The Global Macro Database: A New International Macroeconomic Dataset*

Karsten Müller† Chenzi Xu‡ Mohamed Lehbib § Ziliang Chen¶

January 2025

Abstract

The *Global Macro Database* is an open-source, continuously updated dataset of macroeconomic statistics that unifies and extends existing resources. By harmonizing and integrating data from 32 major contemporary sources—including the IMF, World Bank, and OECD—with historical records from 78 additional datasets, we construct comprehensive annual time series for 46 variables across 243 countries. This database covers global macroeconomic trends from the origins of modern data collection to projected estimates for 2030. Using this extensive database, we study the long-run output losses of financial crises and global temperature shocks, two applications in which historical time series are a crucial input. Our findings show that financial crises are associated with statistically detectable contractions in real GDP for five decades into the future, which are considerably larger than previously estimated. Temperature shocks also predict real GDP contractions up to 30 years ahead, especially in emerging economies.

*Corresponding author: Karsten Müller. Müller acknowledges support from a Singapore MOE PYP grant (WBS A-0003319-01-00 and A-0003319-02-00), Tier 1 grant (A-8001749-00-00), and the NUS Risk Management Institute (A-8002360-00-00). Yaqi Tu provided excellent research assistance. The data, code base, and additional documentation accompanying this paper is available at www.globalmacrodata.com.

†National University of Singapore, Department of Finance and Risk Management Institute; kmueller@nus.edu.sg

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

§National University of Singapore, Department of Finance and Risk Management Institute; lehbib@nus.edu.sg

¶National University of Singapore, Department of Finance and Risk Management Institute; z.chen@nus.edu.sg

1 Introduction

Researchers working with longitudinal macroeconomic data across multiple countries face several significant challenges. First, while international organizations such as the World Bank and International Monetary Fund provide contemporary data, their historical coverage remains limited in both scope and depth. Second, although existing academic databases have made substantial contributions to extend these data, they often suffer from restricted country coverage, a limited selection of variables, or infrequent updates. Finally, harmonizing across the various available sources presents additional methodological challenges that make the cleaning of these data costly for individual researchers, and likely done in an inconsistent manner by different researchers. These constraints frequently lead researchers to rely solely on recent data from individual sources, which can introduce selection bias, limit the robustness of longitudinal analyses, and shrink the scope of research topics.

This paper addresses these methodological challenges by introducing a novel, integrated macroeconomic database. The Global Macro Database synthesizes real-time data from major international organizations with historical series compiled by economic historians. The resulting dataset represents by far the most comprehensive collection of annual macroeconomic statistics available. Our initial release encompasses 46 key macroeconomic variables across 243 countries and territories, with temporal coverage extending from the year 1086 through 2024, including projections through 2030.[1] For critical indicators such as exports and nominal Gross Domestic Product (GDP), we have assembled approximately 20,000 observations, substantially exceeding the coverage of existing databases.

Table 1 demonstrates the comparative advantage of our dataset with its substantially broader coverage relative to existing sources, including the World Bank's Development Indicators and the Macro History Database of Jordà et al. (2017). We have over forty percent more observations than the next most populated source. This expanded coverage stems from our methodological approach of systematically integrating contemporary sources with historical research. Beyond the current release, we will maintain an active data pipeline that updates these time series continuously. Our code infrastructure is further set up to easily allow us to close remaining gaps in data coverage, which we will address on an ongoing basis. Our objective is to establish a comprehensive repository of macroeconomic time series that facilitates cross-country analysis while minimizing the

[1]Some of these variables are constructed based on two or more underlying time series (e.g., real GDP per capita). The number of these individual underlying variables is 35. See Section 2.3.4 for a detailed discussion.

data preparation burden on individual researchers.

The construction of a comprehensive macroeconomic database requires both systematic harmonization of contemporary sources and careful integration of historical records to extend coverage across time and space. Our methodological approach begins with an automated code pipeline that pulls from the most trusted aggregators of data covering multiple countries, including the World Bank's World Development Indicators (WDI) and the International Monetary Fund's World Economic Outlook (WEO). After retrieving raw values and metadata from these providers, we harmonize units and reporting currencies. However, these contemporary sources have important temporal limitations. For example, the WEO reports data from 1980, the WDI from 1960, and the Penn World Tables (PWT) from 1950. It is this limited historical coverage that necessitates integrating historical records.

To extend coverage further into the past, in our second step we systematically incorporate research from economic historians, whose valuable contributions often remain outside major data aggregators. We systematically clean and document statistical breaks in Mitchell (2013), and we collect, clean, and harmonize 38 country-specific datasets compiled by statisticians and historians, so that they can be used in cross-country analyses. In the process, we are able to cross-check and verify the plausibility of estimates produced by international organizations.

Finally, we supplement these sources with several newly digitized datasets, including historical editions of the OECD's Main Economic Indicators, long-run interest rate series from Homer and Sylla (1996), and country-specific time series for Australia (Vamplew, 1987), Iceland (Statistics Iceland, 1997b), Argentina (Nakamura and Zarazaga, 2001), and Portugal (Instituto Nacional de Estatística, 2001).

Our contribution extends beyond digitization and aggregation of time series to address common issues in existing datasets. For instance, the IMF's International Financial Statistics frequently contains incorrectly specified units, erroneously coded missing values, and unaccounted for currency changes. We systematically identify and remedy such issues, maintaining detailed documentation of all adjustments at the level of an individual country-variable-year observation.

To illustrate our methodological approach, consider our treatment of real GDP in local currency. We begin with IMF WEO data, establishing 2019 as a reference year that is unlikely to face major revisions. After careful evaluation of country-specific circumstances, we combine WEO data through 2019 with historical records from other sources. We derive post-2019 values from WEO growth rates, including their nowcasts and forecasts. This approach maintains stability in historical values while incorporating ongoing revisions to recent data as new versions of the WEO are released. For earlier periods, we use ratio-splicing based on the growth rates from each source.

The resulting dataset contains long-run estimates for 46 variables for virtually every economy, including countries such as Czechoslovakia that no longer exist. Since we attempt to integrate all available sources we could identify, it is by definition the most comprehensive possible set of time series on major macroeconomic statistics.

The entire codebase and underlying data are publicly available in a dedicated GitHub repository. We also provide a comprehensive technical documentation exceeding 7,000 pages that details source selection, adjustments, and visualizations for how our data compare with each of the individual underlying sources for each country-year observation. The code is designed to seamlessly integrate new releases from any of the existing sources as well as to incorporate new sources. We consider this current effort as only the beginning of a journey with many improvements over time, and we welcome suggestions for improvements or additional data sources.

The comprehensive nature of our new dataset of macroeconomic statistics enables us to revisit two fundamental questions in macroeconomics that have been constrained by data limitations. The first is an estimation of the long-run consequences of financial crises. A large literature shows that such crises are followed by substantial output losses (e.g., Cerra and Saxena, 2008; Reinhart and Rogoff, 2009; Sufi and Taylor, 2021; Frydman and Xu, 2023). However, these studies have focused on a horizon of usually around five years after a crisis event because, in most samples, there is insufficient data on macroeconomic outcomes to consider longer time windows. When combined with a recent combined chronology of banking crises and bank runs from Jamilov et al. (2024), our data allows us to extend the prediction horizon to many decades.

The principal finding from this exercise is that banking crises are associated with a contraction of GDP that is statistically detectable up to 50 years after the outbreak of a crisis. To the best of our knowledge, this is the first evidence showing just how persistent the output losses of such crises are. Importantly, the longer time horizon we consider leads us to conclude that the GDP losses of banking crises are potentially up to *twice* as large as previously estimated. After crises, the output gap relative to trend grows over time and only peaks after around 30 years at 15–20 percent. These magnitudes are considerably larger than those estimated in existing work, which usually finds average GDP losses on the order of 10 percent. The difference can be explained by our longer forecast horizons combined with more data coverage.

We find similar evidence when considering the GDP losses of bank runs, complementing the evidence in Jamilov et al. (2024). While the average output losses are somewhat smaller than those of banking crises, they also reach around 15 percent after 30 years, and it is only after 50 years that GDP is back to its previous trend. Taken together, these findings provide new evidence suggesting that the costs of these relatively rare crisis events may have been understated due to the lack of long-run data.

Our second application investigates the long-run effects of temperature shocks. Recent work by Bilal and Känzig (2024) suggests that shocks to global (rather than local) temperature are key for subsequent output. In order to differentiate between the impact of country-specific (local) and global temperature shocks, Bilal and Känzig (2024) run regressions in two entire different datasets of real GDP: one that covers many countries but for fewer years, and one that extends farther back in time but is significantly more limited in its coverage of countries. Up to this point, the "ideal" dataset that encompasses a large sample of countries and extends beyond the 20th century has not existed.

We extend their work with the Global Macro Database by considering a considerably longer time horizon, both in terms of sample coverage and prediction horizon. Put differently, our dataset resolves the need to consider several datasets to test the same hypothesis, because we integrate all the underlying sources. This more comprehensive database is particularly relevant given that climate change is a relatively slow-moving phenomenon, which makes an extension of the time series dimension particularly valuable.

We find that global temperature shocks have a significantly larger impact on output than local shocks, and these effects are long-lasting, with statistically significant consequences observed for up to 30 years. Our results reinforce those of Bilal and Känzig (2024), showing that their findings are not only detectable in a larger sample but also even more persistent than previously found.

Our paper contributes to the existing literature in several dimensions. First, we extend existing academic efforts to compile and harmonize macroeconomic statistics for several countries. In spirit, our work is most closely related to Jordà et al. (2017), who construct a close-to-balanced panel dataset on macroeconomic and financial variables covering 18 advanced economies starting in 1870. Other important existing contributions include the *International Historical Statistics* compiled by Mitchell (2013) and the data on macroeconomic disasters constructed by Barro and Ursúa (2008, 2012). We cover a far broader set of countries from a much earlier starting point. Second, we demonstrate the applicability of our dataset in showing the long-run persistence of various shocks on global macroeconomic outcomes. These estimations are only possible with a sufficiently comprehensive and integrated dataset.

The paper proceeds as follows. Section 2 introduces the new dataset, including a description of the variables we include, the source we draw on, a comparison with existing efforts. Section 3 outlines some long-run historical facts about global interest rates, trade patterns, currency returns, and the U.K. economy since medieval times. Section 4 provides evidence on the long-run costs of financial crises and temperature shocks. Section 5 concludes.

A comprehensive documentation page accompanies this paper and serves as a detailed guide to the Global Macro Database. It includes variable definitions, data coverage, methodological details, and instructions for effective use of the dataset.

2 A New Historical International Macroeconomic Dataset

In this section, we overview the Global Macro Database. We begin by detailing our data coverage across variables and countries, including both contemporary indicators and specialized historical series. We then describe our methodological framework for integrating and harmonizing data from diverse sources, with particular attention to addressing common measurement issues and ensuring consistent updates. We provide comprehensive documentation of our approach and all sources, including a description of adjustments to the raw data and time series plots showing how these map into our final estimates, in a series of dedicated technical appendix documents (7,000+ pages) available through www.globalmacrodata.com/documentation. Finally, we demonstrate the dataset's expansive coverage relative to existing sources and how researchers can access and utilize these data.

2.1 Variables

Table 2 provides an overview of the 46 macroeconomic variables most widely used in applied research, along with their coverage in our dataset. These indicators include key measures of economic performance and activity: Gross Domestic Product (GDP, nominal and real), inflation (both a consumer price index (CPI) and a GDP deflator), population, measures of investment and household consumption, exports and imports, current account, the US dollar exchange rate, real effective exchange rate, information on government finances (total government revenues, tax revenues, government deficits, government debt), interest rates (short-term, long-term, central bank policy rate), measures of the money supply (M0, M1, M2, M3, M4), unemployment, and house prices. Additionally, our dataset includes other important variables, such as various measures of financial crises (banking, currency, and sovereign debt).

2.2 Sources

We currently take time series from a total of 110 sources. A full list is in Table OA1. Table 3 shows that these sources can be divided along two dimensions. The first dimension is the number of countries covered by a source. We use data from 72 "aggregators" such as the World Bank, IMF, OECD, or PWT, which provide time series for more than one country. However, not all of these aggregators are international organizations. Jordà et al. (2017) or Barro and Ursúa (2008, 2012) 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 Statistical Bureau of Taiwan or the historical time series on the United Kingdom compiled by Thomas et al. (2010).

The second dimension is whether the data is "current", meaning continuously being updated, or "historical", meaning it is not. The data provided by international organizations, with the exception of some historical sources, almost always falls under the bucket of being "current." These sources are continuously updated as part of our database, as described in section 2.3.5. "Historical" data sources, on the other hand, 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, approximately two thirds of our 110 sources are of the country-specific, historical type (Table 3). 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.

2.3 Data Construction Methodology

Many widely-used datasets require considerable data cleaning to be usable. As such, an important part of our contribution is to apply critical cleaning steps to the raw data that researchers otherwise have to do themselves. Many of the issues in the published raw data are severe enough that they could render any economic analysis that naively uses these data potentially erroneous. Moreover, differences in how individual researchers address these issues can exacerbate differences in reproducibility.

2.3.1 Combining Datasets

We start by assembling the largest possible number of datasets on the key macroeconomic indicators outlined in section 2.1. In contrast to existing aggregators, we do so in a completely transparent fashion. The original raw data is stored and retained, and every single adjustment (e.g., unit issues) is saved in a comment on the observation-level.

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.

The starting point of our data construction is thus a code base that cleans each of the 110 underlying sources. As we outline in the next section, this process is far from trivial, given many issues even in the most existing databases.

2.3.2 Splicing Time Series

When combining time series from different sources, we want to avoid jumps due to series breaks. Such breaks can arise because of slight differences in methodology or because of data revisions in more recent data. To adjust such breaks, we follow common practice and use a transparent ratio-splicing approach.

We begin with the most recent sources and work our way backward. Let $y_{i,t}^{new}$ denote the value from a new source for series i at time t, and $y_{i,t}^{old}$ denote the value from a previous source with older data. The break-adjusted series for the older source is computed as:

$$\hat{y}_{i,t}^{old} = y_{i,t}^{old} \times \theta_{i,t_0} \tag{1}$$

where θ_{i,t_0} is the splicing ratio calculated at the overlap year t_0:

$$\theta_{i,t_0} = \frac{y_{i,t_0}^{new}}{y_{i,t_0}^{old}} \tag{2}$$

The splicing is performed at time t_0, which represents the year where both sources overlap. This adjustment ensures that the level of the old source aligns with the more recent data while preserving the growth rates in the old series.

In practice, we can only apply this splicing method when we have overlapping observations from both sources. In rare cases where no overlap exists but we want to avoid gaps in the data, we use an alternative approach: we calculate the median growth rate of each series over three-year windows before and after the potential splicing point, then use these growth rates to derive the appropriate splicing ratio.

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 1 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 is sufficient to construct the full estimated series.[2] While some sources differ in their reported "level" of investment, the trends are highly similar. We start with the IMF's WEO data as of 2019 and then splice the series backwards using the trends in other historical sources.

The Technical Appendix contains exactly this type of exhibit for every single country and variable, resulting in thousands of pages of documentation. This allows for transparent documentation and quality control for all countries and variables, covering everything from investment in France to Bolivian population to the German current account.

2.3.3 Fixing Common Issues In Widely-Used Datasets

The all-encompassing approach that we take, where we systematically integrate and compare different sources, makes it more straightforward for us to spot errors in our sources. On the Global Macro Database website, we maintain a list of mistakes that we have identified in the datasets of major organizations. We list the name of the dataset, the country and year in question, and the type of error that we have identified. We have also reached out to the publishers of the data via email to alert them of these mistakes. Our hope is that the source organizations will fix them over time.

Wrong or inconsistent units It is surprisingly common that even the datasets of major organizations such as the IMF contain mistakes in how units are reported. For example, at the time of writing, the IMF IFS metadata sometimes incorrectly states that a variable is reported in millions, when it is really reported in billions, which can be easily verified by comparing the numbers to other data publishers, including the national central banks. Other datasets report the data for the same variable in different units depending on the country. The IMF's Government Finance Statistics, for example, report some time series

[2] In this case, the remaining unused sources provide redundant data.

for the same variable in millions and others in billions. Such differences in units are not always mentioned in the documentation, and could lead to potentially erroneous conclusions.

Incorrect or missing values Some values reported in the raw data are simply incorrect or incorrectly recorded. The IMF's International Financial Statistics and World Economic Outlook frequently report missing values as 0 instead of missing. Other values are likely typographical errors that occurred when the data were originally digitized. We fix these errors when the source of the error is obvious. If the data appear clearly erroneous but we cannot identify how to apply a fix, we set them to missing.

