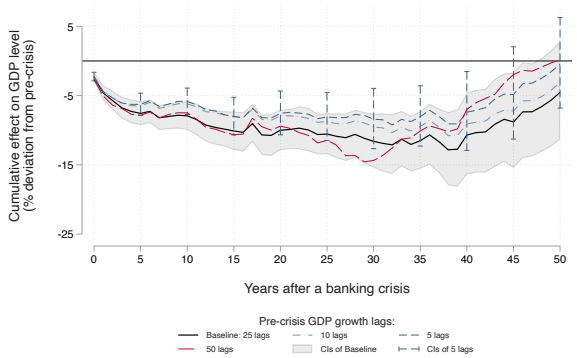
We always welcome additional suggestions. Just open an issue on our GitHub repo or reach out to us at kmueller@globalmacrodata.com.

# B Additional Results

Figure B.1: Baseline with additional lags and pre-trends

*This corresponds to Figure 3 in the main text.*

(a) Full Panel




(b) Balanced panel

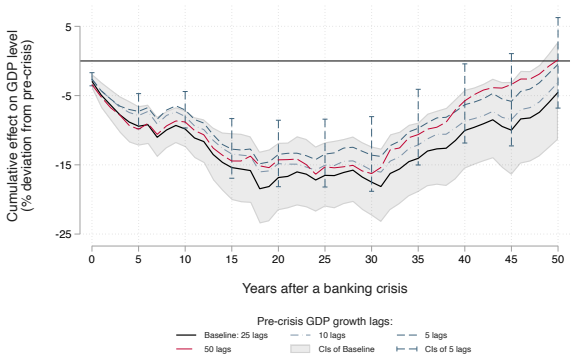



Figure B.2: Crisis Occurrence in Full and Balanced Panels

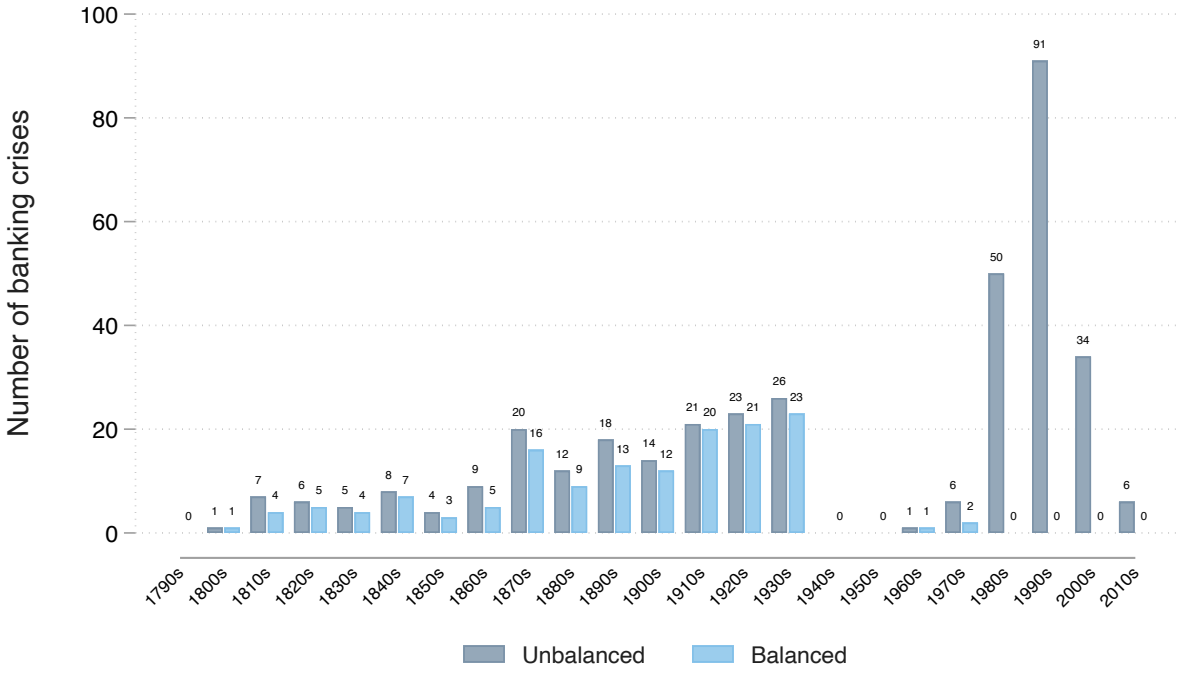
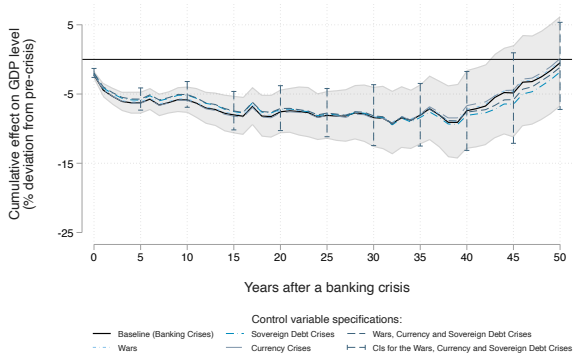


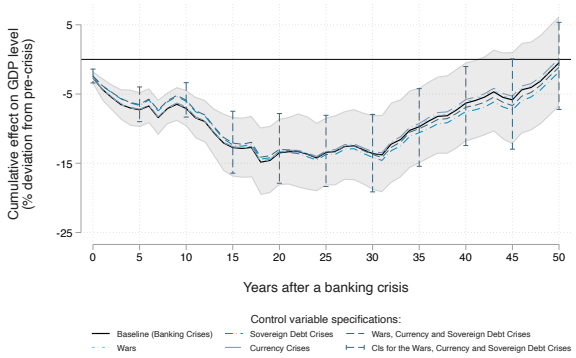
Figure B.3: Wars, Conflicts, and Other Crises

*This corresponds to Figure 4a in the main text.*

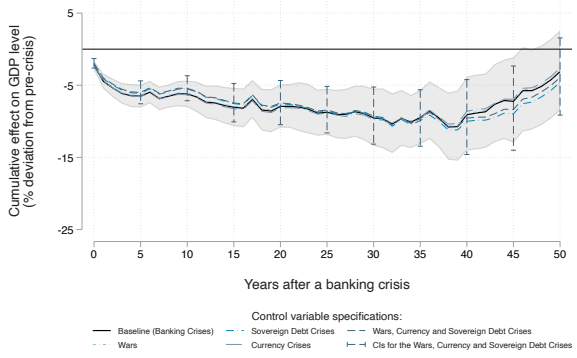
(a) Full Panel - 5 pre-crisis GDP growth lags



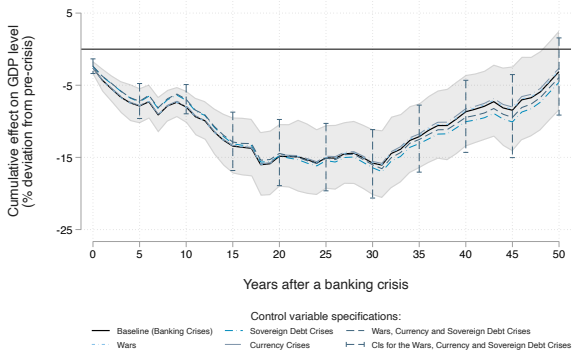

(b) Balanced panel - 5 pre-crisis GDP growth lags



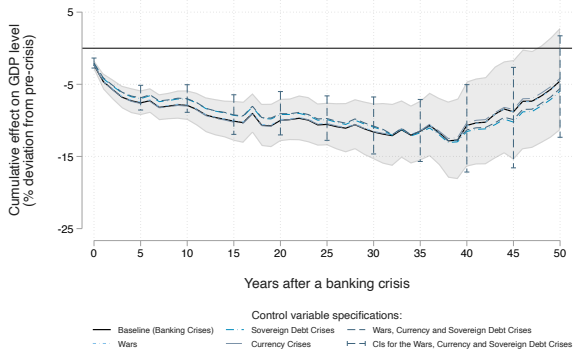

(c) Full Panel - 10 pre-crisis GDP growth lags




(d) Balanced panel - 10 pre-crisis GDP growth lags




(e) Full Panel - 25 pre-crisis GDP growth lags




(f) Balanced panel - 25 pre-crisis GDP growth lags

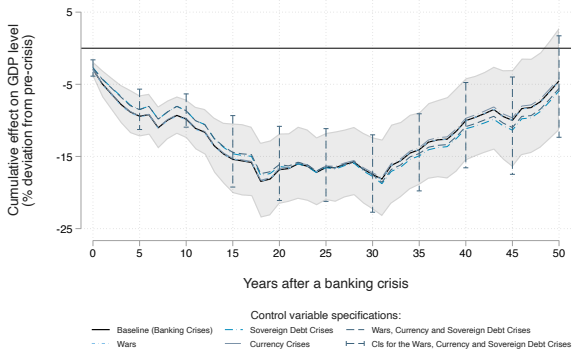
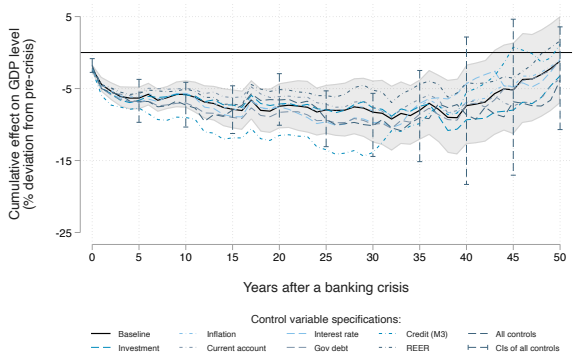


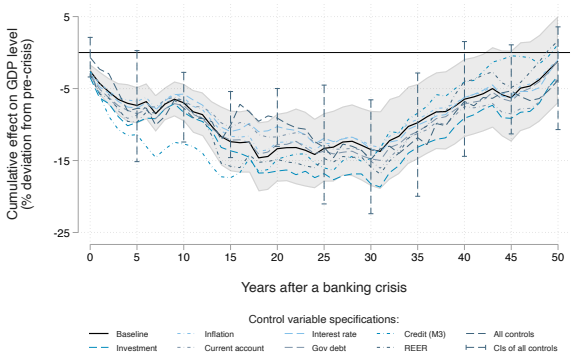
Figure B.4: Additional Macroeconomic Conditions

*This corresponds to Figure 4b in the main text.*

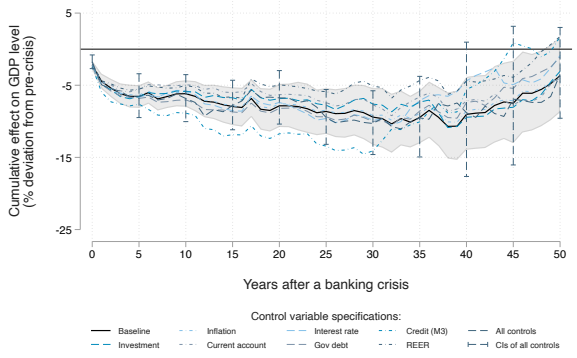
(a) Full Panel - 5 pre-crisis GDP growth lags



