Appendix
New Estimates of Over 500 Years of Historic GDP and Population Data

Christopher J. Fariss
Department of Political Science
University of Michigan
cjf0006@gmail.com

Therese Anders
School of International Relations
University of Southern California
tanders.usc@gmail.com

Jonathan N. Markowitz
School of International Relations
University of Southern California
jnmarkowitz@gmail.com

Miriam Barnum
School of International Relations
University of Southern California
mbarnum@usc.edu

September 22, 2021

Journal of Conflict of Resolution (Forthcoming)

Contents

A Dataset Descriptions **4**
- A.1 Maddison Project Database 2020 (Bolt and van Zanden, 2020; Bolt et al., 2018) . . 4
- A.2 Penn World Tables 10.0 (Feenstra, Inklaar and Timmer, 2015) 5
- A.3 World Development Indicators (World Bank, N.d.) 6
- A.4 Broadberry and Klein (2012) 7
- A.5 Bairoch (1976) 7
- A.6 Broadberry (2015) 8
- A.7 Measuring Worth (2019) 8

B Dataset Coverage **9**
- B.1 Dataset coverage tables 9
- B.2 Proportion of country-year units with missingness for the Maddison Project variables 13
- B.3 Proportion of country-year units with missingness for any variable 14
- B.4 Proportion of country-year units with missingness for each variable 15

C Posterior prediction intervals and observed variables **16**
- C.1 Summary of posterior prediction intervals for cases with and without an observed dataset value 17
- C.2 United States 18
- C.3 Netherlands 19
- C.4 France 20
- C.5 Spain 21
- C.6 Italy 22
- C.7 Russia/Soviet Union 23
- C.8 Sweden 24
- C.9 Albania 25
- C.10 India 26
- C.11 Colombia 27
- C.12 Brazil 28
- C.13 Democratic Republic of Congo/Zaire 29
- C.14 Uganda 30
- C.15 Iran 31
- C.16 Afghanistan 32
- C.17 North Korea 33
- C.18 Pakistan 34
- C.19 Kosovo 35
- C.20 East Timor 36
- C.21 Eritrea 37
- C.22 Cross-sectional comparisons 1500 38
- C.23 Cross-sectional comparisons 1600 39
- C.24 Cross-sectional comparisons 1700 40
- C.25 Cross-sectional comparisons 1800 41
- C.26 Country-year variable examples from 1990 42

D New Results from Expanded Measurement Coverage **46**
- D.1 Measures of Health and Wellbeing 47
  - D.1.1 Correlations between GDP per capita and Life Expectancy, 1700 – Present . 47
  - D.1.2 Correlations between GDP per capita and Child Mortality, 1800 – Present . 48

D.1.3 Correlations between GDP per capita and Literacy Rates, 1500 – Present . . 49
D.2 Measures of Democracy . . . 50
D.2.1 Correlations between GDP per capita and Democracy (VDEM Polyarchy) 1790 – Present . . . 50
D.2.2 Correlations between GDP per capita and Democracy (Polity IV) 1800 – Present . . . 51
D.3 Measures of Repression . . . 52
D.3.1 Correlations between GDP per capita and Repression (VDEM Killing) 1790 – Present . . . 52
D.3.2 Correlations between GDP per capita and Repression (Latent Human Rights Scores) 1946 – Present . . . 53
D.4 Measures of Power Projection . . . 54
D.4.1 Correlations between GDP per capita and Naval Arming (Proportion of Global Ship Count), 1655 – Present . . . 54
D.4.2 Correlations between GDP per capita and Naval Arming (Ship Tonnage), 1870 – Present . . . 55
D.5 Measures of Conflict . . . 56
D.5.1 Correlations between GDP per capita and Militarized Interstate Disputes 1816 – Present . . . 56
D.5.2 Correlations between GDP per capita and Interstate Wars 1816 – Present . . 57
D.5.3 Correlations between GDP per capita and the Total Combat Fatalities 1946 – Present . . . 58