Inconsistent currencies We report all time series in the national currency currently used in a country, which sometimes requires converting units. For example, historical data sources often report data in the currency used at the time. In these cases, we convert the historical currency to the current one using the irrevocable exchange rate imposed during the currency reform (e.g., the introduction of the Euro).

The most difficult issues are undocumented failures to adjust for currency changes within a given time series. For example, in the IMF's International Financial Statistics on Brazil, the data compilers correctly adjusted the time series for the country's many currency changes. However, they do not take into account the 'Plano Real' stabilization program in 1993, which leads to large jumps in the unadjusted time series, because the currency units were not correctly converted.

Missing data across versions The time series published in the statistical portals of the IMF, World Bank, and OECD are sometimes incomplete representations of the data that these organizations have available. For example, the current version of the World Bank's WDI does not comprehensively contain all of the agency's data, which had been published at various points in the past. We thus download and process all archival versions of these data to fill in these gaps. Horn et al. (2024) show that for the case of public debt statistics, changes across database vintages can be substantial.

We have identified a similar issue in the IMF's IFS where some time series values are simply not reported in the version available on the IMF website, both in the "bulk file" and the interactive statistical portal. We fill these gaps by obtaining additional versions of the IFS from the UK Data Service, where they are consistently reported.

A broader issue is that some 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 in PDF format through the Internet Archive and digitized them to fill gaps in the data available online. We are in the process of digitizing the archives of several other major publications, including the IFS, but these are not yet part of the initial release.

Changing base years Some of the time series reported in Mitchell (2013), such as real GDP, change the base years relative to which they are calculated for the same country. Most of the time, no overlapping data are reported. As a result, we do not know how to calculate real GDP growth for the year in which the base year is changed, because we do not have a separate deflator series without such a change in base year.

We remedy this issue by treating these within-series changes in base years like any other statistical break. In cases without overlapping data, we adjust these breaks by imputing the growth rate at the change in base years with the median growth rate of the series in the three years before and after. Note that we implement this procedure for completeness, but that it makes virtually no change to any of our estimates because Mitchell (2013) is almost universally ranked as the source with the lowest priority, and as such is only used when no other data is available.

2.3.4 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, 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 in addressing these questions. We always provide derived measures such as real GDP per capita or government debt to GDP that are internally consistent. However, in the few cases where these ratios differ from what is reported in existing sources, and we have reason to believe that these sources have more correct estimates, we additionally provide time series we believe to be more credible (but that are not internally consistent, for example when real GDP divided by population does not be equal the value reported as real GDP per capita).

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.

2.3.5 Data Updates and Revisions

A critical contribution of our dataset is that it provides automated, continuous updates on an ongoing basis. This is far from trivial: existing academic work such as the Macro History Database (Jordà et al., 2017) or the Penn World Tables (Feenstra et al., 2015) are only updated infrequently, which considerably decreases the usefulness of these important contributions for many users.

We can provide real-time updates by construction. The most recent data we draw on almost universally comes from one of the 24 aggregators that are kept current.[3]

One critical issue arising from the frequency of these updates is data revisions. If we assume this year's real GDP estimates from the IMF are true and splice together historical values back to 1850 from many different sources to match them, a data revision would also change each data point in the 19th century.

To mitigate the impact of data revisions, we implement a "fixing" procedure with a five-year lag. In this initial version, we use 2019 as our reference year. This means that we treat the 2019 values as definitive and adjust all historical time series to align with these reference values. Only data from 2020 onwards are subject to regular revisions as new information becomes available. This approach ensures that historical data (pre-2019) remain stable while allowing recent data to be updated with the latest information.

[3] Our conjecture is that while most researchers would like to have as much historical data as possible, using the shorter periods available from the IMF's WEO or World Bank's WDI is sufficient for most research purposes.

2.3.6 Technical Documentation

The Technical Appendix provides full documentation for the construction of each individual time series and details additional conceptual issues with regard to variable construction.

The first part of this documentation is that all code is publicly available on a dedicated GitHub repository, which we continuously update. We explicitly welcome feedback, including through new issues or pull requests on GitHub, on potential errors and omissions.

The second part is a description of which exact source was used for each value in our estimated series and which adjustments were made to the raw data. These reports contain a table with the sources used to construct a time series and plot both the raw data and our estimates, exactly as in Figure 1.

The entire technical documentation has more than 7,000 pages. Given the massive volume of this information, we offer two additional versions apart from the full documentation: a *country-by-country* version that includes all information on all variables for one specific country, and a *variable-by-variable* version that includes all information on one specific variable for all countries.

As an example, consider the case of Sweden, as highlighted in Figure 2a. If one were to consult only data from international organizations such as the IMF, World Bank, or OECD, one would get data from 1950 onwards, and only for a select group of time series. Our data start in the early 19$^{\text{th}}$ century for many variables, and includes not only historical values but also current estimates and forecasts.

Importantly, the coverage is not limited to advanced economies, for which important contributions such as Jordà et al. (2017) have already collated a wealth of information. Figure 2b underscores this for the case of Chile. Similar to Sweden, we have very long-run time series on key indicators such as real and nominal GDP, interest rates, and government debt starting in the early 1800s. As we describe in more detail in the next section, this represents a major departure from existing datasets.

2.4 Comparison With Existing Sources

By construction, the coverage of our dataset is larger than that of all existing publicly available sources. However, it is worth highlighting just how much more data has previously been collected but has hitherto not been combined into usable time series. Table 4 shows both the number of country-year observations with non-missing data for each variable in our dataset as well as the fraction covered by other datasets. Among other major providers, even the most comprehensive only cover a small fraction of what we incorporate. This issue is pervasive even for relatively basic indicators such as nominal GDP. In addition, most other datasets do not have the variables that we include. The one exception to this general pattern is time series data on population, which are relatively easily available from sources such as Inklaar et al. (2018) and Gapminder (2024).

The comprehensiveness of our data is not restricted to publicly available data sources. When comparing our values with the commercial data provider Global Financial Data (GFD), our coverage still looks favorable. We not only supersede the coverage of these data providers, but also provide a much more comprehensive and transparent documentation.

Figure 3 highlights this point once more by comparing the number of country-year observations for each variable in our database with the next most comprehensive source. Our dataset has a considerably larger number of observations for every variable, and in many cases close to *twice* the observation count relative to the next-best source.

Figure 4 plots the coverage of our dataset by variable for different periods of time. It shows that the improved coverage of our data comes from the addition of historical time series that were in many cases painstakingly constructed by economic historians as well as from recent decades. Figure 5 provides additional evidence on the coverage of the dataset. For each variable, we show the fraction of world GDP covered by the variables in our dataset. These figures highlight the all-encompassing nature of our data efforts. For the majority of variables, by 1900 we cover countries accounting for approximately 80 percent of world GDP. [4]

2.5 Accessing the Data

There are three ways of accessing the data: (1) on www.globalmacrodata.com/data, (2) using static links, or (3) using our dedicated Stata command `GMD`. The full data and documentation are available on www.globalmacrodata.com/documentation. Most users will be interested in our ready-to-use files, which are continuously updated as new data arrive.

These data can also be accessed directly using a static link. When using Stata, for example, one can use the following code:

```
use ‘‘https://www.globalmacrodata.com/GMD’’, clear
```

[4] Appendix section OA.1 contains the plots that show the number of sources typically available for each variable. We provide the total number of sources per variable over three time periods: 1900, 1950, 1980, and 2008. Indicators such as imports and exports often only come from one or two sources, while it is not uncommon that nominal GDP appears in five or more sources.

or, alternatively,

```
   import delimited using ‘‘https://www.globalmacrodata.com/GMD.csv’’, varnames(1)
clear
```

An even more convenient and flexible method is to use our user-written command `GMD`. To get started and download all time series on real GDP in national currency, for example, you can simply type:

```
net install GMD, from(https://www.globalmacrodata.com/package)
GMD rGDP
```

It is also possible to filter for a specific country using its ISO3 code:

```
GMD rGDP, country(FRA)
```

`help GMD` allows users to view the details of the syntax.

For reproduction purposes, you may want to include the version of the data you are using. The ready-to-use files already come with meta data describing the exact version you are downloading and when it was last updated. For ease of reproducibility, researchers will always be able to download a specific version (rather than the most recent one), both when using static links or the `GMD` command. We will maintain links to all previous releases on our GitHub repository; the `GMD` command comes with a `version()` option.

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.

3 Stylized Facts

To showcase the breadth and coverage of our dataset, we describe a number of long-run historical facts about global interest rates, trade patterns, the volatility of currency returns, and the U.K. economy since medieval times, which would be impossible to document using only existing data sources.

3.1 A Long View on Global Interest Rates

How have interest rates developed in the very long-run? We build on the work of Schmelzing (2019) and add estimates of long-term interest rates for six additional countries: Belgium, Switzerland, Sweden, Norway, Denmark, and Canada.

Figure 6 plots the resulting evolution of interest rates from 1875 to 2025. Before World war I, interest rates were broadly stable, and went back to their pre-war average until the 1960s. It was only during the inflationary period of the 1970s that interest rates peaked in an unprecedented manner, although not everywhere. Switzerland, for example, was largely spared the large increases that countries like Denmark or Canada would see.

Since then, as has been widely documented, interest rates have seen a secular decline, with only the most recent data showing a reversal. These decreases have been broad based, and are true even for Switzerland, which had lower initial levels of interest rates.

3.2 Historical Trade Patterns

Figure 7 plots the share of the five largest exporters and the rest of the world in world exports from 1850 to today. Several patterns stand out.

First, the United kingdom made up a staggering 35% of world exports during the mid-19$^{\text{th}}$ century, which has declined to only a minuscule fraction today. France also has seen large decreases, while the share of the United States and Germany has remained relatively stable since 1900.

Japan and China show an entirely different pattern. During its rapid growth period following World War II, Japan saw a large increase in exports and quickly overtook the shares of the United Kingdom, France, and Germany. Following its 1990 banking crisis, however, Japan's importance has decreased. Since the late 1990s and its accession to the WTO, China has seen an unprecedented rise in its export share and quickly overtook essentially all other major economies.

3.3 Historical Exchange Rate Volatility

How does the volatility of exchange rates differ across countries? Our very long-run data on nominal exchange rates against the US dollar provides us with an opportunity to investigate this question.

Figure 8 plots the average standard deviation of log-changes in US dollar exchange rates across countries. We focus on the 20 most and least volatile economies. To aid interpretation, we always include the time period covered by the data. We exclude countries with a standard deviation of zero, which are dollarized countries or those with a constant hard peg to the dollar.

There is considerable variation in the volatility of exchange rates across countries. Small Caribbean islands such as the Bahamas, Bermuda, Netherlands Antilles, and Barbados that are closely integrated with the U.S. economy are markedly overrepresented among low-volatility exchange rates. On the other hand, many Central and Eastern European or Central Asian economies feature among the most volatile currencies, as do countries like Austria(-Hungary) and Germany, for which we have very long time series including periods of major upheaval.

3.4 The UK Economy Since Medieval Times

We start with some stylized facts about the economy of the United Kingdom. The underlying data are collated from Thomas et al. (2010), Jordà et al. (2017), and the most recent edition of the IMF's World Economic Outlook (International Monetary Fund, 2024d). As a result of combining these datasets, we have data for the period 1080–2023.

Figure 9 plots nominal GDP, exports, and government debt to GDP (the first two on a log scale). After a long stretch of stagnation, GDP first showed small signs of increases during the Renaissance period, and increased markedly following the industrial revolution. A similar pattern can be seen in exports, for which several datasets point to broadly similar long-run trends.

An altogether different dynamic can be seen when looking at changes in the ratio of government debt to GDP. After secular increases in debt during the 17th and 18th century, the Battle of Waterloo marked a key turning point. Despite some ups and downs, government debt then declined until World War I, which was followed by a massive increase in indebtedness, similar to World War II. The increase in government debt since the 2000s, while pronounced, pales in comparison with these historical episodes.

4 Financial Crises, Temperature Shocks, and Output Losses

In this section, we showcase the usefulness of the *Global Macro Database* by estimating the aftermath of two types of events that have been found to leave a lasting imprint on the macroeconomy: financial crises and global temperature shocks.

4.1 The Long-Run Costs of Financial Crises

Financial crises have well-documented adverse effects on the real economy. Pioneering work by Cerra and Saxena (2008) and Reinhart and Rogoff (2009) (and many others since) has shown that permanent output losses are the norm rather than the exception after a country's banking sector experiences a crash.

What these analyses have so far missed is to investigate just how long-lasting such effects can be. As a single point of reference, Xu (2022) finds that exposure to the 1866 British banking crisis affected the market share of exporters up to five decades after the initial event. Our long-run data allow us to examine the impact of banking crises on real GDP for similarly long horizons, which is much longer than in the existing literature. We provide new evidence that banking crises and bank runs are also typically followed by a persistent decades-long contraction in output.

Our analysis begins with defining a crisis event in the banking sector. While various studies have proposed different narrative chronologies of banking crises using different criteria, leading to some disagreement between measures, there is substantial overlap in what constitutes a banking crisis. The key indicators, as documented by Reinhart and Rogoff (2009), Jordà et al. (2017), Laeven and Valencia (2013), and more recently Baron et al. (2020a), typically include severe banking sector disruptions manifested through: non-performing loans, large-scale defaults, bank failures, government bailouts and interventions, or widespread bank runs.

Equipped with these data, we estimate local projections (Jordà, 2005), specified as follows:

$$\Delta Y_{i,t+h} = \alpha_i^h + \beta^h Crisis_{i,t} + \varepsilon_{i,t}^h, \tag{3}$$

where $\Delta Y_{i,t+h}$ is the change in log real GDP denominated in common units of US dollars over some forecast horizon h; $Crisis_{i,t}$ an indicator variable equal to 1 if country i experiences the start of a banking crisis (Reinhart and Rogoff, 2009) or bank run (Jamilov et al., 2024) in year t; and $\mathbf{X_{i,t}}$ is a vector of lags of the dependent and independent variables. We double-cluster standard errors by country and year.

While it is unlikely that all banking crises are exogenous to underlying economic conditions, which may themselves impact the future evolution of real GDP, we aim to credibly account for a sufficient amount of autocorrelation in the dependent variable (output) such that we are not merely capturing some form of mean reversion. As such, the sequence of estimates of the coefficient β captures the output losses following crises relative to a counterfactual where they did not occur, independent of their underlying causes.[5]

[5] Because crises are relatively infrequent, we are less worried about applying the same extent of control-

Figure 10 presents the results. We allow for a prediction horizon h of up to 50 years, and similarly, lags of the dependent variable j of varying lengths.[6] Because the estimates stem from a (potentially lengthy) series of individual regressions, and we look far into the future, a question is whether these should all come from the same sample of country-year observations. For simplicity, we report two versions: a "balanced panel" (Figure 10b) that requires the estimates for any forecast horizon h to come from the sample when $h = 50$, and the full "unbalanced panel" (Figure 10a) without this requirement that allows for a different sample for each forecast horizon. Reassuringly, both approaches yield similar conclusions.

Consistent with existing work, we also find that banking crises are costly for the macroeconomy. The magnitudes in Figure 10a show that the output losses following such crises reach 10% of pre-crisis real GDP in the 25 years after the shock, and that in fact the costs continue to intensify in subsequent decades reaching average magnitudes of 15%. They remain statistically detectable throughout the entire 50 year horizon that we examine. The balanced panel estimates (Figure 10b) are even larger in magnitude, with peak output losses of roughly 20-23 percent compared to 15-17 percent in the full sample. In both cases, output displays a prolonged decline that reaches its trough approximately 30-35 years after the crisis onset, without full recovery even 50 years onward. To the best of our knowledge, the extreme persistence of the effects that we find is new to the literature, and suggest that banking crises may be even more detrimental than previously captured.

Figure 11 repeats the same exercise for bank runs. Similar to banking crises, output contracts by around 10% overall and stays depressed for several decades. While Jamilov et al. (2024) makes a crucial distinction between systemic and non-systemic runs, it is clear that the overall impact of bank runs is already significantly negative. Recovery to the previous trend does not emerge until approximately 40 years after the initial shock. In the balanced panel in Figure 11b, the output losses are larger in magnitude and more persistent, as in the case with banking crises.

In sum, these results show that financial crises are costly affairs, and they provide the first evidence that the long-run consequences are considerably larger than previously shown.

ling for lags to our independent variable of interest, and we thus do not add lags in our baseline specification. The results are nearly identical when they are included.

[6] The results are virtually unchanged with lags as long as 50 years.