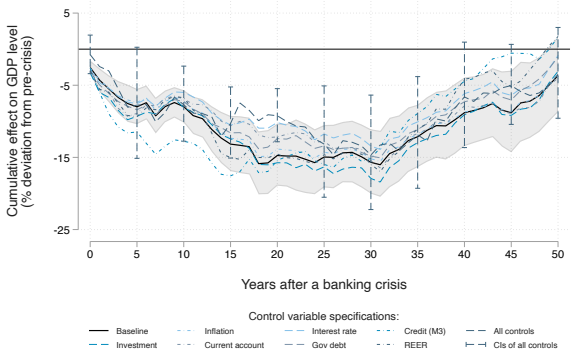

(b) Balanced panel - 5 pre-crisis GDP growth lags



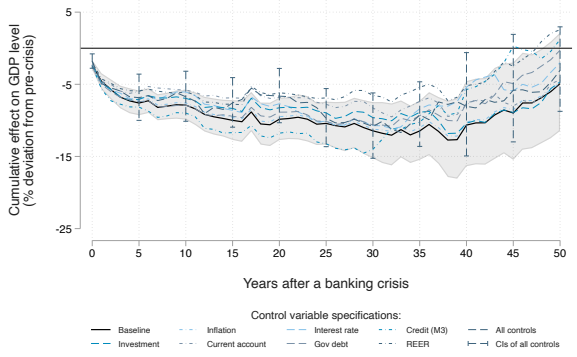

(c) Full Panel - 10 pre-crisis GDP growth lags




(d) Balanced panel - 10 pre-crisis GDP growth lags




(e) Full Panel - 25 pre-crisis GDP growth lags




(f) Balanced panel - 25 pre-crisis GDP growth lags

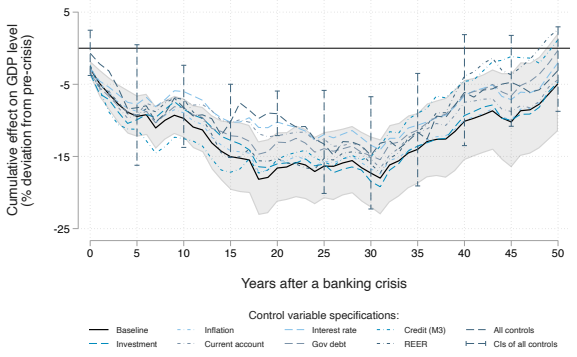
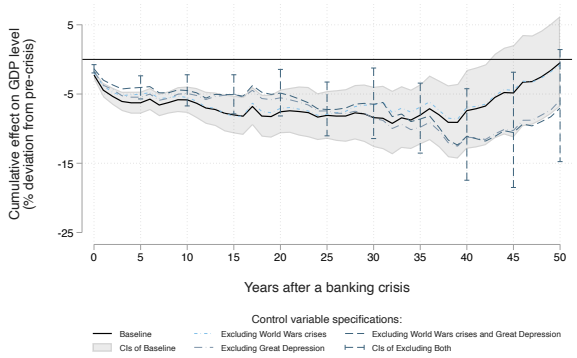


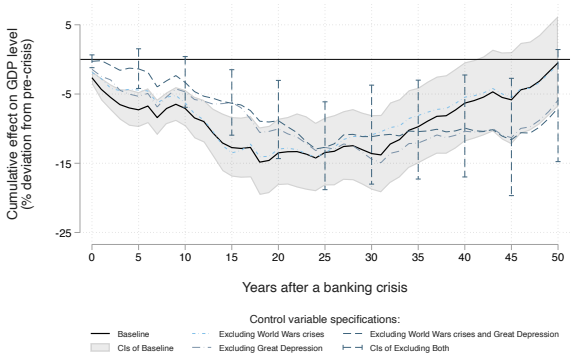
Figure B.5: Excluding World War & Great Depression

*This corresponds to Figure 4c in the main text.*

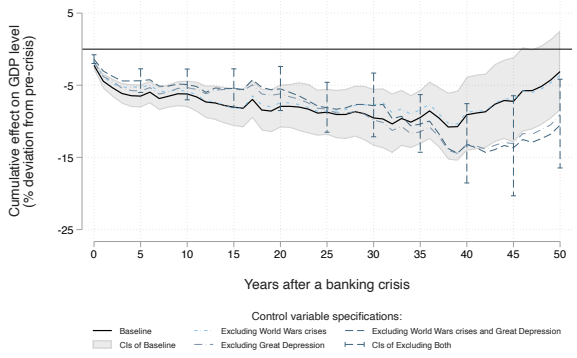
(a) Full Panel - 5 pre-crisis GDP growth lags



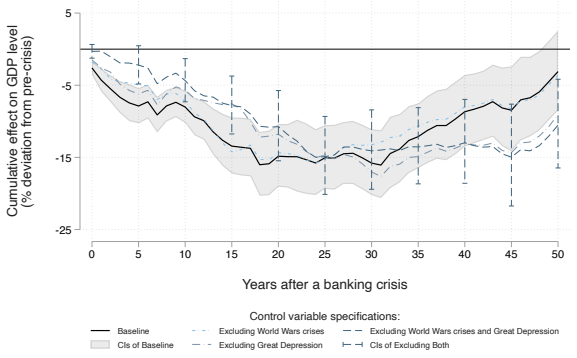

(b) Balanced panel - 5 pre-crisis GDP growth lags



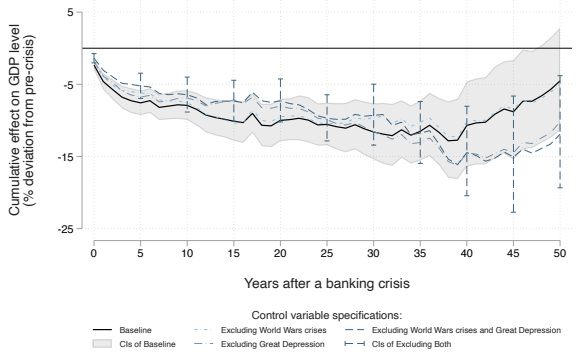

(c) Full Panel - 10 pre-crisis GDP growth lags




(d) Balanced panel - 10 pre-crisis GDP growth lags




(e) Full Panel - 25 pre-crisis GDP growth lags




(f) Balanced panel - 25 pre-crisis GDP growth lags

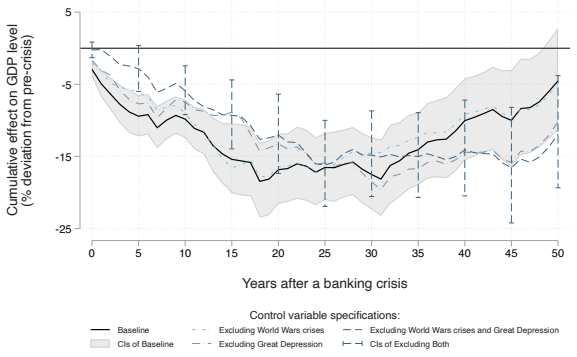
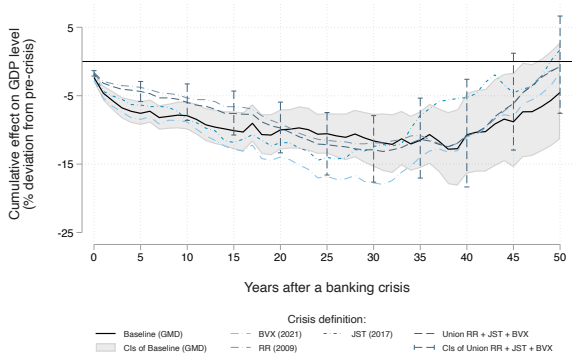


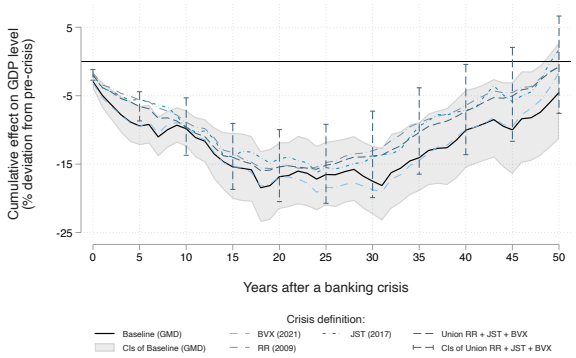
Figure B.6: Alternative Crisis Chronologies

*This corresponds to Figure 4d in the main text.*

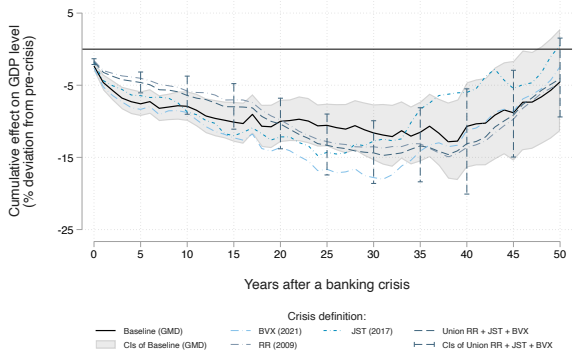
(a) Full Panel - 5 pre-crisis GDP growth lags



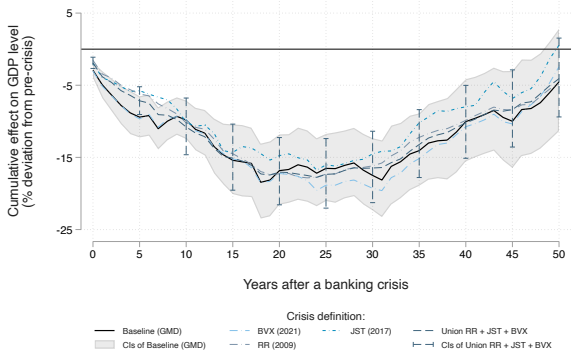

(b) Balanced panel - 5 pre-crisis GDP growth lags



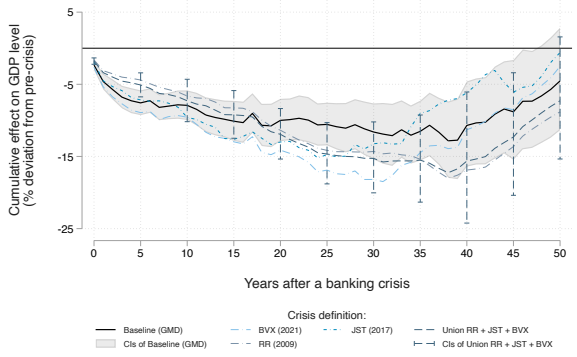

(c) Full Panel - 10 pre-crisis GDP growth lags




(d) Balanced panel - 10 pre-crisis GDP growth lags




(e) Full Panel - 25 pre-crisis GDP growth lags




(f) Balanced panel - 25 pre-crisis GDP growth lags

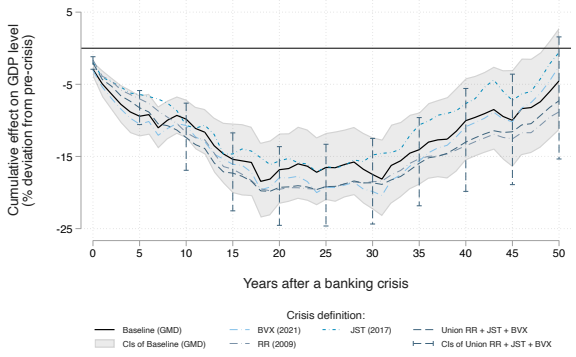
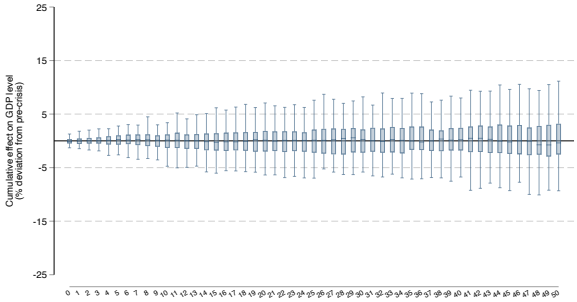


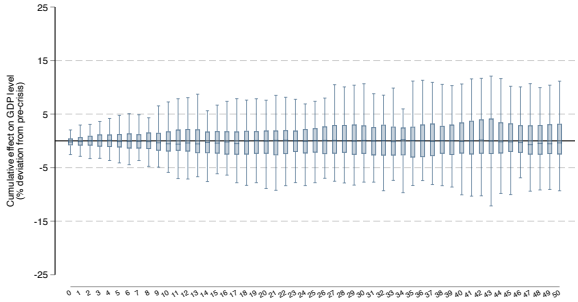
Figure B.7: Placebo Exercise

*This corresponds to Figure 5a in the main text.*

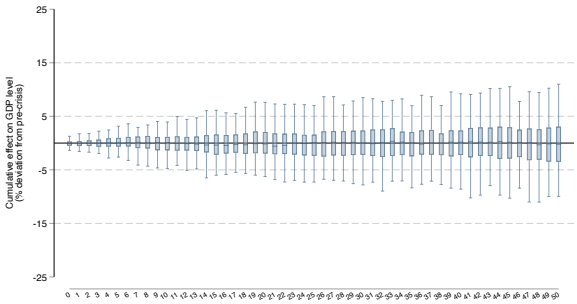
(a) Full Panel - 5 pre-crisis GDP growth lags