E Analysis of Measurement Uncertainty **59**
E.1 Regression analysis with uncertainty . . . 59
E.2 Regression analysis of the VDEM Killing variable with uncertainty from GDP per capita (1946 onwards sample) . . . 59
E.3 Regression analysis of the Battle Deaths variable with uncertainty from GDP per capita (1946 onwards sample) . . . 61
E.4 Additional Suggestions for Assessing the Influence of Missingness in Data . . . 63

F Improving the Validity of Estimates for Missingness **64**
F.1 Gleditsch's Approach to Missingness . . . 64
F.2 Comparing Estimates for Missing Cases . . . 66

G Unique Country Coverage by Variable Type **70**

A Dataset Descriptions

Here we describe in detail each of the datasets we included in our unified measurement model of GDP, GDP per capita, and population. For observed data on GDP see Measuring Worth (2019); World Bank (N.d.); Feenstra, Inklaar and Timmer (2015); Broadberry and Klein (2012); Bairoch (1976). For observed data on GDP per capita see Measuring Worth (2019); Bolt et al. (2018); Bolt and van Zanden (2020); World Bank (N.d.); Broadberry (2015); Broadberry and Klein (2012); Bairoch (1976), and for observed data on population see Bolt and van Zanden (2020); Measuring Worth (2019); World Bank (N.d.); Feenstra, Inklaar and Timmer (2015); Broadberry and Klein (2012); Deng (2004). For each component dataset, we extract relevant indicators, attach unique country identifiers, and reshape the data into a common country-year format. Details on the underlying source materials for each component measure and coding decisions are provided here are documented in the `R` code that we use to merge the constituent datasets together.

A.1 Maddison Project Database 2020 (Bolt and van Zanden, 2020; Bolt et al., 2018)

Maddison's original GDP per capita and population variables are derived from a large number of country-level sources (Maddison, 2003, 2001, 1995). Because the underlying source materials employed by Maddison are expansive and country-specific, we refrain from describing them in detail. The most recent version, the Maddison Project Database (MPD, Bolt and van Zanden 2020), is based on a collaboration of researchers dedicated to continuing Angus Maddison's data collection efforts by extending and, if warranted, revising his estimates. Building on Maddison's original work that relies on a single cross-country income comparison in 1990 using Geary-Khamis multilateral PPPs, the 2018 version of the MPD adds an indicator of GDP per capita (PPP) that incorporates multiple benchmarks, similar to the methodology used in the new Penn World Tables (Feenstra, Inklaar and Timmer, 2015). This includes all rounds of the International Comparison Program (ICP), as well as a variety of historical benchmarks (Bolt et al., 2018). Such an innovation in the construction of per capita GDP allows for income comparisons between countries at a single point in time. With a few exceptions, data from 1990–2010 were revised using figures from the Total Economy Database of the Conference Board (Bolt and van Zanden, 2014). Other estimates

are based on historical national statistics from country-specific sources (Bolt and van Zanden, 2014).

In the latest version of the MPD (Bolt and van Zanden, 2020), chaining PPP benchmarks from multiple rounds of the ICP in the MDP 2020 RGDPpc indicator allows for a comparison of living standards across countries and over time. This methodology is similar to that used in recent versions of the Penn World Tables. For benchmark years, GDP is made comparable across space using PPP exchange rates. Comparability over time is achieved by interpolating between benchmark years—which results in GDP growth rates that are static between benchmark years and typically differ from National Account rates (Feenstra, Inklaar and Timmer, 2015; Bolt and van Zanden, 2020, 3156).

We subset the data from the MPD to include only country-year observations starting in 1500. The original Maddison (2010) data includes both GDP and GDP per capita values. The updated version only includes GDP per capita values.