4.2 Revisiting the Costs of Global Temperature Shocks

Having established the macroeconomic costs of banking crises, we next turn to analyzing the impact of temperature shocks on future growth. Perhaps surprisingly, the economics literature had until recently found relatively limited evidence that temperature shocks play anything close to a catastrophic role in future economic growth. Bilal and Känzig (2024) challenges this notion by pointing out that these studies almost universally relied on variation in *local* rather than global temperatures. Given that climate change is fundamentally a global trend, the relevant estimand of interest should thus be the correlation between a global temperature shock and future local output.

Our objective is to extend the evidence in Bilal and Känzig (2024) by investigating the longer-run aftermath of global temperature shocks. Their study is a prototype of an empirical exercise that would have benefited from our dataset: in order to show the generalizability of their results, they estimate their regressions separately in a "broad but short" panel (taken from the Penn World Tables, covering 173 countries starting in 1960) and a "narrow but long" panel (taken from Jordà et al. (2017), covering 18 countries starting in 1900). As emphasized above, the *Global Macro Database* incorporates both of these sources, apart from many others.

In order to extend the analysis to our broader panel, we reconstruct global and local temperature shocks using the same source as Bilal and Känzig (2024), which is collected from the Berkeley Earth Surface Temperature Database. We use the annual temperature series that combines land and oceanic temperatures, which begins in 1850. The data are at the grid-cell level and were mapped to match the country borders used in our Global Macro Database dataset.

We follow the same procedure as Bilal and Känzig (2024) in calculating "shocks," in which we obtain $\hat{shock}_t$ from the following regression:

$$T_t = \alpha + \beta_1 T_{t-1} + \beta_2 T_{t-2} + shock_t \tag{4}$$

We also follow Bilal and Känzig (2024) in our baseline specification, which includes two lags of real GDP growth per capita and of the temperature shocks.[7]

Figure 12 plots the impulse responses of real GDP to global and local temperature shocks estimated in the panel of countries. These results show even more persistent real GDP losses, with 7 percent lower output at the 10 year horizon. Our long-run dataset allows us to show that far from being a recent phenomenon, the costly nature of global

[7]Since our analysis extends far beyond the most recent decades, we do not include the control variables such as the oil price shock and US Treasury yields. Including them does not change our results.

temperature shocks has been present since the industrial revolution.

In Figure 13, we extend our analysis to a 50 year horizon, and we also separately estimate the impact of global temperature shocks on the set of developed countries that are included in the Jordà et al. (2017) Macrohistory dataset, and those that are not. Decomposing the long-run macroeconomic effects of global temperature shocks allows us to show the cross-sectional heterogeneity in adjustment dynamics. The aggregate response shows a persistent decline in output reaching approximately -15 percent after 25 years. At those horizons, advanced economies (JST countries) and less developed economies (non-JST countries) exhibit almost identical patterns. However, in the very long-run, advanced economies experience recovery with output returning to being statistically indistinguishable from pre-shock levels after 50 years. In contrast, non-JST countries show no signs of recovery, with output remaining depressed by about 12 percent even five decades after the shock. This divergence becomes particularly pronounced around the 35-year horizon, indicating that the recovery in the aggregate response is primarily driven by advanced economies.

All together, these findings indicate that global temperature shocks induce substantial and persistent economic costs across both developed and developing economies. However, by extending the analysis to more countries, we find that advanced economies may be better able to recover from climate-related macroeconomic shocks.

5 Conclusion

The availability of consistent cross-country macroeconomic data represents a fundamental challenge for empirical research. Existing datasets have historically been constrained by limitations in either temporal or geographic coverage, sometimes have major errors, and cannot be easily combined due to differences in variable definitions. The Global Macro Database addresses these constraints by synthesizing data from international organizations—including the World Bank, International Monetary Fund, and Organisation for Economic Co-operation and Development—with contributions from economic historians. This integration is implemented through an open-source framework with comprehensive documentation and regular updates. Researchers will be able to immediately access the most comprehensive set of time series through a Stata command (`GMD`), thereby reducing barriers to cross-country empirical research.

Despite our best efforts, there are limitations to the first version of the database we introduce here. First, it is unavoidable that we have overlooked many data sources, both current and historical. An obvious example is data published by national statistical offices.

We only include such data in a few cases where we otherwise lack historical data, given the disproportionate effort in maintaining automated downloading from many sources. We are currently working on extending the dataset in this direction. A second limitation is with respect to the frequency of the data. For simplicity, we focus entirely on annual information, although the underlying data is often quarterly or even monthly. We are also working on extending the data to be available at the highest possible frequency, too.

Our approach to combining data from such a large number of sources is to follow best practice among researchers (see, e.g., Mitchell, 2013; Jordà et al., 2017), which is to splice data from the most reliable organizations together with estimates by individual researchers, in order to minimize disagreement on comparability across time and countries. We are certain that some of our estimates can be improved and invite, in this regard and otherwise, the input from other researchers through our GitHub repository.

The two applications that we have presented here showcase the value of having access to long-run time series. We show that financial crises and global temperature shocks leave a much longer lasting imprint on the macroeconomy than previously understood, with statistically detectable output losses decades after the initial events. These results reveal previously undocumented long-term patterns in the aftermath of financial crises and temperature shocks, demonstrating the potential for new insights using these data in future research.

References

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

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

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

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

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

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

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

Baffigi, Alberto. "National Accounts, 1861-2011." In Gianni Toniolo, editor, "The Oxford Handbook of the Italian Economy Since Unification," chapter 2. Oxford University Press, Oxford (2013). 157–186. Dataset: La contabilità nazionale dell'Italia, 1861-2011.

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

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

Bank Al-Maghrib. "Séries Statistiques Monétaires." (2001).

Bank for International Settlements. "Bilateral Exchange Rates." (2024a). Accessed on 2024-07-10. Available at https://data.bis.org/topics/XRU.

———. "Central Bank Policy Rates." (2024b). Accessed on 2024-07-10. Available at https://data.bis.org/topics/CBPOL.

———. "Consumer Prices." (2024c). Accessed on 2024-07-10. Available at https://data.bis.org/topics/CPI.

———. "Effective Exchange Rates." (2024d). Accessed on 2024-07-10. Available at https://data.bis.org/topics/EER.

———. "Inflation." (2024e). Accessed on 2024-07-10. Available at https://www.bis.org/topic/inflation.htm.

———. "Residential Property Prices." (2024f). Accessed on 2024-07-10. Available at https://data.bis.org/topics/RPP.

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

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

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

Banque Centrale des Etats de l'Afrique de l'Ouest. "La Base des Données Economiques et Financières." (2024). Accessed on 2024-07-15. Available at https://www.bceao.int/fr/content/la-base-des-donnees-economiques-et-financieres.

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

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

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

———. "Banking Crises Without Panics." *The Quarterly Journal of Economics*, 136(2020b), 51–113.

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

———. "Rare Macroeconomic Disasters." *Annual Review of Economics*, 4(2012), 83–109.

Bilal, Adrien and Diego R Känzig. "The Macroeconomic Impact of Climate Change: Global vs. Local Temperature." Technical report, National Bureau of Economic Research (2024).

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

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

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

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

Central Bank of the Republic of T¸rkiye. "Central Bank of the Republic of T¸rkiye Official Website." (2024). Accessed on 2024-01-16. Available at https://www.tcmb.gov.tr.

Cerra, Valerie and Sweta Chaman Saxena. "Growth Dynamics: the Myth of Economic Recovery." *American Economic Review*, 98(2008), 439–457.

Cha, Myung S., Nak N. Kim, Ki-Joo Park, and Yitaek Park. *Historical Statistics of Korea*. Number 978-981-15-3874-2 in Studies in Economic History. Springer (2022).

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

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

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

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

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

———. "Eurostat Data Explorer." (2024b). Accessed on 2024-04-23. Available at https://ec.europa.eu/eurostat/databrowser/explore/all/all_themes.

Eurostat. *Handbook on Residential Property Prices Indices*. Number 17280 in World Bank Publications - Books. The World Bank Group (2013).

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

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

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

Flandreau, Marc and Frédéric Zumer. *The Making of Global Finance 1880-1913* (2009).

Food and Agriculture Organization of the United Nations. "FAOSTAT: GDP data." https://www.fao.org/faostat/ (2024). Accessed on 2024-12-29. Available at https://www.fao.org/faostat/.

Frydman, Carola and Chenzi Xu. "Banking Crises in Historical Perspective." *Annual Review of Financial Economics*, 15(2023), 265–290.

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

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

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

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

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

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

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

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

Horn, Sebastian, David Mihalyi, Philipp Nickol, and César Sosa-Padilla. "Hidden Debt Revelations." Working Paper 32947, National Bureau of Economic Research (2024).

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

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

Institute of Economic Research, Hitotsubashi University. "Asian Historical Statistics: Basic Data for Asian Economic History." (2020).

Instituto de Pesquisa Econômica Aplicada. "IPEA Data." (2024). Accessed on 2024-10-08. Available at http://www.ipeadata.gov.br.

Instituto Nacional de Estadística. "Spanish Statistical Office." (2024). Accessed on 2024-01-16. Available at https://www.ine.es.

Instituto Nacional de Estatística. *Portuguese Historical Statistics*. INE, Lisboa (2001).

International Institute of Social History. "Clio Infra: Repository of Global Inequality Data." (2024). Accessed on 2024-01-16. Available at https://clio-infra.eu.

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

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

———. *Monetary and Financial Statistics Manual and Compilation Guide*. International Monetary Fund (2016).

———. "Government Finance Statistics." (2024a). Accessed on 2024-08-05. Available at https://data.imf.org/?sk=a0867067-d23c-4ebc-ad23-d3b015045405.

———. "International Financial Statistics." (2024b). Accessed on 2024-08-05. Available at https://www.imf.org/en/Data.

———. "Monetary and Financial Statistics." (2024c). Accessed on 2024-08-05. Available at https://www.imf.org/en/Data.

———. "World Economic Outlook Database." (2024d). Accessed on 2024-08-05. Available at https://www.imf.org/en/Publications/WEO.

Istituto Nazionale di Statistica. "Italian National Institute of Statistics Official Website." (2024). Accessed on 2024-01-16. Available at https://www.istat.it.

Jamilov, Rustam, Tobias König, Karsten Müller, and Farzad Saidi. "Two Centuries of Systemic Bank Runs." ECONtribute Discussion Papers Series 333, University of Bonn and University of Cologne, Germany (2024).

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

Jordà, Òscar. "Estimation and Inference of Impulse Responses by Local Projections." *American Economic Review*, 95(2005), 161–182.

Jordà, Òscar, Moritz Schularick, and Alan M. Taylor. "Macrofinancial History and the New Business Cycle Facts." *NBER Macroeconomics Annual*, 31(2017), 213–263.

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

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

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

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

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

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

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

Mitchell, Brian. *International Historical Statistics*. International Historical Statistics. Palgrave Macmillan London, 1 edition (2013).

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

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

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

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

Organisation for Economic Co-operation and Development. "OECD Main Economic Indicators - Historical Statitics." (1986). Accessed on 2024-10-30. Available at https://archive.org/details/pub_oecd-main-economic-indicators-historical-statistics.

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

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

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

———. "OECD Quarterly National Accounts Data." (2024e). Accessed on 2024-07-10. Available at https://www.oecd-ilibrary.org/economics/data/oecd-national-accounts-statistics/quarterly-national-accounts_data-00017-en.

———. "Revenue Statistics 2024: Health Taxes in OECD Countries." (2024f). Accessed on 2024-07-10. Available at https://www.oecd.org/en/publications/revenue-statistics-2024_c87a3da5-en.html.

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

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

———. "Growth in a Time of Debt." NBER Working Papers 15639, National Bureau of Economic Research, Inc (2010).

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

Schmelzing, Paul. "Eight Centuries of Global Real Interest Rates, R-G, and the 'Suprasecular' Decline, 1311–2018." *RG, and the 'Suprasecular'Decline*, (2019), 1311–2018.

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

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

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

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

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

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

Statistics Iceland. *Hagskinna — Icelandic Historical Statistics from the 17th Century to 1990*. Hagstofa Íslands, Reykjavík (1997a).

———. "Icelandic Historical Statistics." (1997b).

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

Sufi, Amir and Alan M Taylor. "Financial Crises: A Survey." Working Paper 29155, National Bureau of Economic Research (2021).

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

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

United Nations. "System of National Accounts 2008." Technical report, United Nations, New York (2009).

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

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

University of California – Davis. "Government Budget Historical Series." (2024a). Accessed on 2024-10-21. Available at https://gpih.ucdavis.edu/Government.htm.

———. "Nominal GDP Historical Series." (2024b). Accessed on 2024-10-21. Available at https://gpih.ucdavis.edu/GDP.htm.

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

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

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

Xu, Chenzi. "Reshaping Global Trade: The Immediate and Long-Run Effects of Bank Failures." *Quarterly Journal of Economics*, 137(2022), 2107–2161.

Table 1: Comparing the Coverage of Key Macroeconomic Variables

Source	Start Year		Latest		Countries		Variables
	First	Median	Actual	Forecast	Number	Year-Obs.	
GMD	1086	1800	2024	2030	243	55,431	46
GFD	1000	1820	2024	—	236	39,345	37
IFS	1920	1950	2024	—	220	14,179	24
WEO	1960	1970	2024	2029	208	11,086	23
JST	1870	1870	2020	—	18	2,666	29
MAD	1253	1876	2022	—	169	20,590	3
OECD EO	1960	1960	2024	2025	49	2,726	35
PWT	1950	1960	2019	—	183	10,399	3
UN	1970	1970	2020	—	215	10,220	14
WDI	1960	1960	2023	—	222	13,454	36

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 2 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.

Table 2: Variable Overview

Variable	Abbreviation	Unit	From	To	Forecasts	Countries
A. National accounts						
Nominal GDP	nGDP	Millions of LC	1086	2029	5	229
Real GDP	rGDP	Millions of LC	1270	2029	5	194
Real GDP in USD	rGDP_USD	Millions of USD	1791	2024	—	192
Real GDP per capita	rGDP_pc	LCU per capita	1277	2029	5	194
GDP deflator	deflator	Ratio	1270	2029	5	194
Population	pop	Millions	1277	2030	6	239
B. Consumption and investment						
Real final consumption	rcons	Millions of LC	1800	2024	—	213
Final consumption	cons	Millions of LC	1800	2025	1	219
Final consumption in percent of GDP	cons_GDP	%	1800	2025	1	219
Gross capital formation	inv	Millions of LC	1830	2029	5	218
Gross capital formation in percent of GDP	inv_GDP	%	1830	2029	5	218
Gross fixed capital formation	finv	Millions of LC	1800	2025	1	216
Gross fixed capital formation in percent of GDP	finv_GDP	%	1800	2025	1	216
C. External sectors						
Current account	CA	Millions of LC	1772	2029	5	209
Current account in percent of GDP	CA_GDP	%	1772	2029	5	209
Exports	exports	Millions of LC	1280	2029	5	225
Exports in percent of GDP	exports_GDP	%	1280	2029	5	222
Imports	imports	Millions of LC	1560	2029	5	225
Imports in percent of GDP	imports_GDP	%	1560	2029	5	220
Real effective exchange rate	REER	Index, 2010 = 100	1870	2025	1	180
USD exchange rate	USDfx	1 USD in LC	1791	2025	1	233
D. Government finances						
Government debt	govdebt	Millions of LC	1670	2029	5	197
Government debt in percent of GDP	govdebt_GDP	%	1670	2029	5	197
Government deficit	govdef	Millions of LC	1792	2029	5	200
Government deficit in percent of GDP	govdef_GDP	%	1792	2029	5	200
Government expenditure	govexp	Millions of LC	1722	2029	5	203
Government expenditure in percent of GDP	govexp_GDP	%	1650	2029	5	199
Government revenue	govrev	Millions of LC	1722	2029	5	202
Government revenue in percent of GDP	govrev_GDP	%	1650	2029	5	198
Government tax revenue	govtax	Millions of LC	1750	2024	—	197
Government tax revenue in percent of GDP	govtax_GDP	%	1789	2023	—	190
E. Money and interest						
M0	M0	Millions of LC	1619	2024	—	187
M1	M1	Millions of LC	1841	2024	—	185
M2	M2	Millions of LC	1841	2024	—	183
M3	M3	Millions of LC	1819	2024	—	70
M4	M4	Millions of LC	1870	2020	—	4
Central bank policy rate	cbrate	%	1694	2025	1	167
Short-term interest rate	strate	%	1695	2025	1	140
Long-term interest rate	ltrate	%	1310	2024	—	84
F. Prices and labor market						
Consumer price index	CPI	Index, 2010 = 100	1209	2029	5	214
House price index	HPI	Index, 2010 = 100	1819	2024	—	59
Inflation	infl	%	1210	2029	5	215
Unemployment rate	unemp	%	1760	2029	5	220
G. Financial crises						
Banking crisis dummy	BankingCrisis	Dummy	1800	2020	—	163
Sovereign debt crisis dummy	SovDebtCrisis	Dummy	1800	2020	—	160
Currency crisis dummy	CurrencyCrisis	Dummy	1800	2019	—	160

Note: This table presents the variables included in the Global Macro Database (GMD), along with their abbreviations/variable names, units of measurement, temporal coverage, forecast horizons, and country coverage. "LC" refers to local currency units and "USD" to US dollars.