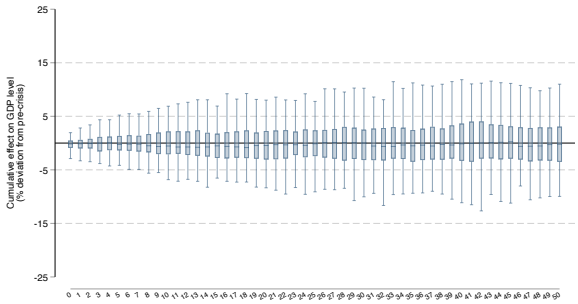

(b) Balanced panel - 5 pre-crisis GDP growth lags



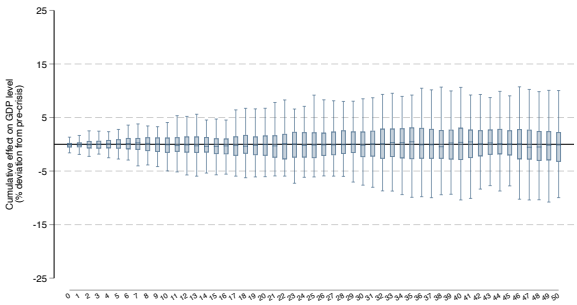

(c) Full Panel - 10 pre-crisis GDP growth lags




(d) Balanced panel - 10 pre-crisis GDP growth lags




(e) Full Panel - 25 pre-crisis GDP growth lags




(f) Balanced panel - 25 pre-crisis GDP growth lags

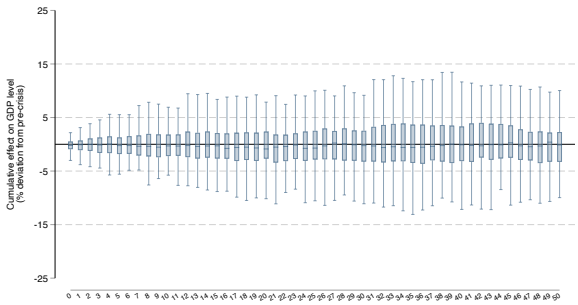
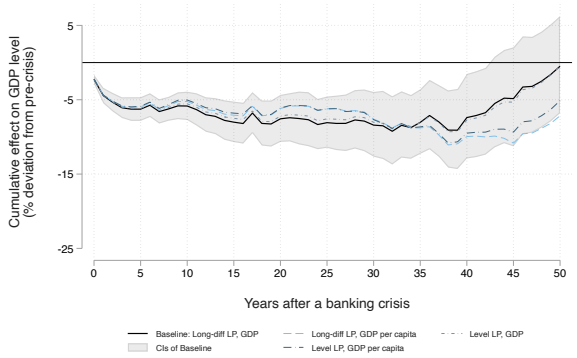


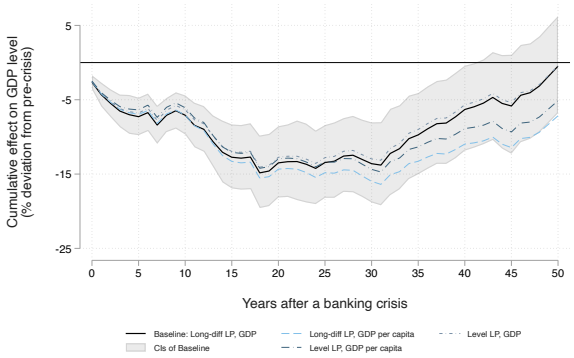
Figure B.8: Alternative LP Specifications

*This corresponds to Figure 5b in the main text.*

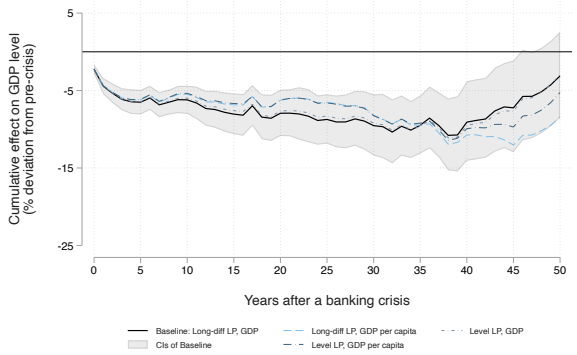
(a) Full Panel - 5 pre-crisis GDP growth lags



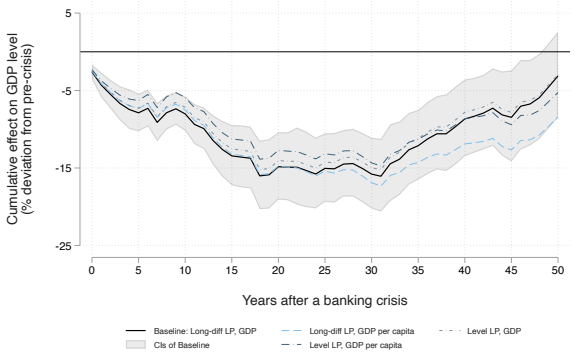

(b) Balanced panel - 5 pre-crisis GDP growth lags



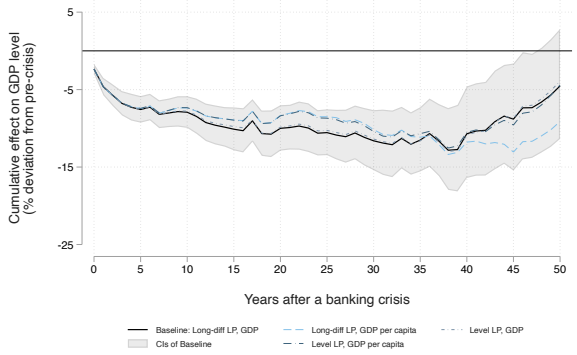

(c) Full Panel - 10 pre-crisis GDP growth lags




(d) Balanced panel - 10 pre-crisis GDP growth lags




(e) Full Panel - 25 pre-crisis GDP growth lags




(f) Balanced panel - 25 pre-crisis GDP growth lags

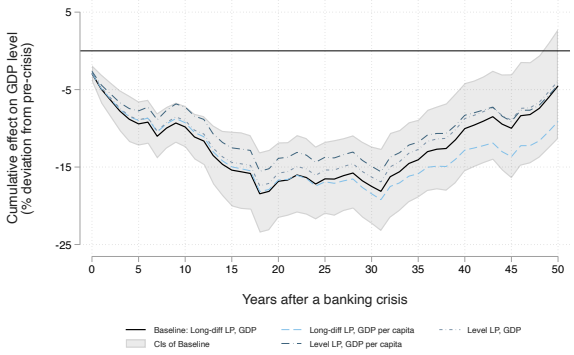
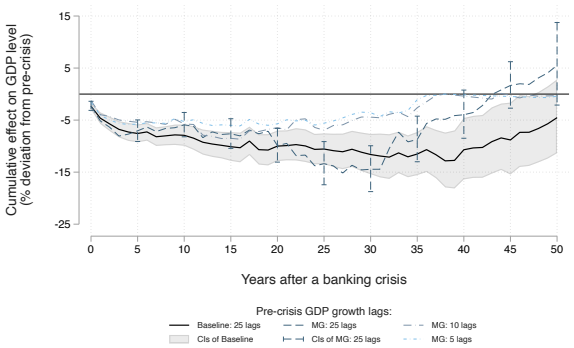


Figure B.9: Mean Group Estimator

*This corresponds to Figure 5c in the main text.*

(a) Full Panel




(b) Balanced panel

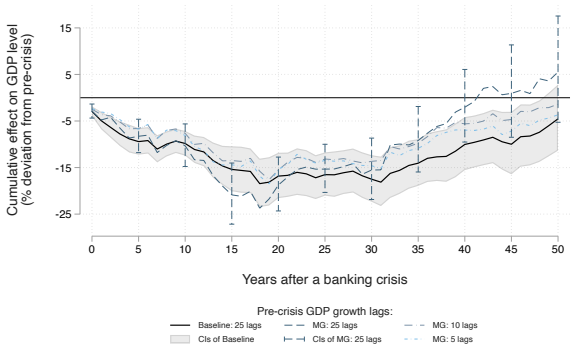
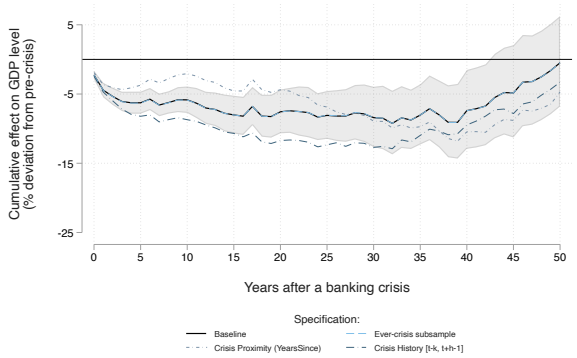


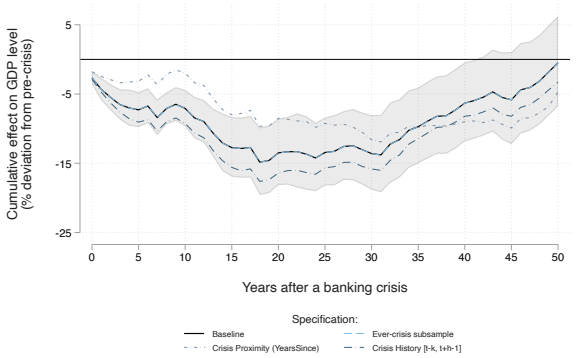
Figure B.10: Accounting for Subsequent Crises

*This corresponds to Figure 6b in the main text.*

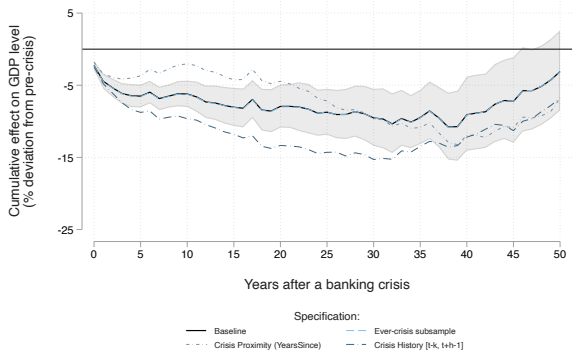
(a) Full Panel - 5 pre-crisis GDP growth lags



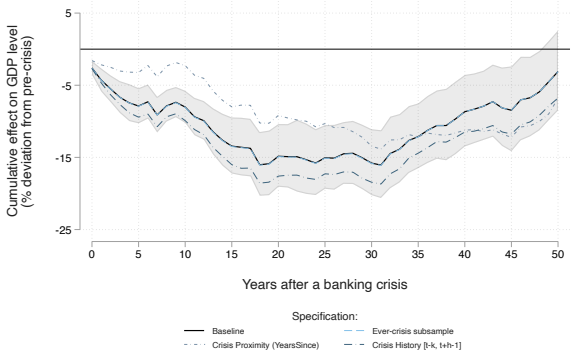

(b) Balanced panel - 5 pre-crisis GDP growth lags



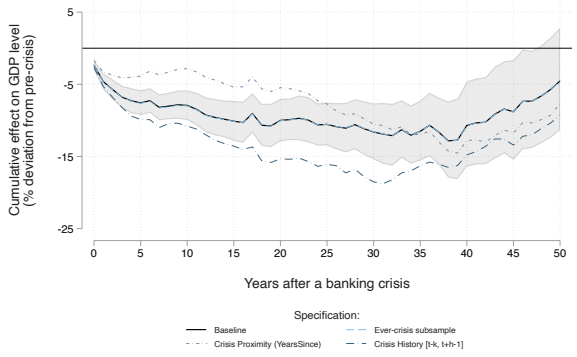

(c) Full Panel - 10 pre-crisis GDP growth lags




(d) Balanced panel - 10 pre-crisis GDP growth lags




(e) Full Panel - 25 pre-crisis GDP growth lags




(f) Balanced panel - 25 pre-crisis GDP growth lags

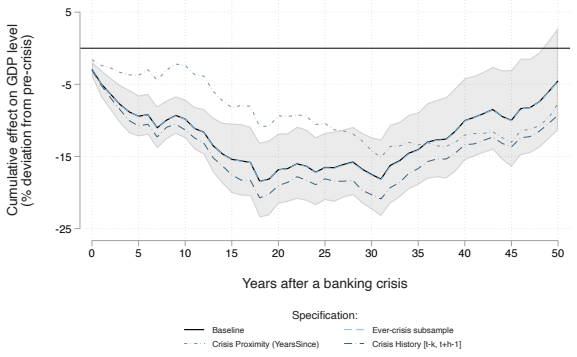
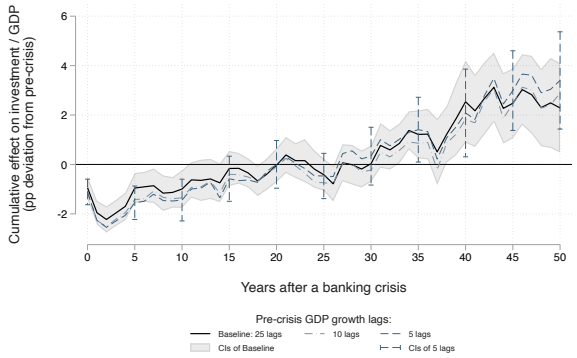