Table 3: Types of Sources in the Global Macro Database

	Current	Historical	**Total**
Aggregators	24	48	**72**
Examples	*WEO*	*JST*	
Country-specific	8	30	**38**
Examples	*FRED*	*Thomas et al. (2010)*	
Total	**32**	**78**	**110**

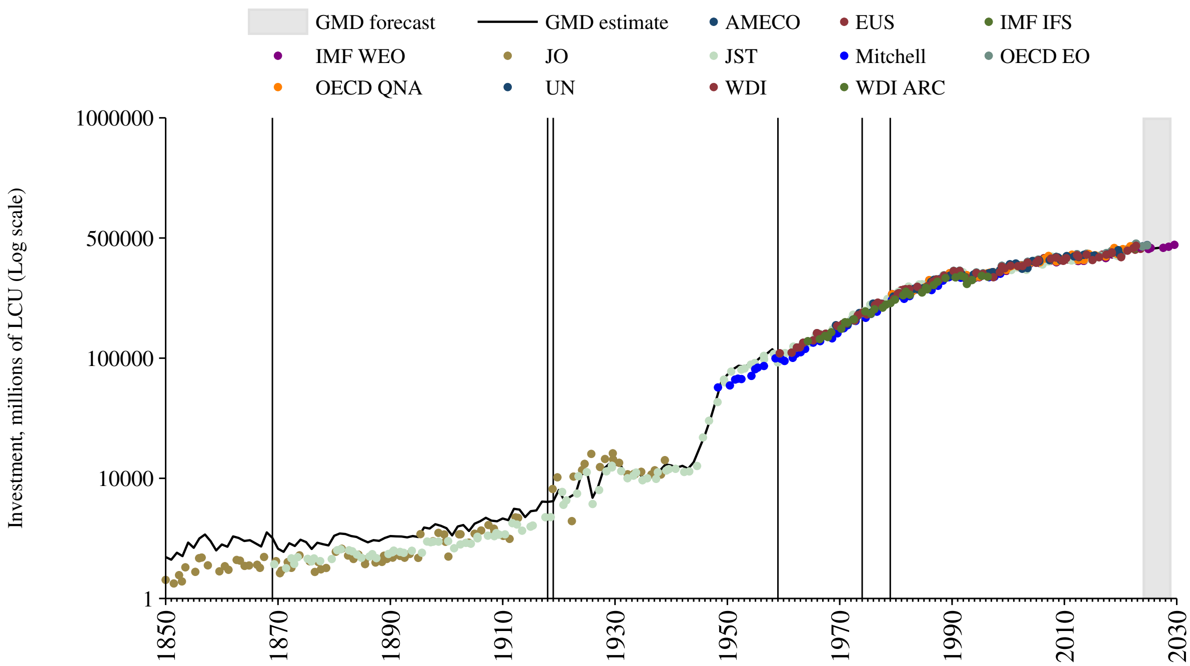
Note: This table plots information on the number of sources used in the Global Macro Database. We differentiate sources along two dimensions: (1) whether they are continuously updated (*current*) or contain only historical data (*historical*), and (2) whether they report information on several countries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

Table 4: 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	8,730	26	—	21	—	—	—	—	73
Short-term interest rate	7,383	24	—	28	—	—	34	—	56
Long-term interest rate	7,750	10	—	23	—	—	34	—	100
Money supply (M0)	14,455	47	—	—	—	—	3	74	93
Money supply (M1)	10,989	61	—	11	—	—	18	53	42
Money supply (M2)	10,509	64	—	—	—	—	14	56	—
Money supply (M3)	3,143	—	—	31	—	—	25	—	—
Money supply (M4)	288	—	—	—	—	—	52	—	—
Real GDP	20,974	25	42	11	53	48	13	31	95
Nominal GDP	19,477	30	45	12	59	52	14	39	84
Consumption	12,752	24	—	18	66	80	—	—	—
Gross capital formation	15,412	27	50	15	54	66	16	36	28
Gross fixed capital formation	14,007	31	—	16	56	72	—	32	28
Current account	13,001	33	66	16	57	—	19	—	21
Exports	25,167	18	39	9	35	35	10	59	73
Imports	24,487	18	40	9	36	36	11	61	75
Real effective exchange rate	11,525	34	—	8	34	—	—	—	39
US dollar exchange rate	23,109	58	—	12	48	—	12	—	98
Government revenue	20,515	20	36	6	22	—	12	60	71
Government tax revenue	14,097	29	—	7	32	—	—	32	1
Government expenditure	19,525	20	37	6	22	—	13	64	72
Government debt	14,689	79	45	10	13	—	17	—	13
Government deficit	12,814	29	57	10	—	—	—	44	100
Unemployment rate	7,645	60	61	26	38	—	25	—	67
Inflation rate	19,671	48	44	3	45	—	13	41	100
Consumer price index	18,657	51	47	3	48	—	14	45	100
House price index	3,517	—	—	—	—	—	58	—	32
Population	51,930	29	17	4	27	32	5	—	53

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 both the IMF MFS and 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.

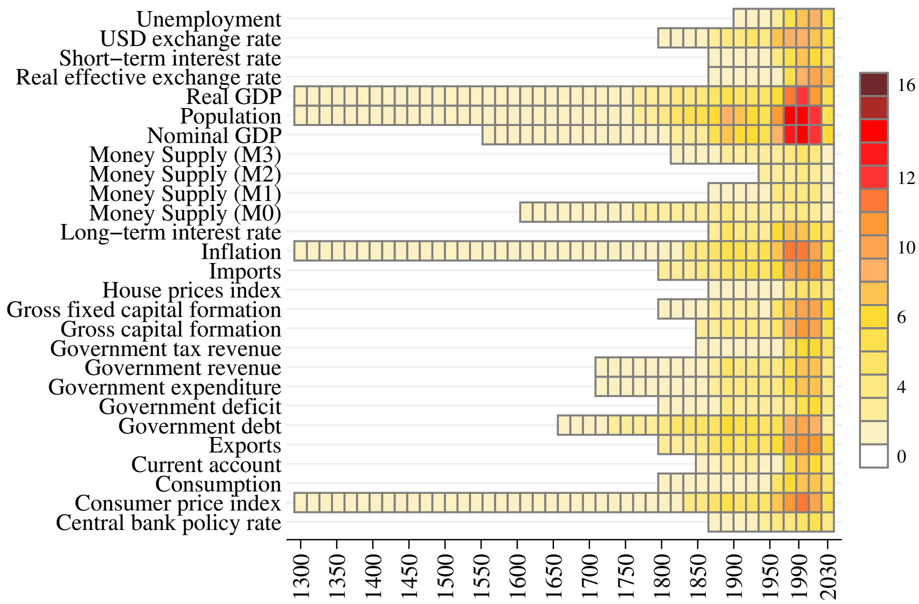
Figure 1: A Splicing Example – Gross Capital Formation in France, 1850-2030



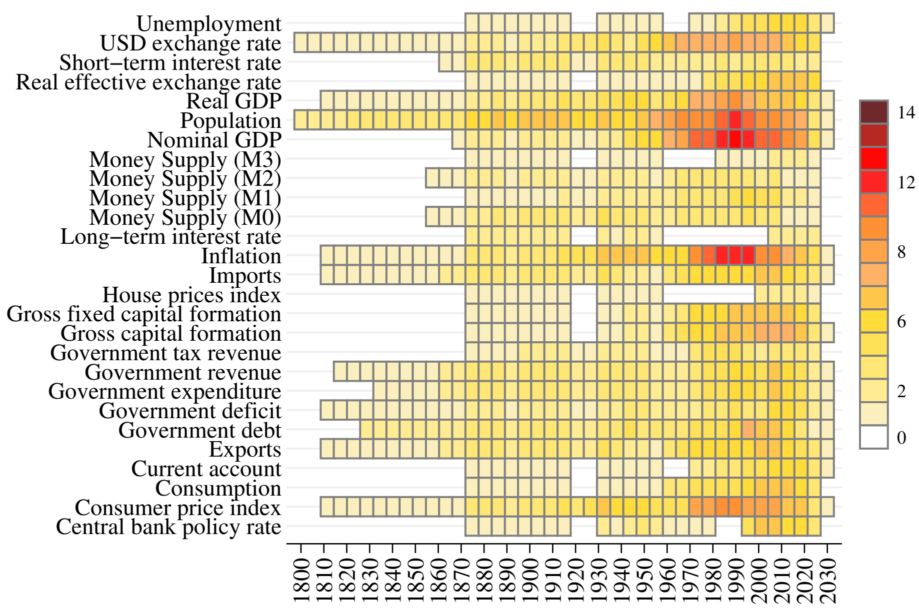
Note: This figure shows gross capital formation (investment) in France from 1850 to 2023 (including forecasts to 2029). We plot the underlying raw data obtained from eleven different data sources (shown as dots), as well as our combined Global Macro Database (GMD) estimates (shown as a line). While the sources show slight differences in levels, they exhibit similar trends. The GMD series starts with the IMF's World Economic Outlook (WEO) data as of 2019 and splices backwards using growth rates from historical sources. Vertical lines indicate a change in sources used to arrive at the GMD estimates.

Figure 2: Examples of Dataset Coverage for Chile and Sweden

(a) Sweden, 1300–2030 (incl. forecasts)

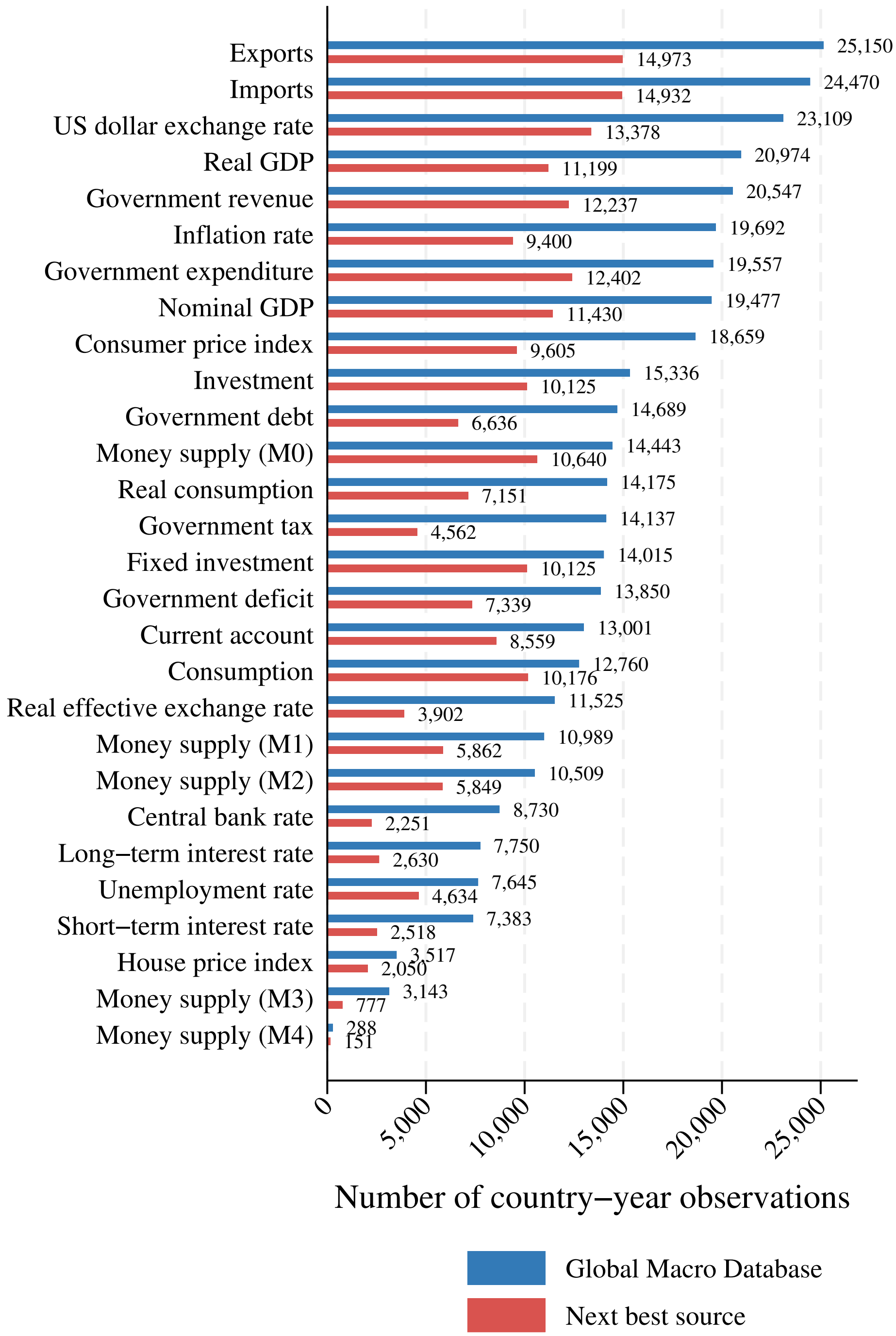


(b) Chile, 1800–2030 (incl. forecasts)



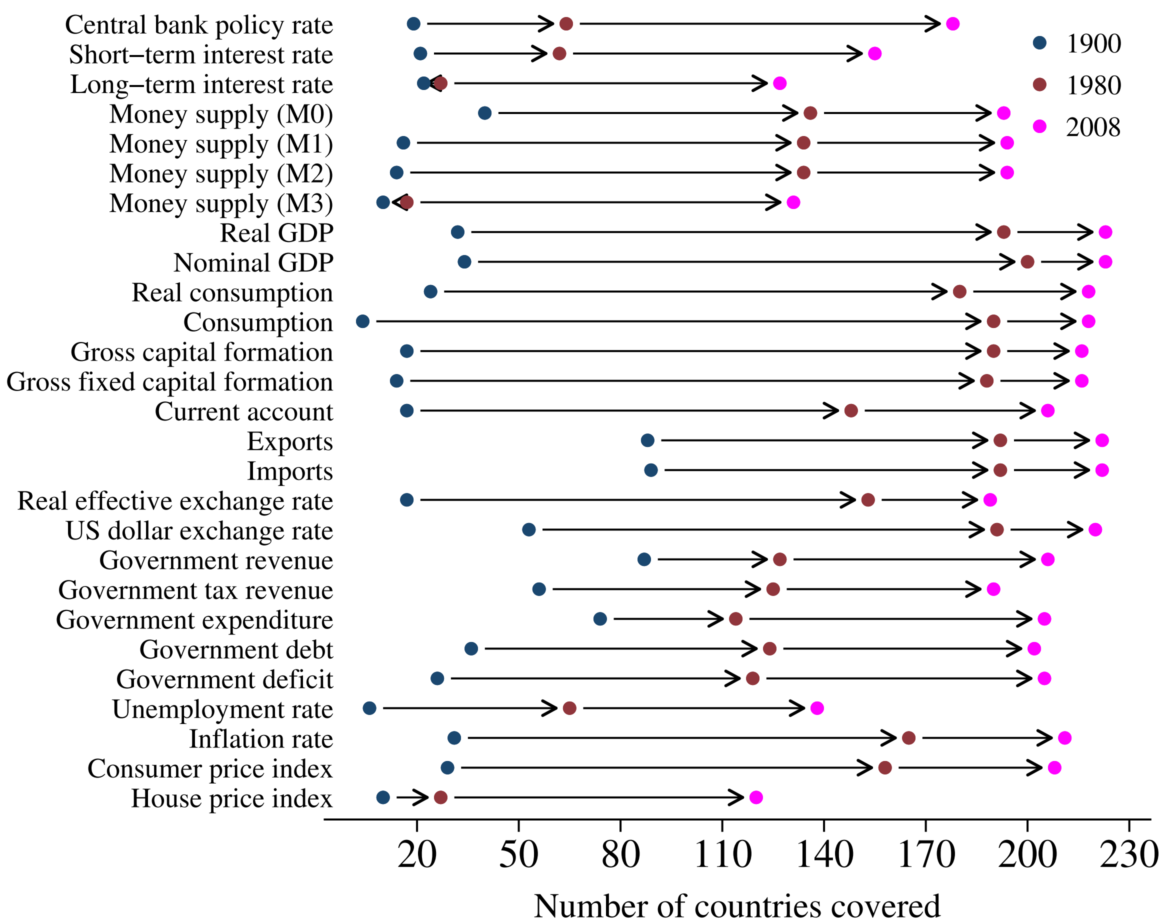
Note: These figures visualize the availability of long-run data on all macroeconomic variables in the Global Macro Database (GMD) and the number of underlying sources for Chile and Sweden. Boxes indicate that a variable is covered in the GMD. Darker, redder colors indicate a larger number of sources for a given time period.

Figure 3: Comparing Dataset Coverage by Variable



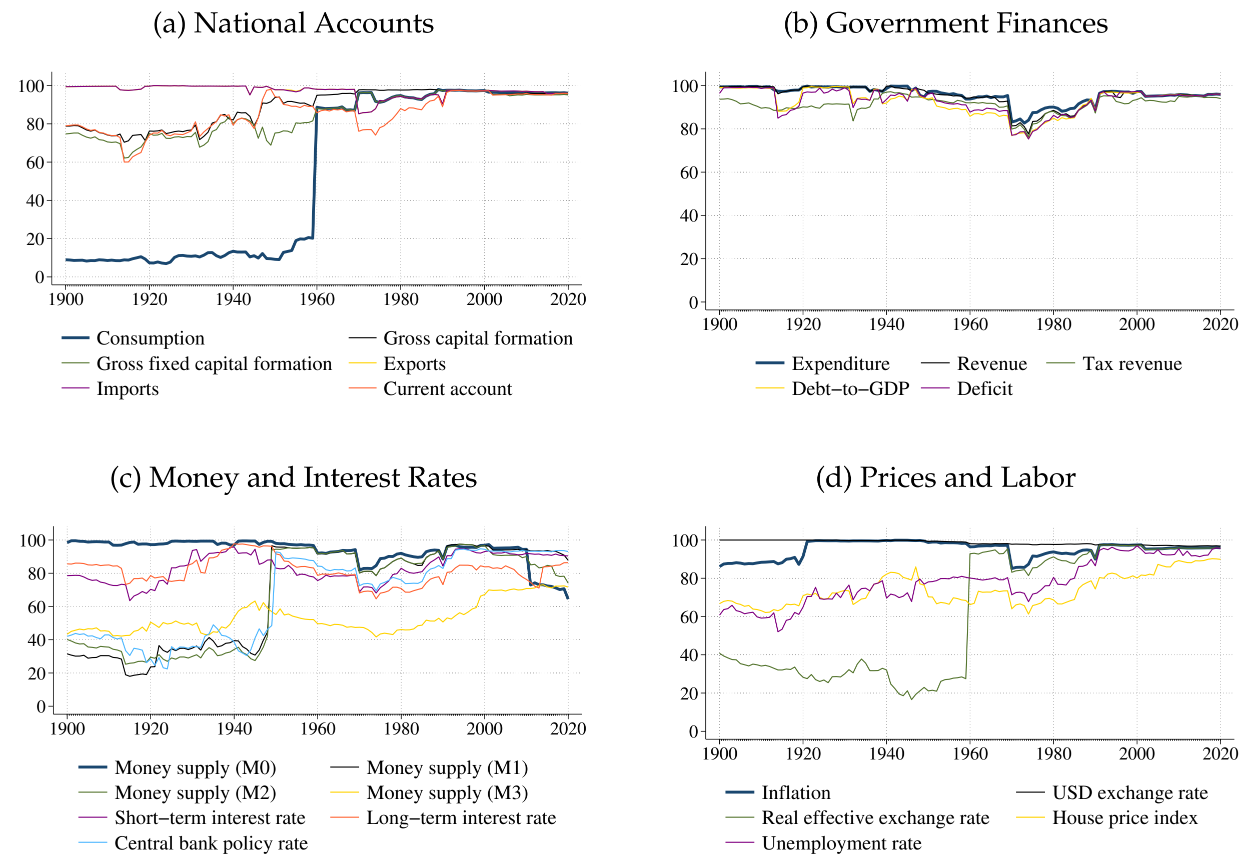
Note: This figure compares, for each variable, the coverage of the Global Macro Database (GMD) with that of the next most comprehensive publicly available source. Blue bars show the number of country-year observations in the Global Macro Database, and the red bars shows the number of country-year observations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

Figure 4: Number of Countries Covered, By Variable



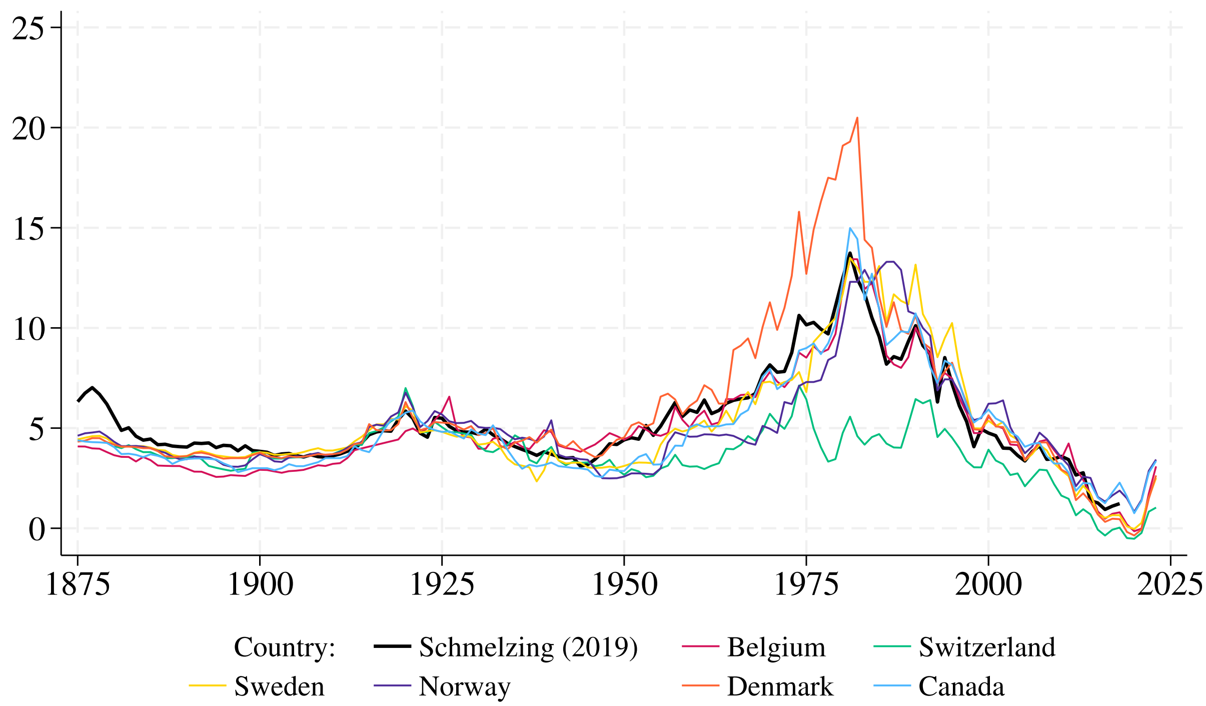
Note: This figure plots the number of countries with available data for all macroeconomic variables contained in the Global Macro Database (GMD) at three points in time: 1900, 1980, and 2008. Each line represents a variable, with dots indicating the coverage at these three dates. The graph reveals substantial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

Figure 5: Share of GDP Covered Over Time, By Variable



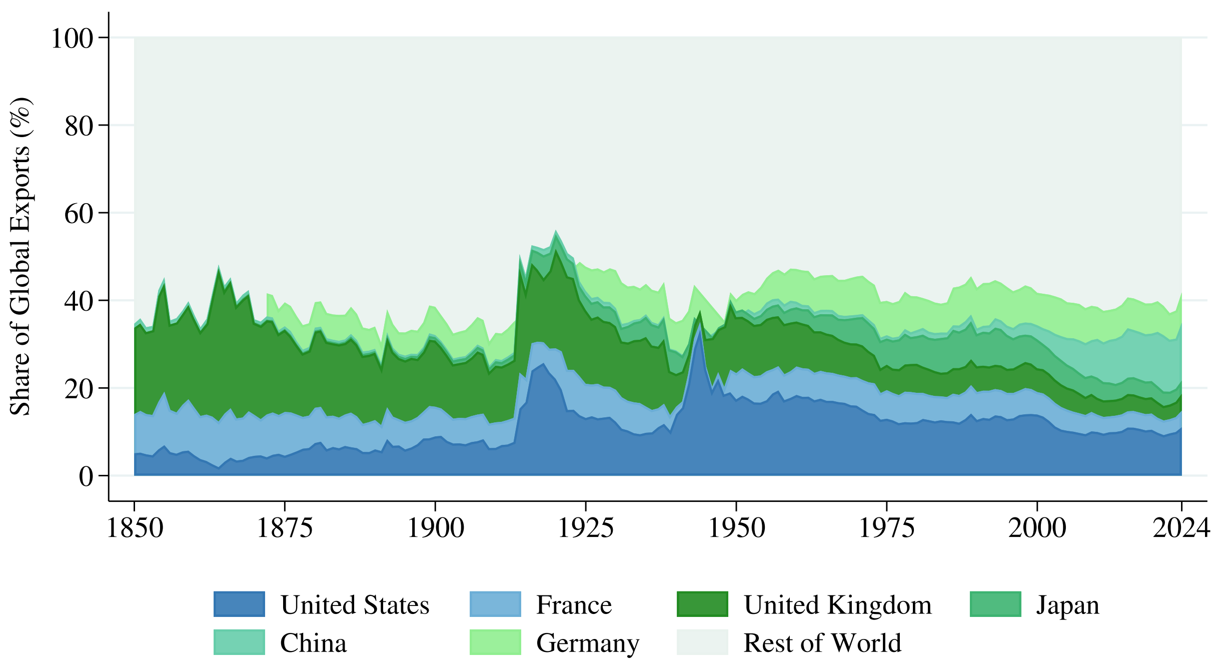
Note: This figure plots the share of GDP covered by each variable in the Global Macro Database (GMD) between 1900 and 2020.

Figure 6: Long-term Interest Rates Over Time



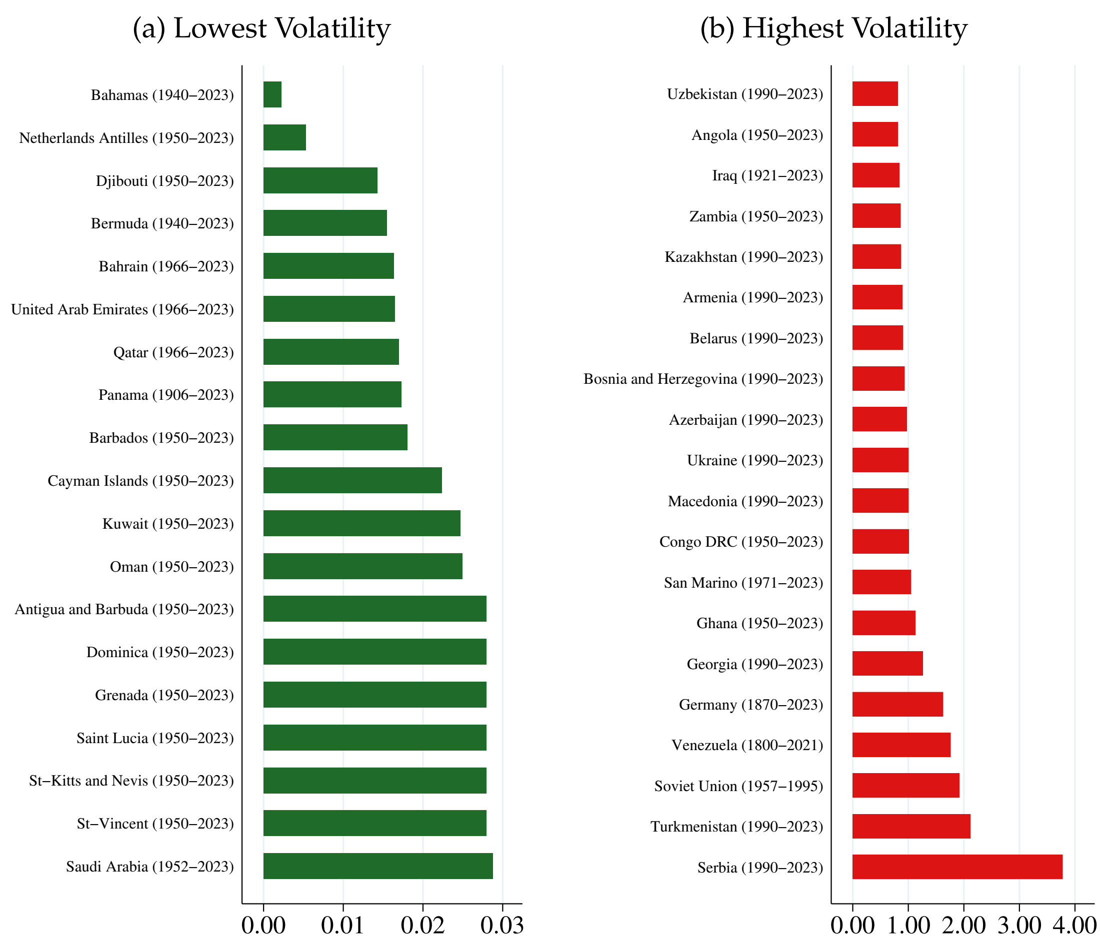
Note: This figure plots the average long-term interest rate for the countries included in Schmelzing (2019) (France, Germany, Italy, Japan, the Netherlands, Spain, the United Kingdom, and the United States), as well as those for 6 additional countries not in his dataset (Belgium, Switzerland, Sweden, Norway, Denmark, and Canada).

Figure 7: Five Largest Exporters Share of Global Exports



Note: This figure shows the share of global exports for some of the largest exporting nations (China, France, Germany, Japan, the United Kingdom, and the United States), plus exports by the Rest of World (ROW). Exports in local currency are converted to USD.

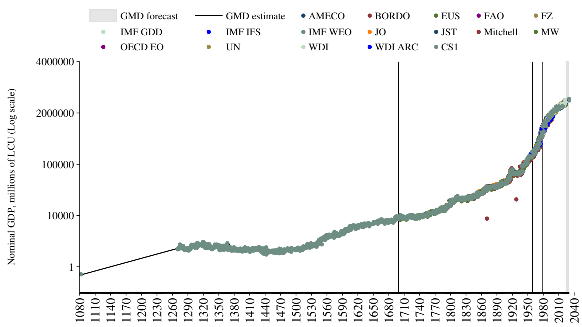
Figure 8: Exchange Rate Volatility against USD



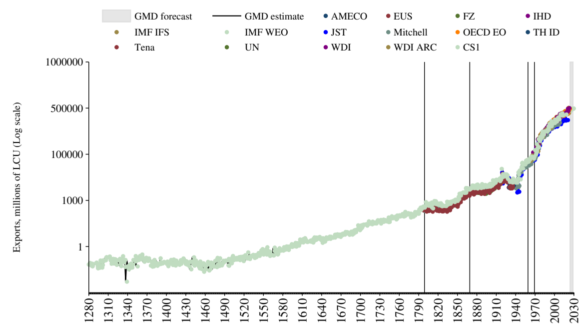
Note: These figures show the volatility of the dollar currency exchange rate over time. Panel A shows the top 20 countries with the most stable exchange rate. Panel B shows the bottom 20 countries that had the most volatile exchange rate over time. We remove countries that have an official or a hard peg to the dollar from our sample.