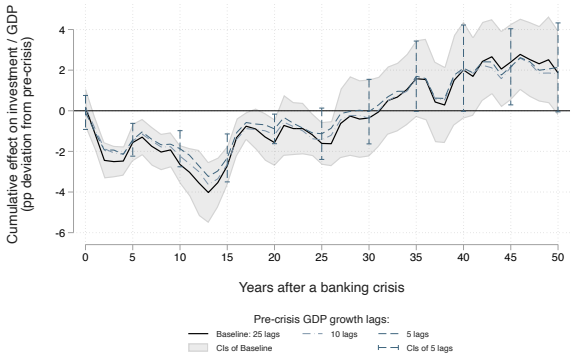
Figure B.11: Capital Accumulation after Banking Crises

*This corresponds to Figure 7 in the main text.*

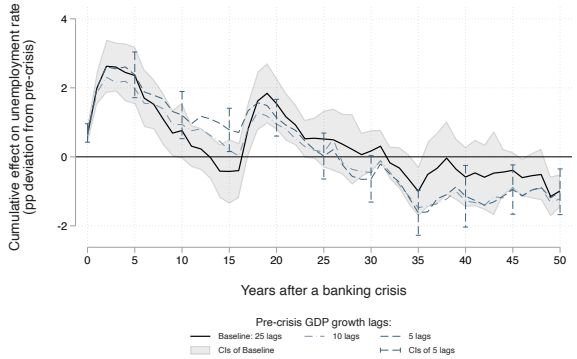
(a) Investment / GDP - Full Panel



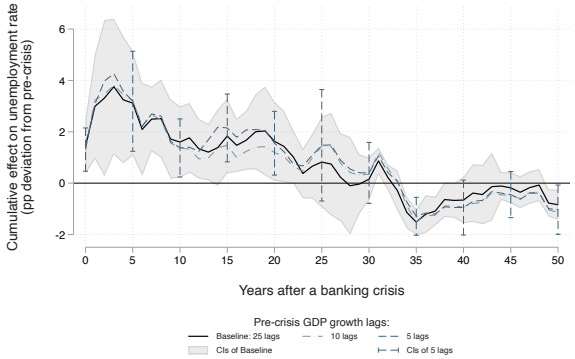

(b) Investment / GDP - Balanced panel



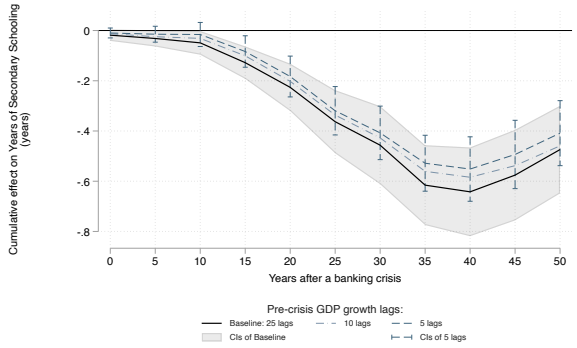

(c) Unemployment - Full Panel



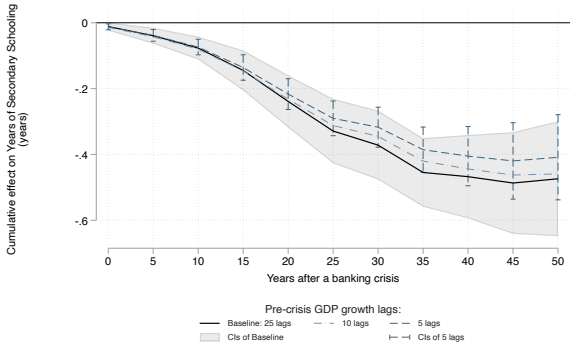

(d) Unemployment - Balanced panel



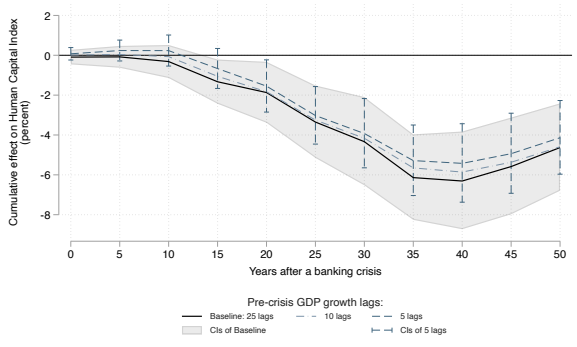

(e) Years of Schooling - Full Panel




(f) Years of Schooling - Balanced panel




(g) Human Capital Index - Full Panel




(h) Human Capital Index - Balanced panel

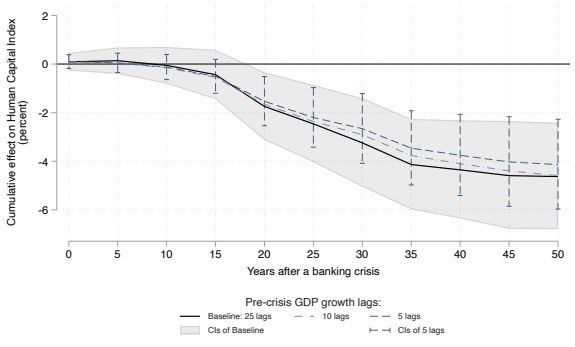
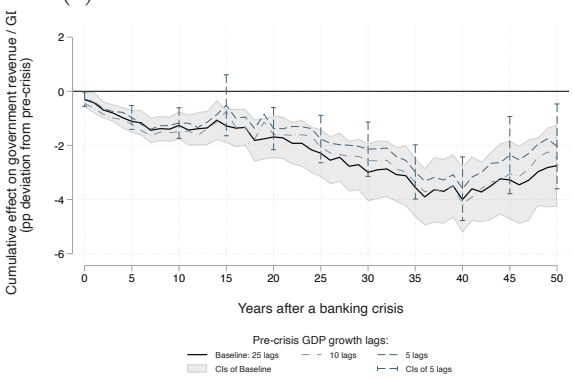


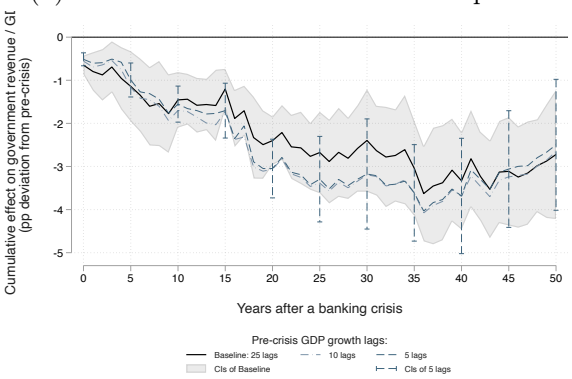
Figure B.12: Fiscal Dynamics Following Banking Crises