Figure 9: The UK Economy, 1080–2023

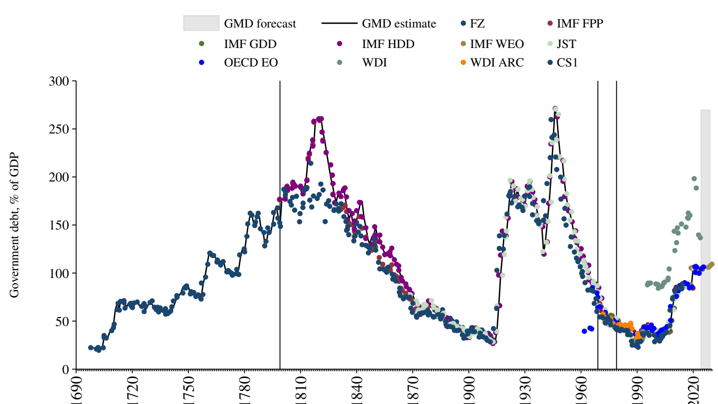
(a) Nominal GDP



(b) Exports



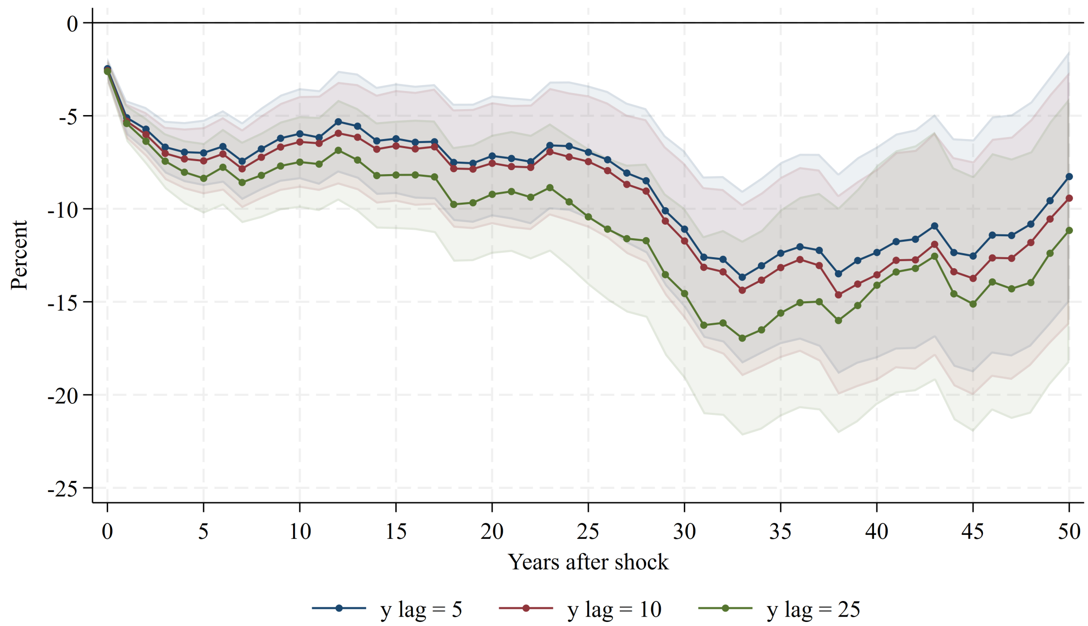
(c) Government Debt to GDP Ratio



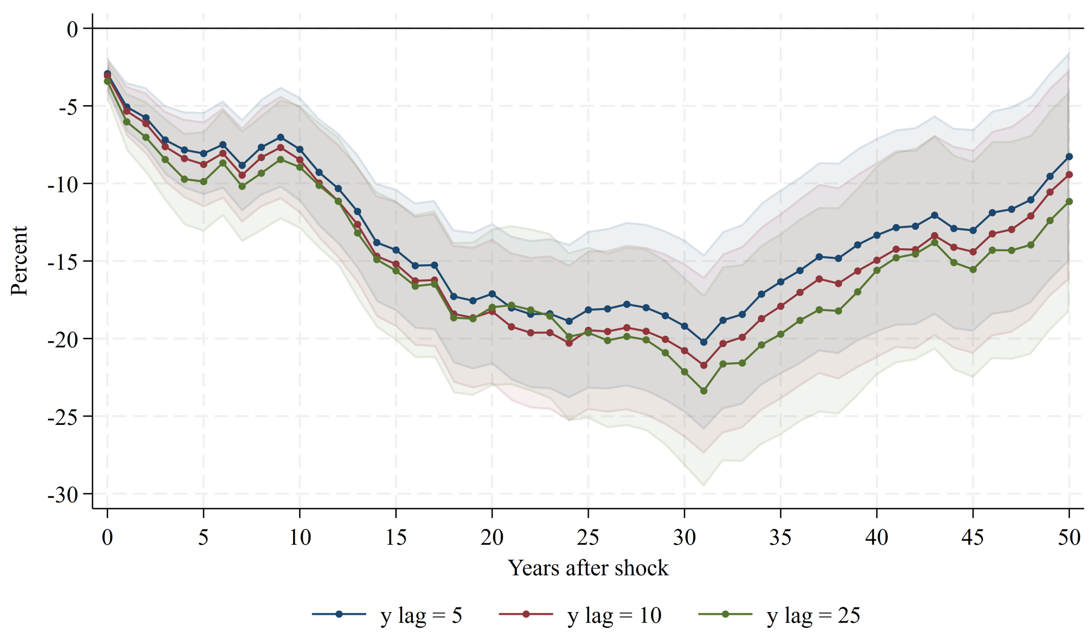
Note: This figure shows three key macroeconomic indicators for the United Kingdom, spanning close to a millenium: nominal GDP, exports, and government debt-to-GDP. The black line shows our final GMD estimates, while colored dots represent data from different sources. The series demonstrates the extensive historical coverage of our dataset, with some series extending back to the medieval period.

Figure 10: The Long-Run Output Losses of Banking Crises

(a) Full sample



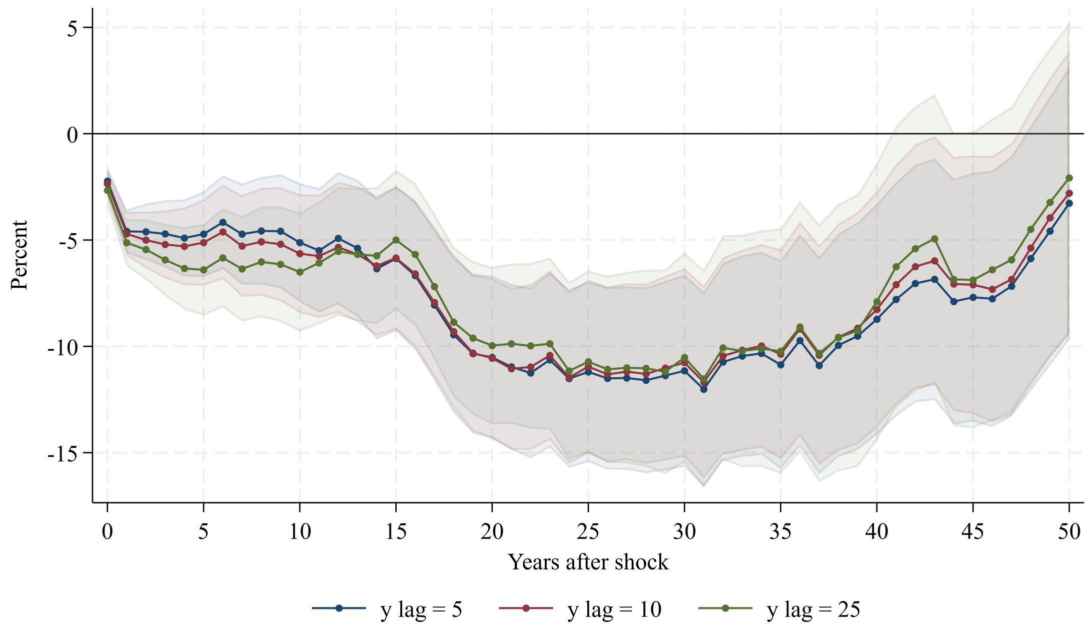
(b) Balanced panel



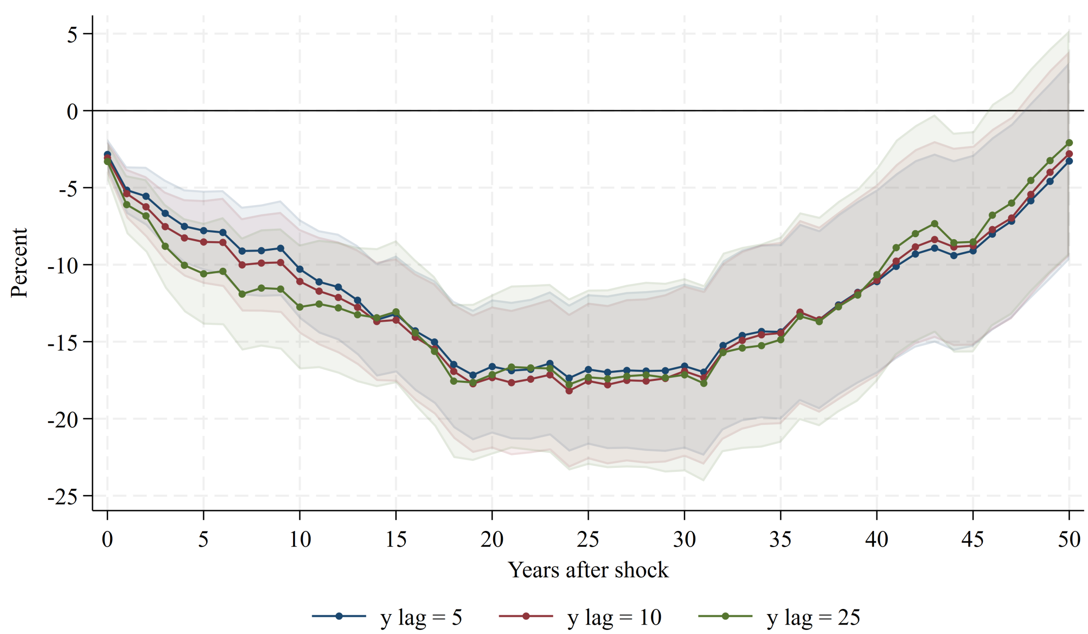
Note: This figure plots the estimated losses in real GDP from banking crises, as defined by Jamilov et al. (2024). We plot the sequence of estimated coefficients β^h from equation 3 for specifications that allow for a different number of lags of the dependent variable. “Full sample” refers to an estimation from the entire sample, and “Balanced panel” to an estimation based on a balanced panel of countries. The shaded areas refer to 95% confidence intervals based on standard errors clustered by country.

Figure 11: The Long-Run Output Losses of Bank Runs

(a) Full sample

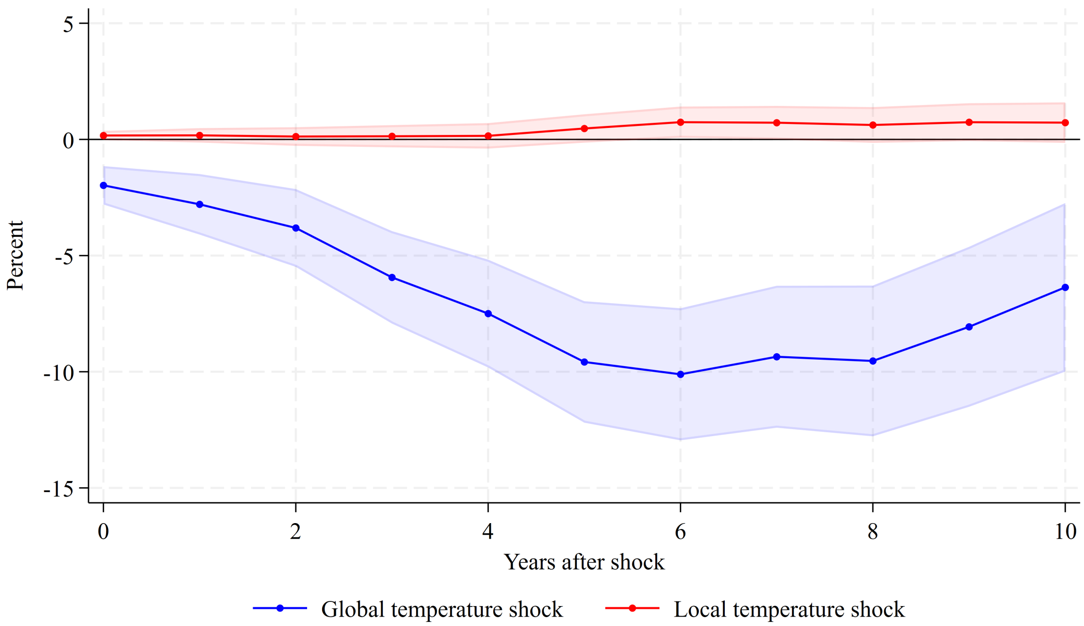


(b) Balanced panel



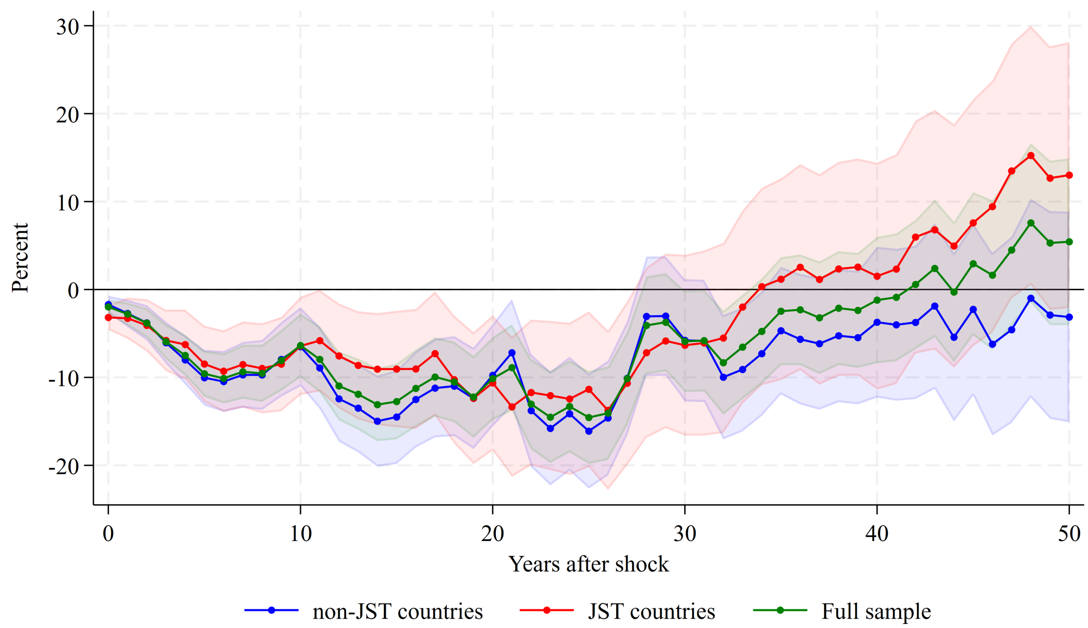
Note: This figure plots the estimated losses in real GDP from bank runs, as defined by Jamilov et al. (2024). We plot the sequence of estimated coefficients β^h from equation 3 for specifications that allow for a different number of lags of the dependent variable. "Full sample" refers to an estimation from the entire sample, and "Balanced panel" to an estimation based on a balanced panel of countries. The shaded areas refer to 95% confidence intervals based on standard errors clustered by country.

Figure 12: Impact of Temperature Shocks on Real GDP, 1850 – 2023



Note: This figure plots the estimated losses in real GDP from global temperature shocks. We plot the sequence of estimated coefficients β^h from equation 3, which closely follows the specification in Bilal and Känzig (2024). The shaded areas refer to 95% confidence intervals based on standard errors clustered by country.

Figure 13: Impact of Global Temperature Shocks on Real GDP



Note: This figure plots the estimated losses in real GDP from global temperature shocks. We plot the sequence of estimated coefficients β^h from equation 3, which closely follows the specification in Bilal and Känzig (2024). "JST countries" are the countries in the Macrohistory Database of Jordà et al. (2017), which Bilal and Känzig (2024) use as their "narrow but long" panel for estimating the output losses of temperature shocks; "non-JST countries" are all other countries in our dataset. The shaded areas refer to 95% confidence intervals based on standard errors clustered by country.