*This corresponds to Figure 8 in the main text.*

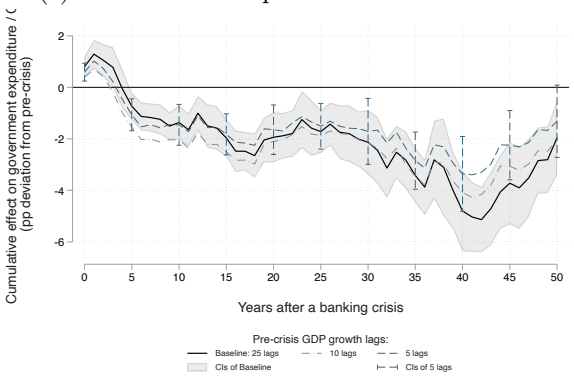
(a) Government Revenues - Full Panel



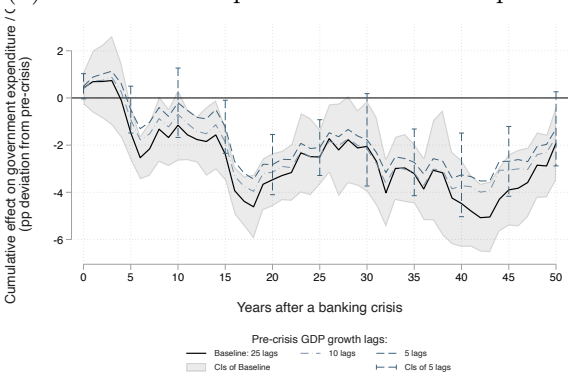

(b) Government Revenues - Balanced panel



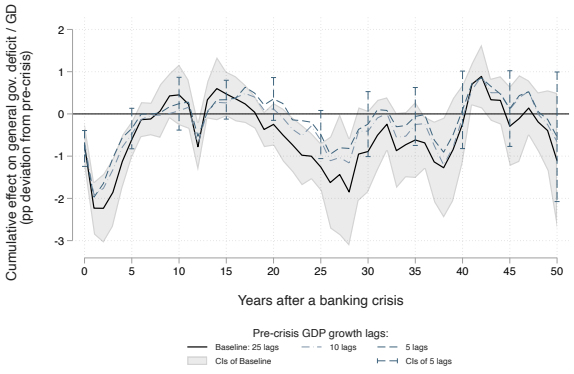

(c) Government Expenditures - Full Panel



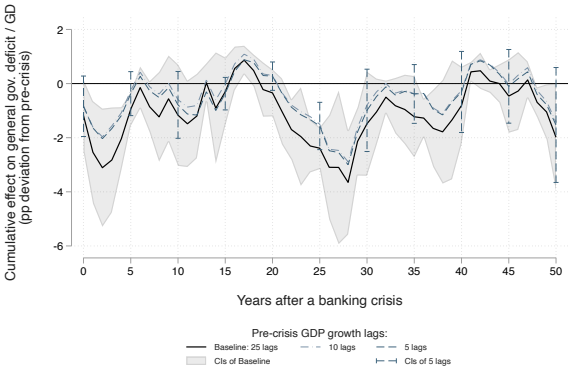

(d) Government Expenditures - Balanced panel



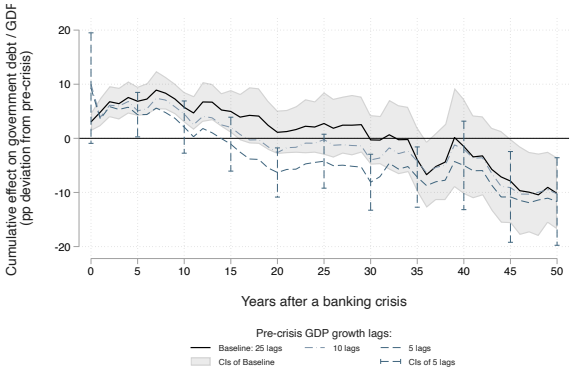

(e) Fiscal Balance - Full Panel




(f) Fiscal Balance - Balanced panel




(g) Government Debt - Full Panel




(h) Government Debt - Balanced panel

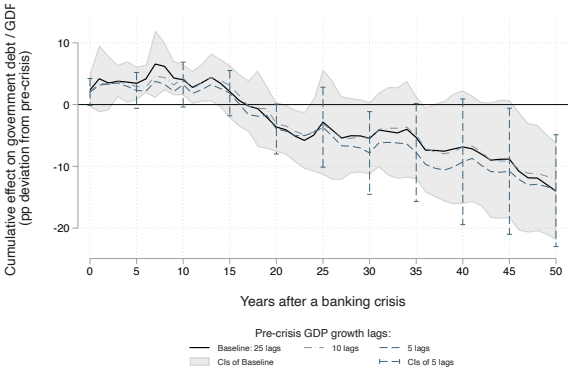
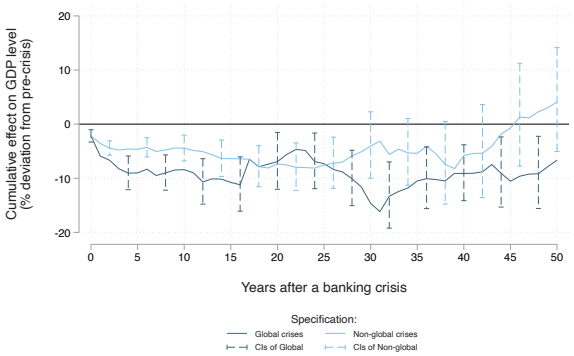
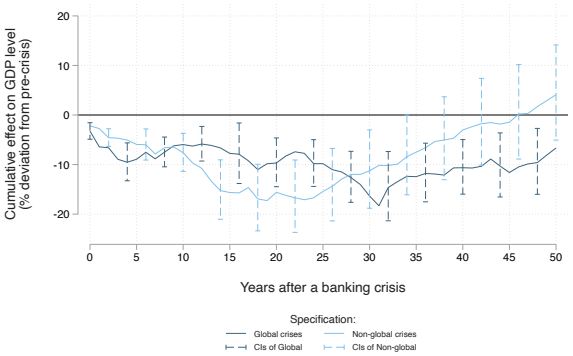


Figure B.13: Global Crises by GDP Share

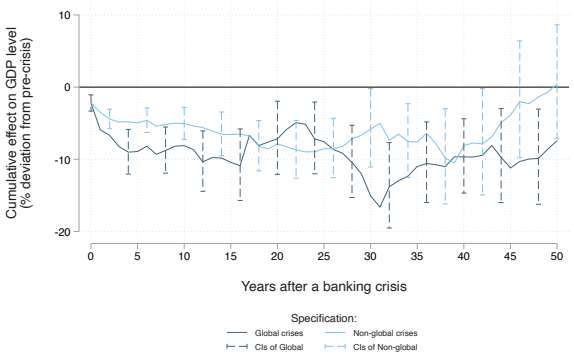
(a) Full Panel - 5 pre-crisis GDP growth lags



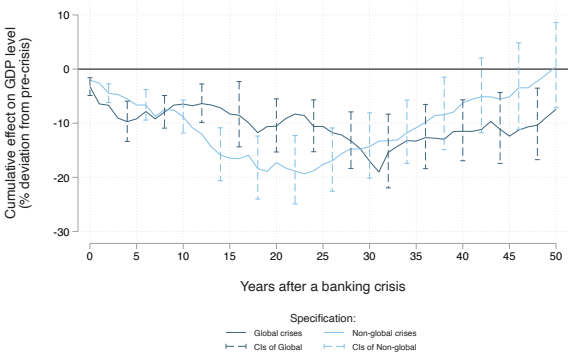

(b) Balanced panel - 5 pre-crisis GDP growth lags



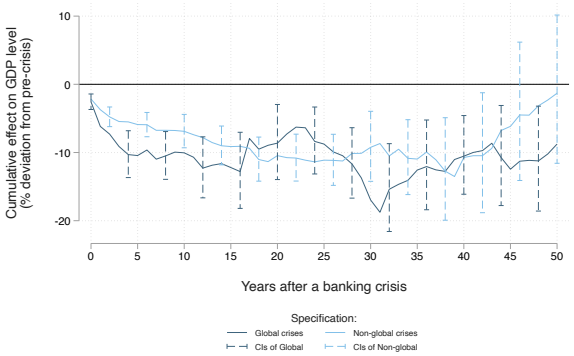

(c) Full Panel - 10 pre-crisis GDP growth lags




(d) Balanced panel - 10 pre-crisis GDP growth lags




(e) Full Panel - 25 pre-crisis GDP growth lags




(f) Balanced panel - 25 pre-crisis GDP growth lags

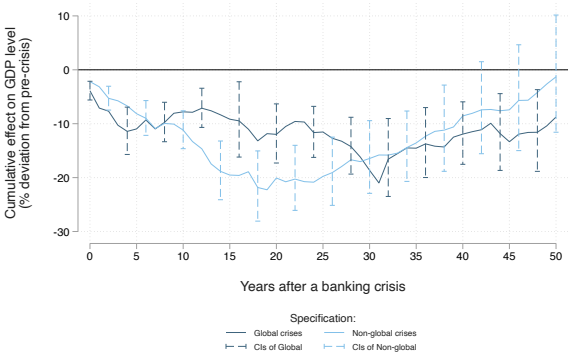
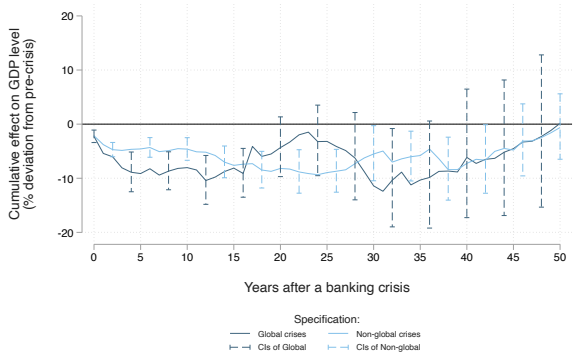
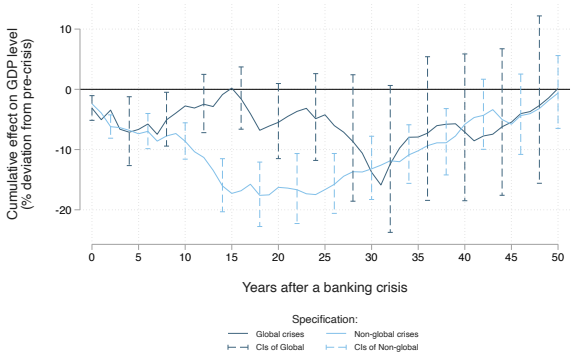


Figure B.14: Global Crises by Financial Center

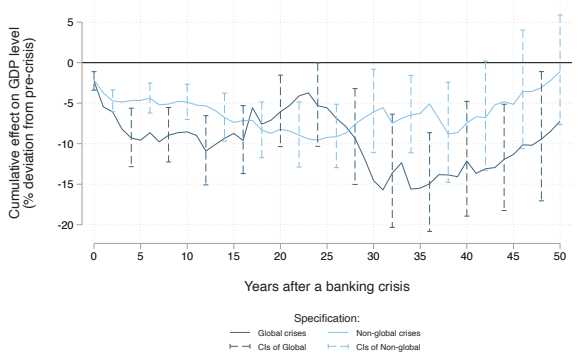
(a) Full Panel - 5 pre-crisis GDP growth lags



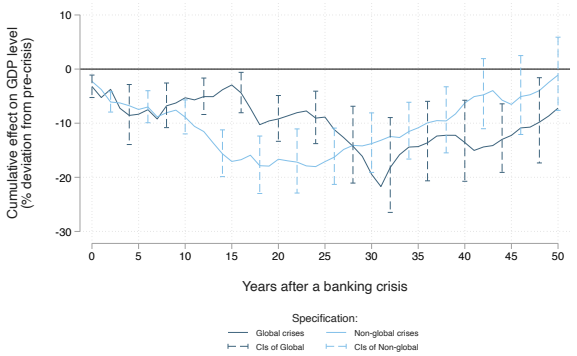

(b) Balanced panel - 5 pre-crisis GDP growth lags



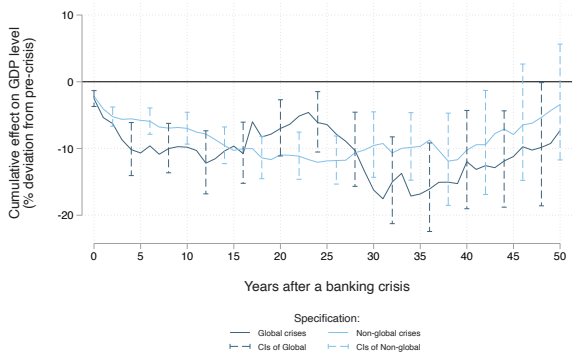

(c) Full Panel - 10 pre-crisis GDP growth lags




(d) Balanced panel - 10 pre-crisis GDP growth lags




(e) Full Panel - 25 pre-crisis GDP growth lags




(f) Balanced panel - 25 pre-crisis GDP growth lags

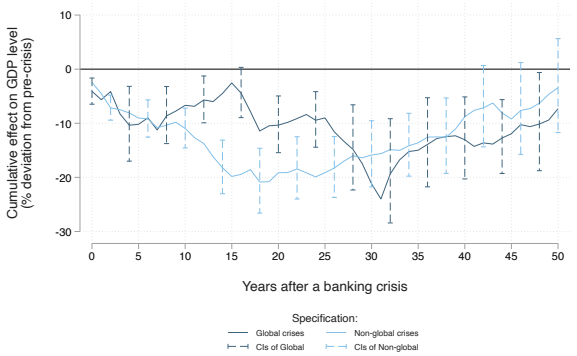
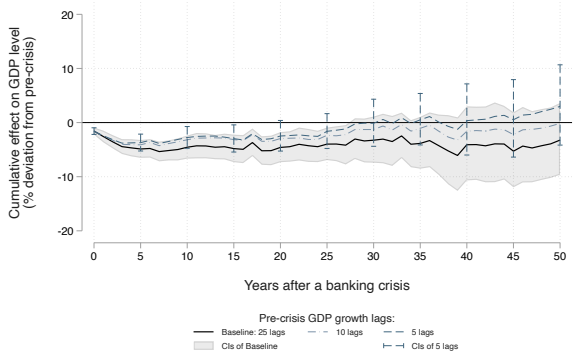
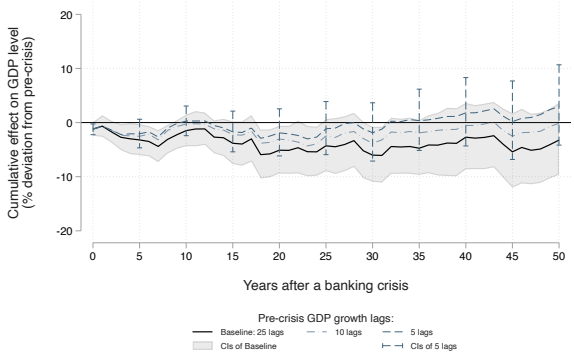