Global Macro Database
Online Appendix

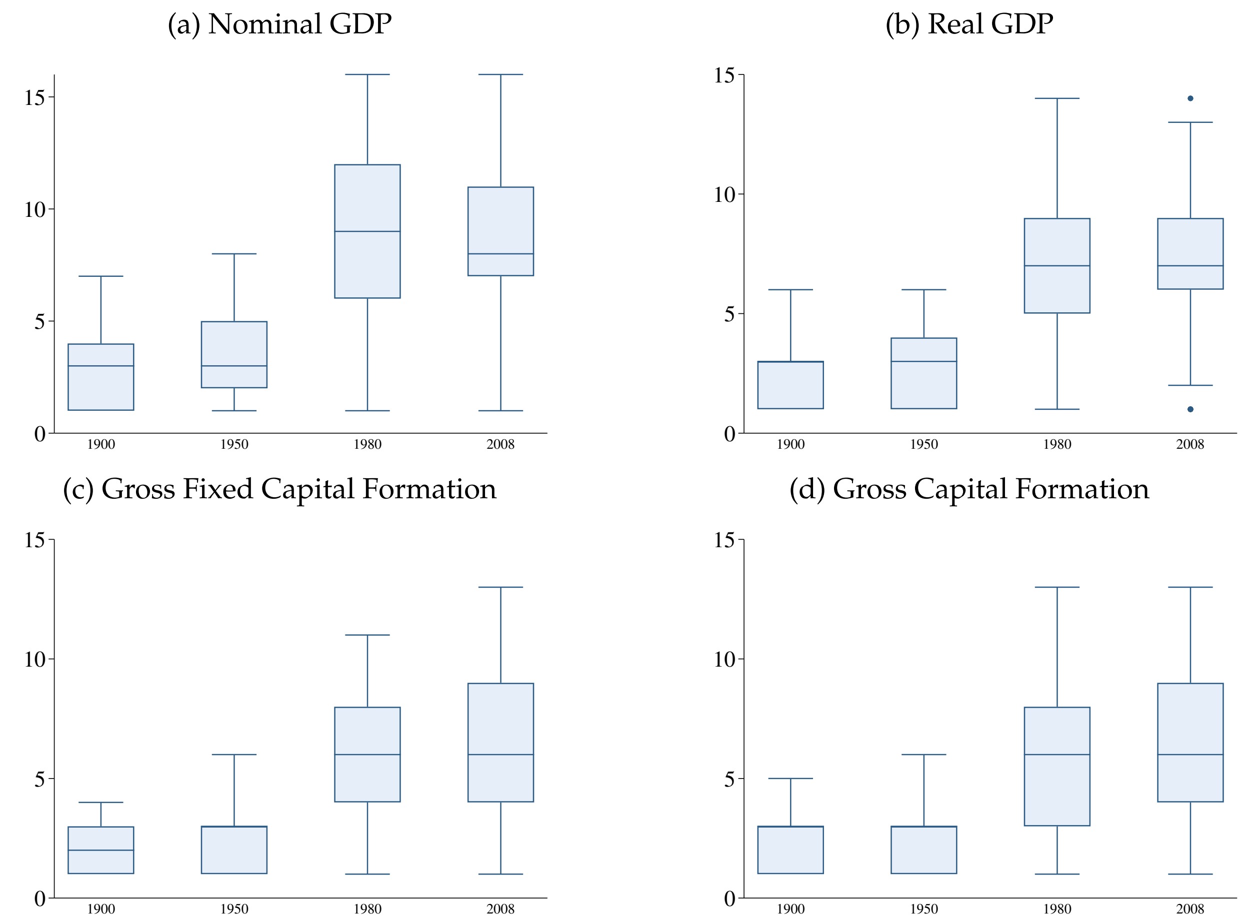
This online appendix contains exhibits that complement the paper *The Global Macro Database: A New International Macroeconomic Dataset* by Karsten Müller, Chenzi Xu, Mohamed Lehbib, and Ziliang Chen.

OA.1 Number of Sources per Variable

This section provides additional information on the number of sources typically available for each variable over time. We present this data in a series of illustrative box plots that show how the number of sources we splice together changes between 1900 and 2008.

The broad pattern we observe is that the number of sources we can draw on, perhaps unsurprisingly, tends to increase over time. For variables such as GDP or gross (fixed) capital formation, exports and imports, exchange rates, or unemployment, we have five or more sources for the typical country from 1980 onwards. The data on house prices or policy rates draws on fewer sources.

Figure OA1: Number of Sources per Variable – I/V



(a) Nominal GDP

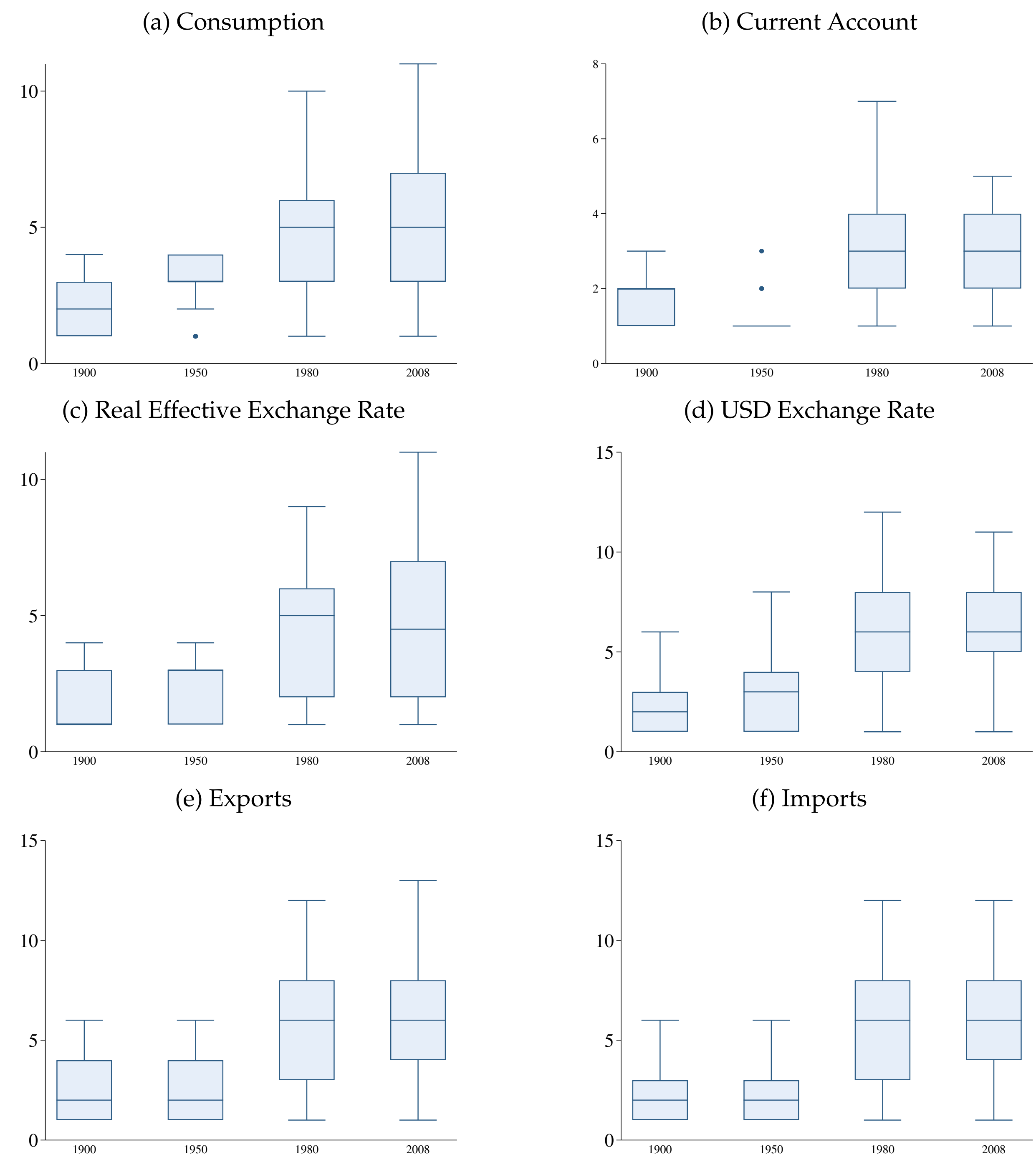
(b) Real GDP

(c) Gross Fixed Capital Formation

(d) Gross Capital Formation

Note: These figures visualize the number of sources we draw on to splice the final values in the Global Macro Database (GMD) for four benchmark years (1900, 1950, 1980, and 2008). These box plots show the median, interquartile range, upper and lower adjacent values, and outliers if any). For any variable x, the upper adjacent value is defined as $x[p75] + \frac{3}{2}(x[p75] - x[p25])$, where $p75$ and $p25$ refers to the 75$^{\text{th}}$ and 25$^{\text{th}}$ percentile, respectively.

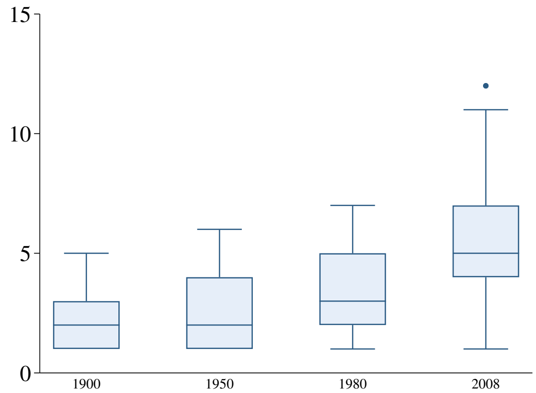
Figure OA2: Number of Sources per Variable – II/V



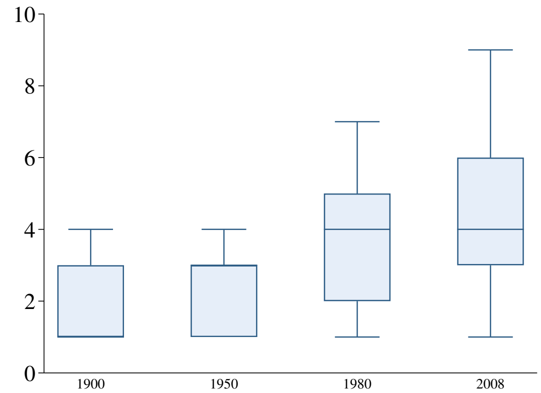
Note: These figures visualize the number of sources we draw on to splice the final values in the Global Macro Database (GMD) for four benchmark years (1900, 1950, 1980, and 2008). These box plots show the median, interquartile range, upper and lower adjacent value, and outliers if any). For any variable x, the upper adjacent value is defined as $x[p75] + \frac{3}{2}(x[p75] - x[p25])$, where $p75$ and $p25$ refers to the 75$^{\text{th}}$ and 25$^{\text{th}}$ percentile, respectively.

Figure OA3: Number of Sources per Variable – III/V

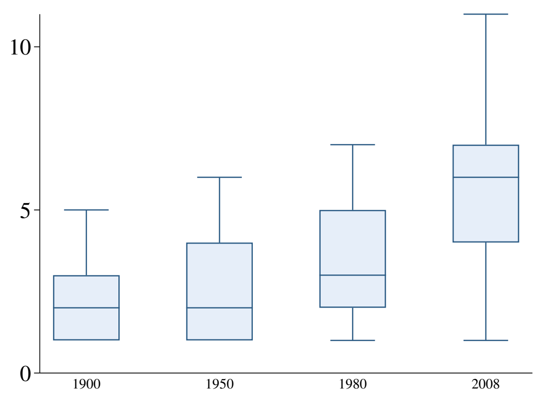
(a) Government Revenue



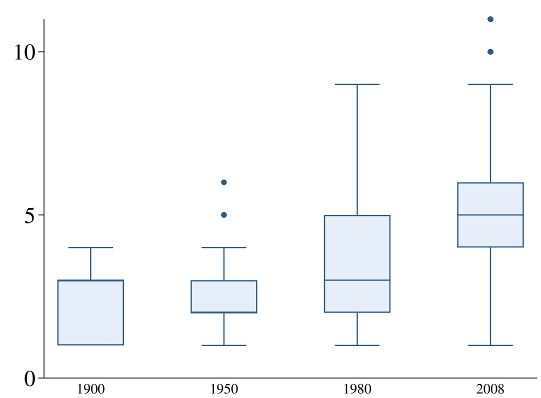
(b) Government Tax Revenue



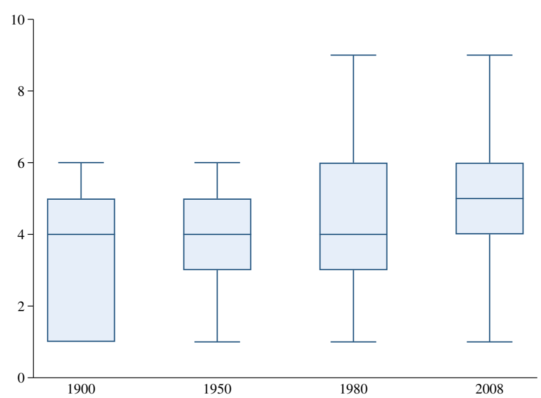
(c) Government Expenditure



(d) Government Deficits

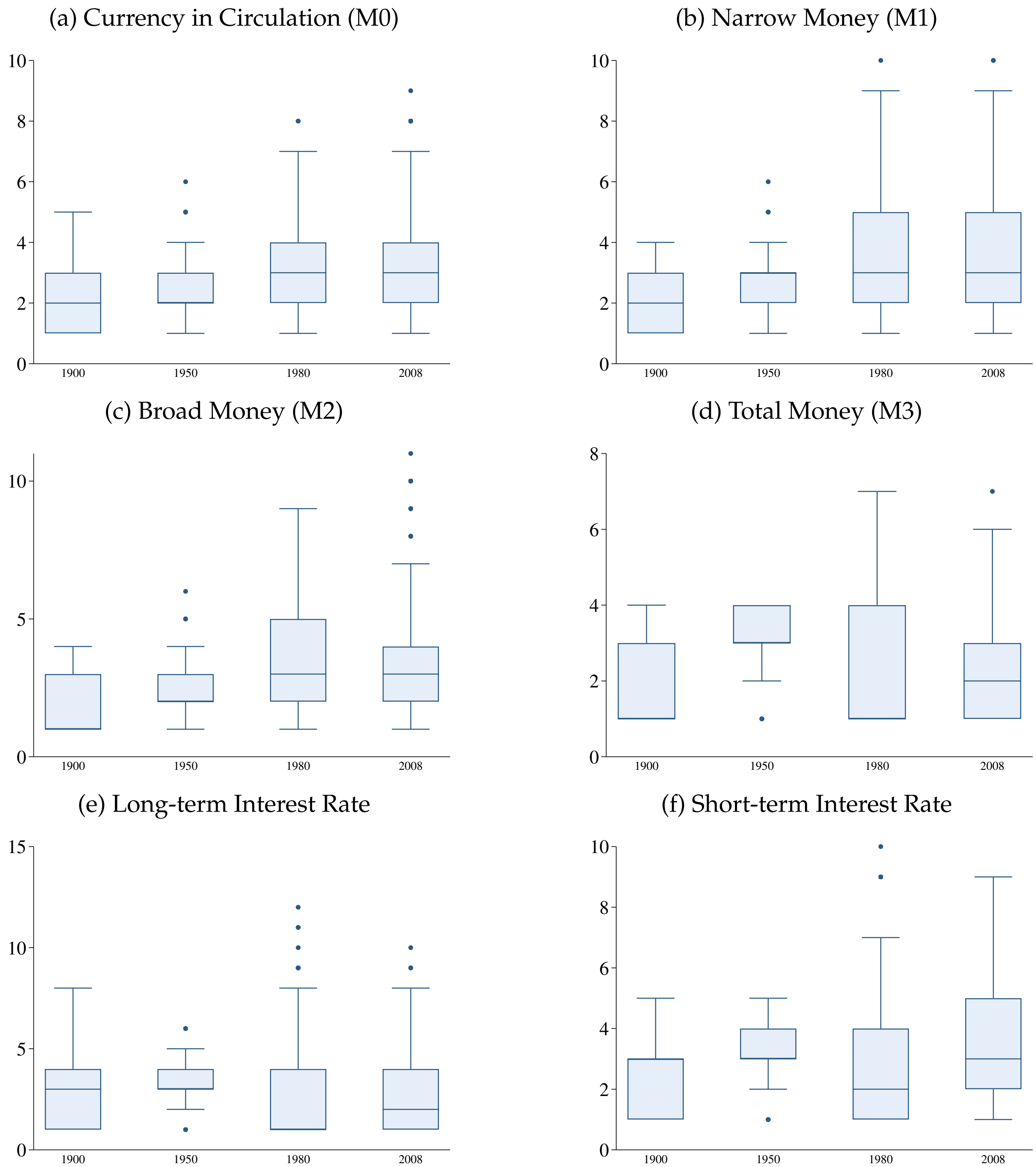


(e) Government Debt-to-GDP



Note: These figures visualize the number of sources we draw on to splice the final values in the Global Macro Database (GMD) for four benchmark years (1900, 1950, 1980, and 2008). These box plots show the median, interquartile range, upper and lower adjacent value, and outliers if any). For any variable x, the upper adjacent value is defined as $x[p75] + \frac{3}{2}(x[p75] - x[p25])$, where $p75$ and $p25$ refers to the 75$^{\text{th}}$ and 25$^{\text{th}}$ percentile, respectively.