Figure B.15: Years Fixed Effects

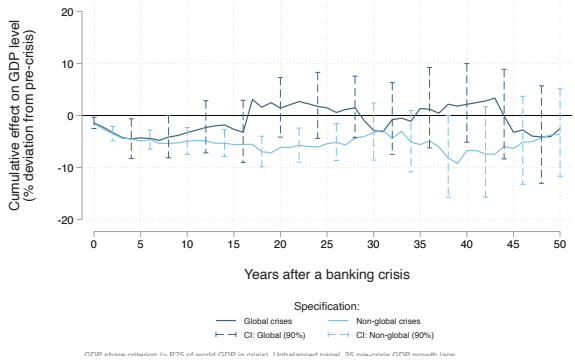
(a) Full Panel



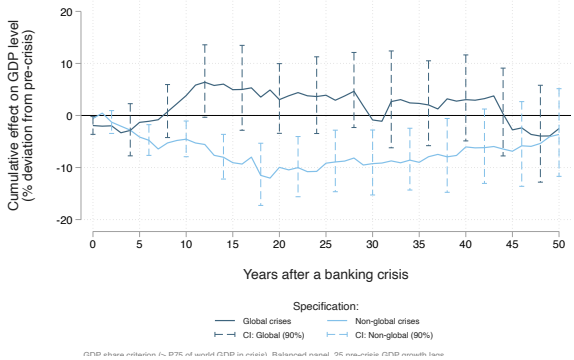

(b) Balanced panel



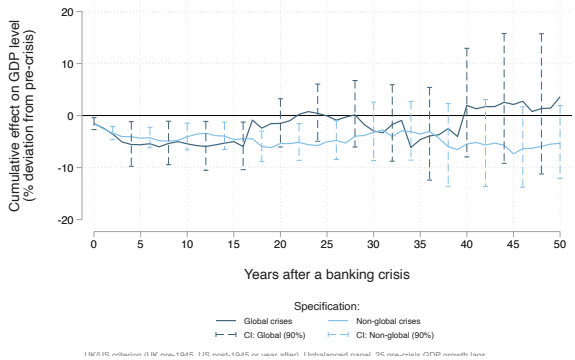

(c) Full Panel - Global Crisis (P75)



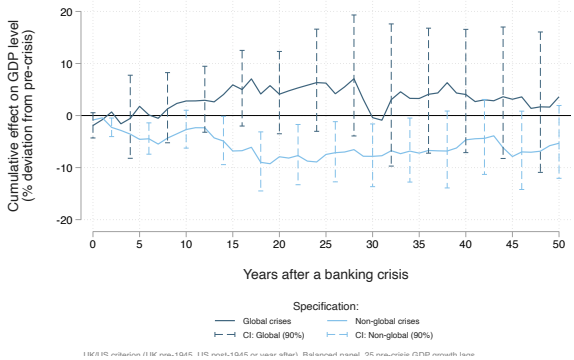

(d) Balanced panel - Global Crises (P75




(e) Full Panel - Global Crises (UK-US)




(f) Balanced panel - Global Crises (UK-US)



# References for Online Appendix


**Abdulai, Ibrahim.** 2022. "Monetary Policy, Inflation Stabilisation and Output Growth in Ghana, 1980–2017." PhD diss. University of Ibadan, Table 2.2. Data from Bank of Ghana Annual Reports and Statistical Bulletins.

**Abildgren, Kim.** 2017. "A Chart & Data Book on the Monetary and Financial History of Denmark." *Working Paper*.

**African Development Bank.** 2018. "AFDB Socio Economic Database." Accessed on 2024-07-20. Available at https://dataportal.opendataforafrica.org/nbyenxf/afdb-socio-economic-database-1960-2022.

**African Union.** 2024. "AFRISTAT: The Economic and Statistical Observatory for Sub-Saharan Africa." Accessed on 2024-07-20. Available at https://afristat.opendataforafrica.org.

**Albers, Thilo N. H.** 2018. "The Prelude and Global Impact of the Great Depression: Evidence from a New Macroeconomic Dataset." *Explorations in Economic History*, 70: 150–163.

**Albers, Thilo N.H., Morten Jerven, and Marvin Suesse.** 2023. "The Fiscal State in Africa: Evidence from a Century of Growth." *International Organization*, 77(1): 65–101.

**Andersson, Per F.** 2026. "Government Revenue Data." Accessed on 2026-03-23. Available at https://www.perfandersson.com/data.html.

**Arab Monetary Fund.** 2024. "Arab Economic Database." Accessed on 2024-07-10. Available at https://www.amf.org.ae/ar.

**Arroyo Abad, Leticia, Peter H. Lindert, et al.** 2015. "Latin American Fiscal History: Country Budget Datasets (Argentina, Colombia, Costa Rica, Peru, Uruguay)." Compiled for the UC Davis Global Price and Income History Group.

**Asian Development Bank.** 2024. "Key Indicators Database." Accessed on 2024-07-20. Available at https://kidb.adb.org.

**Australian Bur. of Statistics.** 2008. "Australian Historical Population Statistics, 2008." Catalogue 3105.0.65.001, released 5 August 2008. Accessed on 2026-06-11. Available at https://www.abs.gov.au/AUSSTATS/abs@.nsf/mf/3105.0.65.001.

**Baffigi, Alberto.** 2013. "National Accounts, 1861–2011." In *The Oxford Handbook of the Italian Economy Since Unification.* , ed. Gianni Toniolo, Chapter 2, 157–186. Oxford:Oxford University Press. Dataset: La contabilitá nazionale dell'Italia, 1861–2011.

**Banca d'Italia.** 2011. "Supplements to the Statistical Bulletin – Monetary and Financial Indicators – The Public Finances, borrowing requirement and debt." Volume XXI, 15 December 2011. Annual general and central government gross debt 1861–2010.

**Banca d'Italia.** 2024. "Historical Exchange Rates Database." Accessed on 2024-01-16. Available at https://tassidicambio.bancaditalia.it/terzevalute-wf-ui-web/converter.

**Banco Central de Reserva, Perú.** 2022. "BCRP 100 Años: Series Estadísticas." Accessed on 2026-03-31. Available at https://www.bcrp.gob.pe/.

**Banco de España.** 2024. "Historical and Statistical Database - Historical Series of the Spanish Economy." Accessed on 2024-01-16. Available at https://repositorio.bde.es/handle/123456789/23358.

**Banco de la República, Colombia.** 2024. "Series Estadísticas Históricas." Accessed on 2026-03-31. Available at https://uba.banrep.gov.co/htmlcommons/SeriesHistoricas/.

**Banco de la República, Colombia.** 2026. "Series Estadísticas Históricas — Unidad de Bases de Datos Apertura (UBA)." Accessed on 2026-06-11. Available at https://uba.banrep.gov.co.

**Bank Al-Maghrib.** 2024. "Séries Statistiques Monétaires." *https://www.bkam.ma/Statistiques/Statistiques-monetaires/Series-statistiques-monetaires*, Accessed: 2024-07-24.

**Bank Indonesia.** 2023. "Indonesian Economic and Financial Statistics." Accessed on 2024-11-15. Available at https://www.bi.go.id/en/statistik/ekonomi-keuangan/seki/Default.aspx#headingOne.

**Bank of Algeria.** 2023. "Indicateurs Monétaires." Accessed on 2024-07-24. Available at https://www.bank-of-algeria.dz/situation-monetaire.

**Bank of Israel.** 2001. "2000 Statistical Data." Historical statistical appendix tables. Accessed on 2026-04-26. Available at https://www.boi.org.il/en/communication-and-publications/regular-publications/annual-report-statistical-appendix-and-complementary-data/2000-statistical-data/.

**Bank of Japan.** 2024. "Historical Statistics." Accessed on 2024-10-07. Available at https://www.imes.boj.or.jp/en/historical/hstat/hstat.html.

**Banque de France.** 2024*a*. "Le Portail Statistique de la Banque de France." Accessed on 2024-07-23. Available at https://webstat.banque-france.fr/.

**Banque de France.** 2024*b*. "Statistical Series of Member Countries from Monetary Cooperation Africa-France." Accessed on 2024-10-21. Available at https://www.banque-france.fr/en/publications-and-statistics/statistics.

**Banushi, Xhesika, Alexandra Bykova, Magdalena Frei, Artem Kochnev, Isilda Mara, Manuel Neubauer, Renate Prasch, Hana Ruskova, Monika Schwarzhappel, and David Zenz.** 2024. "Introducing the wiiw COMECON Dataset." The Vienna Institute for International Economic Studies (wiiw) wiiw Statistical Report 13. Accessed on 2026-03-06. Available at https://comecon.wiiw.ac.at/.

**Baptista, Asdrúbal.** 1997. "Bases Cuantitativas de la Economía Venezolana 1830-1995." Programmatic table extraction via pdfplumber. All monetary values converted to current Bolívar Digital (Bs.D, post-Oct-2021) via cumulative reform factor $10^{14}$.

**Baron, Matthew, Emil Verner, and Wei Xiong.** 2020. "Banking Crises Without Panics*." *The Quarterly Journal of Economics*, 136(1): 51–113.

**Baron, Matthew, Emil Verner, and Wei Xiong.** 2021. "Banking crises without panics." *The Quarterly Journal of Economics*, 136(1): 51–113.

**Barro, Robert J, and José F Ursúa.** 2008. "Macroeconomic Crises Since 1870." National Bureau of Economic Research.

**BCEAO.** 2024. "La Base des Données Economiques et Financiéres." Accessed on 2024-07-15. Available at https://www.bceao.int/fr/content/la-base-des-donnees-economiques-et-financieres.

**Bértola, Luis, and María Rey.** 2018. "The Montevideo-Oxford Latin American Economic History Database (MOxLAD): Origins, Contents and Sources." *Economic History of Developing Regions*, 33(3): 209–224.

**BIS.** 2025*a*. "Bilateral Exchange Rates." *https://data.bis.org/topics/XRU*, Data retrieved via BIS Statistics API.

**BIS.** 2025*b*. "Central Bank Policy Rates." *https://data.bis.org/topics/CBPOL*, Data retrieved via BIS Statistics API.

**BIS.** 2025*c*. "Consumer Prices." *https://data.bis.org/topics/CPI*, Data retrieved via BIS Statistics API.

**BIS.** 2025*d*. "Effective Exchange Rates." *https://data.bis.org/topics/EER*, Data retrieved via BIS Statistics API.

**BIS.** 2025*e*. "Residential Property Prices." *https://data.bis.org/topics/RPP*, Data retrieved via BIS Statistics API.

**Bordo, Michael, Barry Eichengreen, Daniela Klingebiel, and Maria S. Martinez-Peria.** 2001. "Is the Crisis Problem Growing More Severe?" *Economic Policy*, 16(32): 52–82.

**Broadberry, Stephen, and Leigh Gardner.** 2022. "Economic Growth in Sub-Saharan Africa, 1885–2008: Evidence from Eight Countries." *Explorations in Economic History*, 83(C).

**Carter, Susan B., Scott S. Gartner, Michael R. Haines, Alan L. Olmstead, Richard Sutch, and Gavin Wright.** 2006. *Historical Statistics of the United States: Millennial Edition.* Vol. 3, Cambridge University Press New York.

**CBS, and Univ. Groningen.** 2001. "Tweehonderd jaar statistiek in tijdreeksen 1800–1999." *https://www.cbs.nl/-/media/imported/documents/2000/51/200-jaar-stat-tijdreeksen.pdf*, Accessed: 2026-04-28.

**Central Bank of Türkiye.** 2024. "Central Bank of the Republic of Türkiye Official Website." Accessed on 2024-01-16. Available at https://www.tcmb.gov.tr.

**Cha, Myung Soo, Nak Nyeon Kim, Ki-Joo Park, and Yitaek Park.** 2022. *Historical Statistics of Korea.* Springer.

**Cogneau, Denis, Yannick Dupraz, and Sandrine Mesplé-Somps.** 2021. "Fiscal Capacity and Dualism in Colonial States: The French Empire 1830–1962." *The Journal of Economic History*, 81(2): 441–480.

**Consejo Monetario Centroamericano.** 2017. "Estadísticas Macroeconómicas Históricas 1960-2017." *Per-country Excel workbooks (CR/DO/GT/HN/NI/SV) plus the regional comparative file COMPARATIVO_PAISES.xlsx*, Files: CR_HISTORICO.xls, DO_HISTORICO.xls, GT_HISTORICO.xls, HN_HISTORICO.xls, NI_HISTORICO.xls, SV_HISTORICO.xls, COMPARATIVO_PAISES.xlsx, hosted at https://www.secmca.org/wp-content/uploads/2022/09/.

**Consejo Monetario Centroamericano.** 2026. "SECMCA Public API and Estadísticas Macroeconómicas Históricas 1960–2017." Hybrid source built from the public API and the historical workbook. Accessed on 2026-04-21. Available at https://secmca-api.secmca.org/simafir_api/ws/swagger-ui/index.html and https://www.secmca.org/estadisticas-historicas-1960-2017/.

**Cox, Gary W, and Mark Dincecco.** 2021. "The Budgetary Origins of Fiscal-Military Prowess." *The Journal of Politics*, 83(3): 851–866.