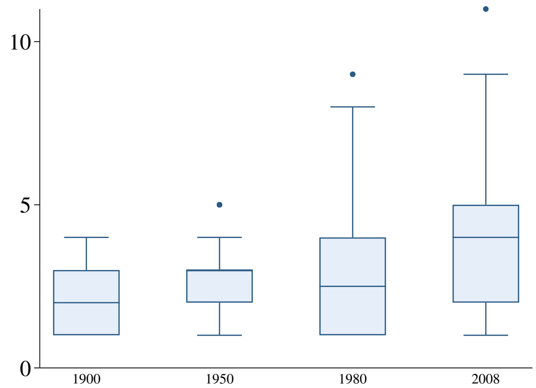
Figure OA4: Number of Sources per Variable – IV/V



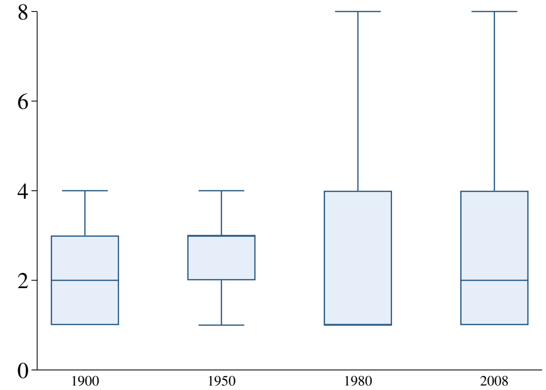
Note: These figures visualize the number of sources we draw on to splice the final values in the Global Macro Database (GMD) for four benchmark years (1900, 1950, 1980, and 2008). These box plots show the median, interquartile range, upper and lower adjacent value, and outliers if any). For any variable x, the upper adjacent value is defined as $x[p75] + \frac{3}{2}(x[p75] - x[p25])$, where $p75$ and $p25$ refers to the 75$^{\text{th}}$ and 25$^{\text{th}}$ percentile, respectively.

Figure OA5: Number of Sources per Variable – V/V

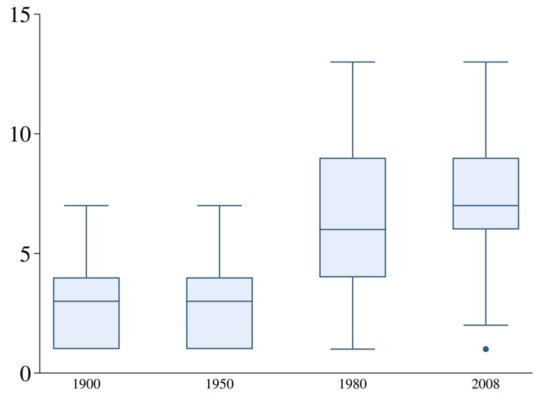
(a) Central Bank Policy Rate



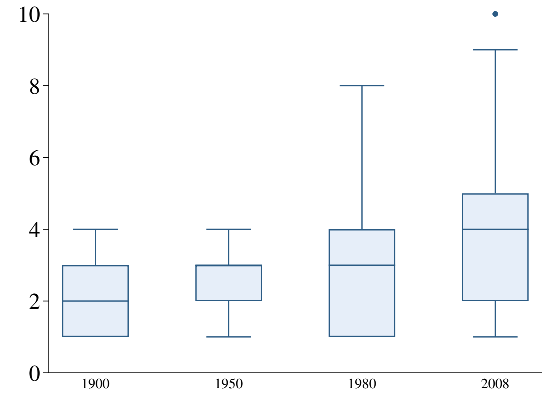
(b) House Price Index



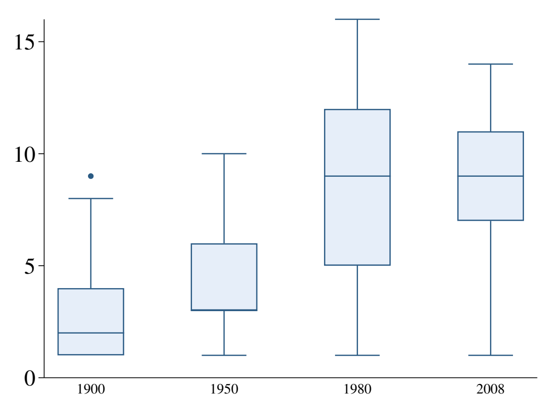
(c) Inflation



(d) Unemployment Rate



(e) Unemployment Rate



Note: These figures visualize the number of sources we draw on to splice the final values in the Global Macro Database (GMD) for four benchmark years (1900, 1950, 1980, and 2008). These box plots show the median, interquartile range, upper and lower adjacent value, and outliers if any). For any variable x, the upper adjacent value is defined as $x[p75] + \frac{3}{2}(x[p75] - x[p25])$, where $p75$ and $p25$ refers to the 75$^{\text{th}}$ and 25$^{\text{th}}$ percentile, respectively.

OA.2 Data Sources

Table OA1: Dataset Overview

Source	Abbreviation	Updated	Digitized	From	To	Forecasts	Variables	Countries	Historical
Panel A: Aggregator Sources									
Asian Development Bank (2024)	ADB	2024/10/14	No	2000	2023	0	26	49	No
African Development Bank (2018)	AFDB	2025/01/18	No	1980	2020	0	14	53	No
African Union (2024)	AFRISTAT	2024/07/20	No	1990	2023	0	8	22	No
Institute of Economic Research, Hitotsubashi University (2020)	AHSTAT	2024/05/04	No	1860	2013	0	20	6	Yes
European Commission (2024a)	AMECO	2025/01/18	No	1960	2025	1	16	43	No
Arab Monetary Fund (2024)	AMF	2024/07/10	No	1971	2021	0	18	22	No
Barro and Ursúa (2012)	BARRO	2024/05/04	No	1800	2009	0	2	42	Yes
Banque Centrale des Etats de l'Afrique de l'Ouest (2024)	BCEAO	2025/01/18	No	1960	2024	0	23	8	No
Broadberry and Gardner (2022)	BG	2024/09/30	No	1885	2008	0	2	8	Yes
Bank for International Settlements (2024c)	BIS CPI	2025/01/18	No	1661	2023	0	1	62	No
Bank for International Settlements (2024f)	BIS HPI	2025/01/18	No	1927	2024	0	1	58	No
Bank for International Settlements (2024d)	BIS REER	2025/01/18	No	1994	2024	0	1	63	No
Bank for International Settlements (2024a)	BIS USDfx	2025/01/18	No	1791	2024	0	1	189	No
Bank for International Settlements (2024b)	BIS cbrate	2025/01/18	No	1945	2024	0	1	45	No
Bank for International Settlements (2024e)	BIS infl	2025/01/18	No	1662	2023	0	1	62	No
Banca d'Italia (2024)	BIT	2025/01/10	No	1955	2024	0	1	9	Yes
Bordo et al. (2001)	BORDO	2024/04/21	No	1880	1997	0	9	56	Yes
Darvas (2021)	BRUEGEL	2024/04/21	No	1960	2023	0	1	178	Yes
Baron et al. (2020b)	BVX	2024/04/21	No	1870	2016	0	7	48	Yes
United Nations (2024a)	CEPAC	2024/07/11	No	1950	2023	0	23	36	No
International Institute of Social History (2024)	CLIO	2025/01/10	No	1727	2011	0	1	42	Yes

Continued on next page

Continued from previous page

Source	Abbreviation	Update	Digitized	From	To	Forecasts	Variables	Countries	Historical
Mack and Martínez-García (2011)	DALLASFED	2024/04/21	No	1975	2023	0	2	26	Yes
University of California – Davis (2024b)	Davis	2024/06/01	No	1818	2012	0	1	55	Yes
European Commission (2024b)	EUS	2025/01/18	No	1960	2024	0	24	50	No
Food and Agriculture Organization of the United Nations (2024)	FAO	2024/12/10	No	1970	2023	0	3	214	Yes
University of California – Davis (2024a)	FLORA	2024/06/01	No	1799	1975	0	2	12	Yes
Banque de France (2024b)	FRANC ZONE	2025/01/18	No	1991	2019	0	6	15	Yes
Flandreau and Zumer (2009)	FZ	2024/06/24	No	1880	1913	0	14	16	Yes
Smits et al. (2009)	GNA	2024/06/28	No	1800	2005	0	3	17	Yes
Gapminder (2024)	Gapminder	2024/08/14	No	1800	2030	6	1	197	Yes
Grimm (2024)	Grimm	2024/10/30	No	1945	2023	0	1	166	Yes
Schuler (2015)	HFS	2024/06/28	No	1800	2008	0	30	64	Yes
Homer and Sylla (1996)	Homer Sylla	2024/10/28	No	1798	1989	0	3	26	Yes
Ellison et al. (2024)	IHD	2024/10/07	No	1925	1936	0	7	39	Yes
International Labour Organization (2024)	ILO	2024/12/10	No	2000	2023	0	1	215	Yes
Mauro et al. (2015)	IMF FPP	2024/10/28	No	1800	2022	0	4	151	Yes
Mbaye et al. (2018)	IMF GDD	2024/04/21	No	1950	2018	0	3	185	Yes
International Monetary Fund (2024a)	IMF GFS	2025/01/18	No	1972	2023	0	16	158	No
International Monetary Fund (2010)	IMF HDD	2024/06/16	No	1800	2015	0	1	188	Yes
International Monetary Fund (2024b)	IFS	2025/01/18	No	1920	2024	0	19	222	No
International Monetary Fund (2024c)	IMF MFS	2025/01/18	No	1950	2023	0	7	162	No
International Monetary Fund (2024d)	WEO	2025/01/18	No	1960	2029	5	20	208	No
Albers et al. (2023)	JERVEN	2024/10/07	No	1890	2015	0	4	50	Yes
Jones and Obstfeld. (1997)	JO	2024/06/07	No	1850	1945	0	5	13	Yes
Jordà et al. (2017)	JST	2024/06/25	No	1870	2020	0	26	18	Yes
Ljungberg (2019)	LUND	2024/10/14	No	1870	2016	0	1	27	Yes

Continued on next page

Continued from previous page

Source	Abbreviation	Update	Digitized	From	To	Forecasts	Variables	Countries	Historical
Laeven and Valencia (2013)	LV	2024/10/17	No	1970	2017	0	4	155	Yes
Inklaar et al. (2018)	MAD	2024/04/04	No	1253	2022	0	3	169	Yes
Cox and Dincecco (2021)	MD	2024/06/20	No	1650	2010	0	6	31	Yes
Bértola and Rey (2018)	MOXLAD	2024/04/21	No	1870	2010	0	7	20	Yes
Officer and Williamson (2024)	MW	2024/10/02	No	1209	2023	0	10	41	Yes
Mitchell (2013)	Mitchell	2024/04/21	Yes	1750	2010	0	19	142	Yes
National Bank of Serbia (2024)	NBS	2024/06/24	No	1833	1950	0	23	8	Yes
Organisation for Economic Co-operation and Development (2024a)	OECD EO	2025/01/18	No	1960	2025	1	30	49	No
Organisation for Economic Co-operation and Development (2024b)	OECD HPI	2025/01/18	No	1960	2023	0	2	47	No
Organisation for Economic Co-operation and Development (2024c)	OECD KEI	2025/01/18	No	1914	2023	0	6	47	No
Organisation for Economic Co-operation and Development (2024d)	OECD MEI	2025/01/18	No	1935	2023	0	5	47	No
Organisation for Economic Co-operation and Development (1986)	OECD MEI ARC	2024/10/30	No	1955	1984	0	3	18	No
Organisation for Economic Co-operation and Development (2024e)	OECD QNA	2025/01/18	No	1947	2023	0	5	48	No
Organisation for Economic Co-operation and Development (2024f)	OECD REV	2025/01/18	No	1970	2022	0	1	38	No
Feenstra et al. (2015)	PWT	2024/04/21	No	1950	2019	0	3	183	Yes
Reinhart and Rogoff (2009)	RR	2024/04/21	No	1719	2016	0	5	71	Yes
Reinhart and Rogoff (2010)	RR debt	2024/04/21	No	1719	2010	0	1	68	Yes
Schmelzing (2019)	Schmelzing	2024/07/10	No	1310	2018	0	1	8	Yes
Albers (2018)	TH ID	2024/10/07	No	1925	1936	0	2	28	Yes
Federico and Tena-Junguito (2019)	Tena	2024/09/25	No	1800	1938	0	6	150	Yes
United Nations (2024b)	UN	2025/01/18	No	1950	2020	0	8	239	No
Ha et al. (2023)	WB CC	2024/09/05	No	1970	2023	0	2	207	Yes
World Bank (2024)	WDI	2025/01/18	No	1960	2023	0	30	223	No
World Bank (1999)	WDI ARC	2024/07/24	No	1960	1997	0	33	209	Yes

Continued on next page

Continued from previous page

Source	Abbreviation	Update	Digitized	From	To	Forecasts	Variables	Countries	Historical
Panel B: Country Specific Sources									
Nakamura and Zarazaga (2001)	ARG 1	2024/06/24	Yes	1901	1935	0	1	1	Yes
Ministerio de Economia de la Republica Argentina (2024)	ARG 2	2024/06/24	No	1940	2023	0	6	1	Yes
Hutchinson and Ploeckl (2024)	AUS 1	2024/04/21	No	1789	2020	0	6	1	Yes
Vamplew (1987)	AUS 2	2024/07/04	Yes	1788	1917	0	13	1	Yes
Schulze (2000)	AUT 1	2024/04/21	No	1870	1913	0	2	1	Yes
Instituto de Pesquisa Econômica Aplicada (2024)	BRA 1	2024/10/08	No	1872	2023	0	3	1	Yes
Statistics Canada (2024)	CAN 1	2024/10/02	No	1867	1977	0	19	1	Yes
Swiss National Bank (2009)	CHE 1	2024/06/26	No	1907	2005	0	5	1	Yes
Historical Statistics of Switzerland (2012)	CHE 2	2024/10/02	No	1851	1992	0	7	1	Yes
Abildgren (2017)	DNK 1	2024/06/24	No	1487	2023	0	14	1	Yes
Bank of Algeria (2023)	DZA 1	2024/07/24	No	1974	2023	0	3	1	Yes
Instituto Nacional de Estadística (2024)	ESP 1	2025/01/18	No	1995	2023	0	6	1	No
Banco de España (2024)	ESP 2	2024/12/29	No	1277	2014	0	11	1	Yes
Banque de France (2024a)	FRA 1	2025/01/18	No	1970	2024	0	3	1	No
Levy-Garboua and Monnet (2016)	FRA 2	2024/09/30	No	1800	2015	0	3	1	Yes
Thomas et al. (2010)	GBR 1	2024/06/18	No	1086	2016	0	17	1	Yes
Bank Indonesia (2023)	IDN 1	2025/01/18	No	2008	2018	0	14	1	No
Statistics Iceland (1997b)	ISL 1	2024/04/21	Yes	1870	2016	0	3	1	Yes
Statistics Iceland (1997a)	ISL 2	2024/09/26	No	1625	1990	0	28	1	Yes
Baffigi (2013)	ITA 1	2024/07/04	No	1861	2011	0	9	1	Yes
Piselli and Vercelli (2023)	ITA 2	2024/07/04	No	1861	2016	0	14	1	Yes
Istituto Nazionale di Statistica (2024)	ITA 3	2025/01/18	No	1995	2023	0	7	1	No
Bank of Japan (2024)	JPN 1	2024/10/08	No	1882	2017	0	7	1	Yes
Cha et al. (2022)	KOR 1	2024/05/13	No	1911	2016	0	4	1	Yes

Continued on next page

Continued from previous page

Source	Abbreviation	Update	Digitized	From	To	Forecasts	Variables	Countries	Historical
Gardner (2022)	LBR 1	2024/09/30	No	1845	1979	0	7	1	Yes
Bank Al-Maghrib (2001)	MAR 1	2024/07/24	No	1985	2024	0	4	1	Yes
Grytten (2022)	NOR 1	2024/09/25	No	1816	2019	0	3	1	Yes
Eitrheim et al. (2023)	NOR 2	2024/07/08	No	1516	2022	0	15	1	Yes
Statistics Poland (2024)	POL 1	2025/01/18	No	1989	2020	0	15	1	No
Instituto Nacional de Estatística (2001)	PRT 1	2024/07/08	Yes	1549	1998	0	12	1	Yes
Saudi Central Bank (2024)	SAU 1	2025/01/18	No	1973	2020	0	7	1	No
Schön and Krantz (2017)	SWE 1	2024/04/21	No	1290	2020	0	20	1	Yes
Central Bank of the Republic of T¸rkiye (2024)	TUR 1	2025/01/18	No	1994	2024	0	1	1	No
National Statistics, Republic of China (Taiwan) (2024)	TWN 1	2024/05/13	No	1951	2021	0	4	1	Yes
Federal Reserve Bank of St. Louis (2024)	FRED	2025/01/18	No	1929	2024	0	22	1	No
Carter et al. (2006)	USA 2	2024/09/26	No	1774	2003	0	19	1	Yes
South African Reserve Bank (2024)	ZAF 1	2025/01/18	No	1959	2020	0	17	1	No