**Darvas, Zsolt.** 2021. "Timely Measurement of Real Effective Exchange Rates." Bruegel Bruegel Working Papers.

**Deutsche Bundesbank.** 1976. "Deutsches Geld- und Bankwesen in Zahlen 1876-1945." HISTAT / GESIS digitized tables, ZA-Nr. 8222.

**Deutsche Bundesbank.** 1998. "50 Jahre Deutsche Mark." HISTAT / GESIS digitized tables, ZA-Nr. 8186.

**Diebolt, Claude.** 2003. "Accounting Expenditures on Education. Japan from Meiji Restoration to the Second World War." *Historical Social Research*, 28(1/2): 290–305.

**Eichholtz, Piet M.A.** 1997. "A Long Run House Price Index: The Herengracht Index, 1628-1973." *Real Estate Economics*, 25(2): 175–192.

**Eitrheim, Øyvind, Jan T. Klovland, and Jan F. Qvigstad.** 2023. "Historical Monetary and Financial Statistics for Norway." Norges Bank 57/2022.

**Ellison, Martin, Sang Seok Lee, and Kevin Hjortshøj O'Rourke.** 2024. "The Ends of 27 Big Depressions." *American Economic Review*, 114(1): 134–168.

**European Commission.** 2024. "AMECO Database." Accessed on 2024-11-06. Available at https://economy-finance.ec.europa.eu/economic-research-and-databases/economic-databases/ameco-database_en.

**Eurostat.** 2013. *Handbook on Residential Property Prices Indices. World Bank Publications - Books*, The World Bank Group.

**Eurostat.** 2025. "Eurostat Database." *https://ec.europa.eu/eurostat/databrowser/explore/all/all_themes*, Data retrieved via Eurostat API.

**FAO.** 2024. "FAOSTAT: GDP Data." *https://www.fao.org/faostat/*, Accessed on 2024-12-29. Available at https://www.fao.org/faostat/.

**Federico, G., and A. Tena-Junguito.** 2025. "World Population 1800-1938." IFCS - Working Papers in Economic History 2025-1.

**Federico, Giovanni, and Antonio Tena-Junguito.** 2019. "World Trade, 1800–1938: A New Synthesis." *Revista de Historia Económica / Journal of Iberian and Latin American Economic History*, 37(1): 9–41.

**Fed. Reserve Bank of St. Louis.** 2024. "Federal Reserve Economic Data." Accessed on 2024-05-13. Available at https://fred.stlouisfed.org/.

**Feenstra, Robert C., Robert Inklaar, and Marcel P. Timmer.** 2015. "The Next Generation of the Penn World Table." *American Economic Review*, 105(10): 3150–3182.

**Feinstein, Charles H.** 1972. *National Income, Expenditure and Output of the United Kingdom 1855–1965*. Cambridge University Press.

**Ferreres, Orlando J, et al.** 2005. "Dos Siglos de Economía Argentina."

**Flandreau, Marc, and Frédéric Zumer.** 2004. *The Making of Global Finance 1880-1913*. OECD.

**Fortin-Gagnon, Olivier, Maxime Leroux, Dalibor Stevanovic, and Stéphane Surprenant.** 2022. "A Large Canadian Database for Macroeconomic Analysis." *Canadian Journal of Economics/Revue canadienne d'économique*, 55(4): 1799–1833.

**Foster, R. A.** 1996. "Australian Economic Statistics 1949–50 to 1994–95." Reserve Bank of Australia Occasional Paper 8, Sydney. rev 1997.

**Fouquin, Michel, and Jules Hugot.** 2016. "Two Centuries of Bilateral Trade and Gravity Data: 1827-2014." CEPII Working Paper 2016-14.

**Gapminder.** 2024. "Gapminder: Total Population." Accessed on 2024-08-14. Available at http://gapm.io/dpop.

**Gardner, Leigh.** 2022. *Sovereignty Without Power: Liberia in the Age of Empires, 1822-1980*. Cambridge University Press.

**Global Financial Data.** 2024. "Global Financial Data Database." Accessed on 2024-10-22. Available at https://www.globalfinancialdata.com.

**Government of India.** 1922. "Statistical Abstract relating to British India." 55th number, 1910-11 to 1919-20. Machine-readable Excel from Digital South Asia Library, University of Chicago.

**Grimm, Maximilian.** 2024. "The Effect of Monetary Policy on Systemic Bank Funding Stability." University of Bonn and University of Cologne, Germany ECONtribute Discussion Papers Series 341.

**Grytten, Ola H.** 2022. "Revising Growth History: New Estimates of GDP for Norway, 1816–2019." *Economic History Review*, 75(1): 181–202.

**Ha, Jongrim, M. Ayhan Kose, and Franziska Ohnsorge.** 2023. "One-Stop Source: A Global Database of Inflation." *Journal of International Money and Finance*, 137: 102896.

**Herranz-Loncan, Alfonso, and Jose Alejandro Peres-Cajias.** 2016. "Tracing the Reversal of Fortune in the Americas. Bolivian GDP per capita since the mid-nineteenth century." *Cliometrica*, 10(1): 99–128. Accessed on 2026-03-19. Available at https://joseperescajias.com/data/.

**Hist. Statistics of Switzerland.** 2012. "Historical Statistics of Switzerland - National Accounting." Accessed on 2024-10-02. Available at https://hsso.ch.

**HM Treasury.** 2026. "Historical Public Finances Database." Accessed on 2026-03-11. Available at https://www.gov.uk/government/statistics/historical-public-finances-database.

**Holtfrerich, Carl-Ludwig.** 1980. "Die deutsche Inflation 1914 bis 1923." HISTAT / GESIS digitized tables, ZA-Nr. 8374.

**Homer, S., and R.E. Sylla.** 1996. *A History of Interest Rates.* Rutgers University Press.

**Hong Kong Institute.** 2024. "Hong Kong Economic History Database." Accessed on 2026-03-31. Available at https://www.aof.org.hk/research/HKIMR/publications-and-research/hong-kong-economic-history-database.

**Hubmann, Georg, Clemens Jobst, and Michael Maier.** 2020. "A new long-run consumer price index for Austria, 1800–2018." *OeNB Monetary Policy & the Economy*, , (Q3/2020). Accessed on 2026-03-19. Available at https://journaldata.zbw.eu/dataset/ein-neuer-langer-verbraucherpreisindex-fur-osterreich-1800-2018-replication-data.

**Hutchinson, Diane, and Florian Ploeckl.** 2024. "What Was the Australian GDP or CPI Then?" Accessed on 2024-04-21. Available at https://www.measuringworth.com/datasets/australiadata/.

**IEC, Hitotsubashi University.** 2020. "Asian Historical Statistics: Basic Data for Asian Economic History."

**IER, Hitotsubashi University.** 2019. "Asian Historical Statistics, Volume 4: Korea." *https://d-infra.ier.hit-u.ac.jp/English/ltes/a000-asia-long-kr.html*, Database compiled by Mizoguchi, T., Pyo, H.K., and Moon, H. Accessed: December 28, 2025.

**Inklaar, Robert, Harmen de Jong, Jutta Bolt, and Jan van Zanden.** 2018*a*. "Rebasing 'Maddison': New Income Comparisons and the Shape of Long-Run Economic Development." Groningen Growth and Development Centre, University of Groningen GGDC Research Memorandum GD-174.

**Inklaar, Robert, Harmen de Jong, Jutta Bolt, and Jan van Zanden.** 2018*b*. "Rebasing 'Maddison': New Income Comparisons and the Shape of Long-Run Economic Development." Groningen Growth and Development Centre, University of Groningen GGDC Research Memorandum GD-174.

**Inst. Nacional de Estadística.** 2024. "Spanish Statistical Office." Accessed on 2024-01-16.

**Inst. Nacional de Estatística.** 2001. *Portuguese Historical Statistics.* Lisboa:INE.

**International Monetary Fund.** 2010. "A Historical Public Debt Database." International Monetary Fund IMF Working Papers 2010/245.

**International Monetary Fund.** 2016. "Monetary and Financial Statistics Manual and Compilation Guide." International Monetary Fund.

**International Monetary Fund.** 2024. "Government Finance Statistics." Accessed on 2024-08-05. Available at https://data.imf.org/?sk=a0867067-d23c-4ebc-ad23-d3b015045405.

**International Monetary Fund.** 2025*a*. "IMF Data Portal." *https://data.imf.org/*, Accessed: 2025-09-22.

**International Monetary Fund.** 2025*b*. "World Economic Outlook Database." *https://data.imf.org/*, Accessed: 2025-09-22.

**Internl Inst. Social History.** 2024. "Clio Infra: Repository of Global Inequality Data." Accessed on 2024-01-16. Available at https://clio-infra.eu.

**Internl Labour Organization.** 2024. "Unemployment Rate - ILO Modelled Estimates." Accessed on 2024-01-07. Available at https://ilostat.ilo.org/data.

**IPEA.** 2024. "IPEA Data." Accessed on 2024-10-08. Available at http://www.ipeadata.gov.br.

**Ist. Naz. di Statistica.** 2024. "Italian National Institute of Statistics Official Website." Accessed on 2024-01-16. Available at https://www.istat.it.

**Ivanov, Martin.** 2012. *The Gross Domestic Product of Bulgaria 1870–1945*. National Statistical Institute of Bulgaria.

**Jones, Matthew T., and Maurice Obstfeld.** 1997. "Saving, Investment, and Gold: A Reassessment of Historical Current Account Data." University of California at Berkeley Center for International and Development Economics Research (CIDER) Working Papers C97-094.

**Jónsson, Guðmundur, Magnús S Magnússon, and Hallgrímur Snorrason.** 1997. *Hagskinna: Sögulegar Hagtölur um Ísland*. Hagstofa Íslands.

**Jordà, Òscar, Moritz Schularick, and Alan M Taylor.** 2017*a*. "Macrofinancial history and the new business cycle facts." *NBER macroeconomics annual*, 31(1): 213–263.

**Jordá, Óscar, Moritz Schularick, and Alan M. Taylor.** 2017*b*. "Macrofinancial History and the New Business Cycle Facts." *NBER Macroeconomics Annual*, 31(1): 213–263.

**Kehoe, Timothy J., and Juan Pablo Nicolini.** 2022. *A Monetary and Fiscal History of Latin America, 1960–2017*. University of Minnesota Press. Accessed on 2026-03-16. Available at https://bfilatinamerica.uchicago.edu/entities/monetary-and-fiscal-history-of-latin-america/.

**Laeven, Luc, and Fabian Valencia.** 2013*a*. "Systemic banking crises database." *IMF Economic Review*, 61(2): 225–270.

**Laeven, Luc, and Fabián Valencia.** 2013*b*. "Systemic Banking Crises Database." *IMF Economic Review*, 61(2): 225–270.

**Lehbib, Mohamed, and Karsten Müller.** 2025. "GMD: The Easy Way to Access the World's Most Comprehensive Macroeconomic Database." Working Paper. Accessed: 2025-12-29.

**Levy-Garboua, Vivien, and Eric Monnet.** 2016. "Les Taux d'Intérêts en France: Une Perspective Historique." *Revue d'économie financiére*, 1(121): 35–58.

**Ljungberg, Jonas.** 2019. "Nominal and Real Effective Exchange Rates for Europe, 1870-2016: Some Methodological Issues." Lund University, Department of Economic History Lund Papers in Economic History 200.

**Lobell, Håkan.** 2025. "Swedish Historical National Accounts 1300–2022: 2025 Update." Lund University Department of Economic History. Based on Schön and Krantz (2015).

**Mack, Adrienne, and Enrique Martínez-García.** 2011. "A Cross-Country Quarterly Database of Real House Prices: A Methodological Note." Federal Reserve Bank of Dallas Globalization Institute Working Papers 99.

**Markevich, Andrei, and Mark Harrison.** 2011. "Great War, Civil War, and Recovery: Russia's National Income, 1913 to 1928." *Journal of Economic History*, 71(3): 672–703.

**Martín-Aceña, Pablo.** 2018. "Money in Spain. New Historical Statistics 1830–1998." Banco de España Documentos de Trabajo 1806.

**Martín Moreno, Sergio (curator).** 2026. "Archivo de Historia Económica de México — Base de Datos Histórica Siglo XX, 1876-2025." *File ABASEXX01.xlsx (Base de datos general)*, Curated compilation drawing on Cárdenas (1987), Solís (Colegio Nacional Vol. 10), Gómez Galvarriato & Musacchio (2000), Banco de México archives (Informes Anuales 1925-2010, Boletín Estadístico), Isoard (1979) SHCP, Aspe (1993), INEGI publications. Monetary values converted from old pesos to current new pesos via 1 N$ = 1000 old $ (Jan 1993 redenomination).

**Mauro, Paolo, Rafael Romeu, Ariel Binder, and Asad Zaman.** 2015. "A Modern History of Fiscal Prudence and Profligacy." *Journal of Monetary Economics*, 76(C): 55–70.

**Mbaye, Samba, Marialuz Moreno-Badia, and Kyungla Chae.** 2018. "Global Debt Database: Methodology and Sources." International Monetary Fund IMF Working Papers 2018/111.

**Ministerio de Economia.** 2024. "Informacion Economica al Dia - Apendice 8." Accessed on 2024-01-16. Available at https://www.economia.gob.ar/download/infoeco/apendice8.xlsx.

**Min. of Economy and Finance, Korea.** 2026. "Open Fiscal Data." Accessed on 2026-04-26. Available at https://www.openfiscaldata.go.kr/op/en/us/UOPENUSA02.

**Mitchell, Brian.** 2013. *International Historical Statistics. International Historical Statistics*. 1 ed., Palgrave Macmillan London.

**Müller, Karsten, Chenzi Xu, Mohamed Lehbib, and Ziliang Chen.** 2025. "The Global Macro Database: A New International Macroeconomic Dataset." National Bureau of Economic Research Working Paper 33714.

**Nakamura, Leonard I., and Carlos E. Zarazaga.** 2001. "Banking and Finance in Argentina in the Period 1900–35." Federal Reserve Bank of Dallas Center for Latin America Working Papers 0501.

**National Bank of Belgium.** 2024. "Statistical Data." Accessed on 2024-11-15. Available at https://stat.nbb.be/?lang=en.

**National Bank of Serbia.** 2024. "South-Eastern European Monetary and Economic Statistics from the 19th Century to World War II." Accessed on 2024-06-24. Available at https://www.nbs.rs/en/drugi-nivo-navigacije/publikacije-i-istrazivanja/seemhn/seemhn-dctf/.

**National Statistics, Taiwan.** 2024. "Taiwan's Statistical Bureau." Accessed on 2024-10-02. Available at https://eng.stat.gov.tw/ct.asp?xItem=37408&CtNode=5347&mp=8.

**Natl Bureau of Statistics, China.** 2024. "National Data." Accessed on 2024-11-15. Available at https://data.stats.gov.cn/english/index.htm.

**New Zealand Treasury.** 2026. "Fiscal Time Series — Historical Fiscal Indicators." Accessed: 2026-04-30.

**Ngakegni, Antoine.** 1991. "Agrégats Monétaires sur Longue Période en France, 1876-1990." *Journal de la Société de statistique de Paris*, 132(3): 19–60.

**NZ Inst. Economic Research.** 2026. "Data1850: Long-Run Statistics for New Zealand." Accessed: 2026-04-30.

**OECD.** 1986. "OECD Main Economic Indicators - Historical Statitics." Accessed on 2024-10-30. Available at https://archive.org/details/pub_oecd-main-economic-indicators-historical-statistics.

**OECD.** 2024*a*. "OECD Economic Outlook Data." Accessed on 2024-06-05. Available at https://www.oecd.org/en/topics/sub-issues/economic-outlook.html.

**OECD.** 2024*b*. "OECD Housing Prices Indicators Data." Accessed on 2024-07-10. Available at https://www.oecd.org/en/data/indicators/housing-prices.html.

**OECD.** 2024*c*. "OECD Key Economic Indicators Data." Accessed on 2024-07-10. Available at https://www.oecd-ilibrary.org/economics/data/main-economic-indicators/key-short-term-indicators_data-00039-en.

**OECD.** 2024*d*. "OECD Main Economic Indicators Data." Accessed on 2024-07-04. Available at https://web-archive.oecd.org/temp/2023-09-05/76384-oecdmaineconomicindicatorsmei.htm.

**Officer, Lawrence H., and Samuel H. Williamson.** 2024. "Measures of Worth." *MeasuringWorth*, Accessed on 2024-12-02. Available at https://www.measuringworth.com.

**Pamuk, Şevket.** 2004. "Prices in the Ottoman Empire, 1469-1914." *International Journal of Middle East Studies*, 36(3): 451–468. Mirror of IISG dataset at http://www.iisg.nl/hpw/istanbul.xls. Underlying methodology in Ozmucur and Pamuk (2002), Journal of Economic History.

**Peres-Cajias, Jose Alejandro.** 2014. "Bolivian Public Finances, 1882–2010. The Challenge to Make Social Spending Sustainable." *Revista de Historia Economica / Journal of Iberian and Latin American Economic History*, 32(1): 77–117. Accessed on 2026-03-19. Available at https://joseperescajias.com/data/.

**Piketty, Thomas, and Gabriel Zucman.** 2014. "Capital is Back: Wealth-Income Ratios in Rich Countries 1700-2010." *The Quarterly Journal of Economics*, 129(3): 1255–1310.

**Piselli, Paola, and Fabio Vercelli.** 2023. "I Titoli Pubblici nei Bilanci delle Banche 1890-2016." Banca d'Italia. Di prossima pubblicazione (forthcoming).

**Reinhart, Carmen M., and Kenneth S. Rogoff.** 2009*a*. "The Aftermath of Financial Crises." *The American Economic Review*, 99(2): 466–472.

**Reinhart, Carmen M, and Kenneth S Rogoff.** 2009*b*. *This time is different: Eight centuries of financial folly*. princeton university press.

**Reinhart, Carmen M., and Kenneth S. Rogoff.** 2010. "Inflation: Historical Consumer Price and Cost-of-Living Indices." *thttps://carmenreinhart.com/inflation/*, Data files: Inflation_Country_A-E.xls, Inflation_Country_F-M.xls, Inflation_Country_M-S.xls, Inflation_Country_T-Z.xls. Accessed: 2025-06-30.

**Ritschl, Albrecht.** 2002. *Deutschlands Krise und Konjunktur. Binnenkonjunktur, Auslandsverschuldung und Reparationsproblem zwischen Dawes-Plan und Transfersperre 1924-1934*. Berlin:Akademie-Verlag.

**Ritschl, Albrecht, and Mark Spoerer.** 1997. "Das Bruttosozialprodukt in Deutschland nach den amtlichen Volkseinkommens- und Sozialproduktsstatistiken 1901-1995." *Jahrbuch für Wirtschaftsgeschichte*, 1997(2): 27–54.

**Saudi Central Bank.** 2024. "Saudi Central Bank Official Website." Accessed on 2024-01-16. Available at https://www.sama.gov.sa.

**Schön, Lennart, and Olle Krantz.** 2017. "Swedish Historical National Accounts 1300-2020." School of Economics and Management, Lund University.

**Schuler, Kurt.** 2015. "Establishing an International Data Archive on Free Banking." The Johns Hopkins Institute for Applied Economics, Global Health, and the Study of Business Enterprise Studies in Applied Economics 24.

**Schulze, Max-Stephan.** 2000. "Patterns of Growth and Stagnation in the Late Nineteenth Century Habsburg Economy." *European Review of Economic History*, 4(3): 311–340.

**Sefton, James, and Martin Weale.** 1995. *Reconciliation of National Income and Expenditure: Balanced Estimates of National Income for the United Kingdom, 1920–1990.* Cambridge University Press.

**Sensch, Jurgen.** 2012. "Die Struktur der kassenmassigen Steuereinnahmen im Deutschen Reich und in der Bundesrepublik Deutschland (1872-2009)." HISTAT / GESIS digitized tables, ZA-Nr. 8530.

**Sivasubramonian, S.** 2000. *The National Income of India in the Twentieth Century.* New Delhi:Oxford University Press.

**Smits, Jan-Pieter, Pieter J. Woltjer, and Donghui Ma.** 2009. "A Dataset on Comparative Historical National Accounts, ca.1870-1950: A Time-series Perspective." Groningen Growth and Development Centre, University of Groningen GGDC Research Memorandum GD-107.

**South African Reserve Bank.** 2024. "South African Reserve Bank Official Website." Accessed on 2024-01-16. Available at https://www.resbank.co.za.

**Sprenger, Bernd.** 1982. "Geldmengenanderungen in Deutschland seit 1835." HISTAT / GESIS digitized tables, ZA-Nr. 8231.

**Statistics Canada.** 2024. "Historical Statistics of Canada: Section F: Gross National Product, the Capital Stock, and Productivity." Accessed on 2024-10-02. Available at https://www150.statcan.gc.ca/n1/pub/11-516-x/sectionf/4057751-eng.htm.

**Statistics Iceland.** 1997. *Hagskinna — Icelandic Historical Statistics from the 17th Century to 1990.* Reykjavík:Hagstofa Íslands.

**Statistics Poland.** 2024. "Bank Danych Makroekonomicznych (BDM)." Accessed on 2024-11-15. Available at https://bdm.stat.gov.pl.

**Statistisches Reichsamt.** 1925. "Zahlen zur Geldentwertung in Deutschland 1914 bis 1923." HISTAT / GESIS digitized tables, ZA-Nr. 8354.

**Stuart, Rebecca.** 2018. "A Long Run Macro Database for Ireland: 1922-2012."

**Swiss National Bank.** 2009. "Swiss National Bank Historical Time Series." Accessed on 2024-10-02. Available at https://www.snb.ch/en/the-snb/mandates-goals/statistics/statistics-pub/publication-history/historical-time-series.

**Taiwan Statistical Office.** 2025. "Macroeconomic Data." *https://eng.stat.gov.tw/Default.aspx*.

**Thomas, Ryland, Sally Hills, and Nicholas Dimsdale.** 2010. "The UK Recession in Context — What Do Three Centuries of Data Tell Us?" *Bank of England Quarterly Bulletin*, 50(4): 277–291.

**United Nations.** 2009. "System of National Accounts 2008." United Nations, New York.

**United Nations.** 2024*a*. "Economic Commission for Latin America and the Caribbean Statistics." Accessed on 2024-06-18. Available at https://www.cepal.org/en.

**United Nations.** 2024*b*. "National Accounts - Analysis of Main Aggregates (AMA)." Accessed on 2024-02-02. Available at https://unstats.un.org/unsd/snaama/Downloads.

**Univ. of California, Davis.** 2022. "Global Price and Income History Group (GPIH)." Accessed on 2022-06-01. Available at https://gpih.ucdavis.edu/GDP.htm.

**Univ. of California, Davis.** 2024. "Government Budget Historical Series." Accessed on 2024-10-21. Available at https://gpih.ucdavis.edu/Government.htm.

**UN Statistics Division.** 2025. "UN Comtrade: Historical Trade Statistics 1900-1960." *https://unstats.un.org/unsd/trade/data/tables.asp*.

**Vamplew, Wray.** 1987. "Australians: Historical Statistics." Vol. 10. Broadway, N.S.W., Australia:Fairfax, Syme & Weldon Associates.

**Villa, Pierre.** 1994. *Une analyse macroéconomique de la France au XXème siècle.* Paris:CNRS Éditions.

**World Bank.** 1999. "World Development Indicators (Archives)." Accessed on 2024-07-21. Available at https://databank.worldbank.org/source/wdi-database-archives.

**World Bank.** 2024. "World Development Indicators." Accessed on 2024-04-21. Available at https://data.worldbank.org/indicator.

**World Bank.** 2026. "Global Economic Monitor." Accessed on 2026-04-19. Available at https://databank.worldbank.org/source/global-economic-monitor-(gem).

**Xu, Yi, Zhihong Shi, Bas Van Leeuwen, Yuping Ni, Zipeng Zhang, and Yan Ma.** 2015. "Chinese National Income, ca. 1661–1933." University Library of Munich MPRA Paper 70873.