A.2 Penn World Tables 10.0 (Feenstra, Inklaar and Timmer, 2015)

The Penn World Tables (PWT) are among the most widely-used sources of data on GDP and population in the post-WWII period. The PWT relies on national accounts as the original data source for GDP in each country. The major contribution of these datasets to our understanding of economic wealth worldwide is their careful consideration of differences in prices in either the temporal dimension, the cross-country dimension, or both. Starting with version 8.0, the PWT employs a new methodology to make national income figures comparable across time, countries, or both using Purchasing Power Parities (PPP) that are constructed with a wide range of historical price surveys. The two latest releases of the PWT distinguish between expenditure-side and output-based versions of GDP. Expenditure-side GDP is suitable to measure standard of living, while output-side GDP should be used when assessing productive capacity (Feenstra, Inklaar and Timmer, 2015, 3151). Output-based computations of GDP differ from the expenditure approach in that they account for relative prices in imports and exports, not just consumption and investment (Feenstra, Inklaar and Timmer, 2015, 3153). By including only expenditure-based GDP variables from the Penn World Tables in our model, we ensure that the series are conceptually similar to earlier data sources that were unable to account for price differences in traded goods. One of

the most commonly used GDP datasets for scholars of international relations is data produced and updated by Gleditsch (2002), which are based on the PWT dataset. We include this dataset because it is central to the study of international relations. As we describe in more detail below, our model produces estimates that are based on the methodology of each variable included in our model. Interested readers can select the most appropriate dataset and employ our estimated distributions of that original variable in their analysis.

A.3 World Development Indicators (World Bank, N.d.)

We include data on GDP, GDP per capita, and population from the World Bank (N.d.). We include both the World Bank's (N.d.) GDP indicator measured in exchange rate based constant 2010 US dollars as well as a GDP indicator in 2011 international dollars (PPP). The figures are compiled from the World Bank and OECD national accounts data. The PPP series leverages the 2011 ICP round to incorporate cross-country price comparisons. The documentation in the metadata file indicates that the series is based on an underlying interpolation of component data upon aggregating it to a "gap-filled total." Unfortunately, we do not have information on the details of this aggregation process. We therefore use the full series of GDP as provided by the World Bank's (N.d.) online data portal DataBank. In future versions of our model, we plan to identify these interpolated cases when possible and adjust our model accordingly.

The GDP per capita data are based on the World Bank's (N.d.) total population figures and either GDP in constant 2010 US dollars or GDP in 2011 international dollars (PPP). The constant series compares values across countries using market exchange rates while the PPP series relies on price comparisons.

Finally, the population figures from World Bank (N.d.) are based on national population censuses. The census data that informs this measure stem from a variety of sources, including the United Nations World Population Prospects (for the majority of developing countries), Eurostat (for European countries), and national statistical agencies. The data are interpolated for all years between census years. Since we do not have information on the years that a census was conducted for each country, we retain the interpolated data for the use in the latent variable model.

A.4 Broadberry and Klein (2012)

The GDP, GDP per capita, and population variables in Broadberry and Klein (2012) are limited to European countries (including Russia and Turkey) as well as the United States. A detailed list of underlying source material is available in the paper's appendix (Broadberry and Klein, 2012, pp. 105). For GDP, these sources include the data from Maddison (2010), official national account statistics, and the work of country-expert historians. Data on population are drawn mainly from Mitchell (2003) and Maddison (2010), and supplemented with country-specific data from official national statistics and historians.

A.5 Bairoch (1976)

The underlying source material for the data by Bairoch is detailed in the paper's methodological appendix. For GNP, these sources include the work of historians and official national statistics for earlier country-years as well as OECD figures for years starting in 1950 (Bairoch, 1976, 329 et seq.). For the 19th century and the year 1900, three-year annual averages are available for every decade starting from 1830 (Bairoch, 1976, 286). For the 20th century, data are available for select years between 1913 and 1973 (Bairoch, 1976, 297). The GNP values of individual countries are converted into 1960 US dollars using price comparison data from Milton Gilbert, not market exchange rates (Bairoch, 1976, 318, appendix E3). We therefore consider the data from Bairoch (1976) to be analogous to PPP series. With the exception of the data from Bairoch (1976), the data on total economic size are measured as the gross *domestic* product (GDP). Bairoch (1976) uses gross *national* product (GNP) instead. While the GNP excludes value added by foreign firms, this measure is highly correlated with GDP.

For population, Bairoch relies on United Nations Demographic yearbooks, data from the League of Nations, and national statistical agencies to assemble his data (321). We incorporate all of Bairoch's estimates in our model, including the country-year cases flagged as having a larger-than-average margin of error (the figures presented in parentheses).

A.6 Broadberry (2015)

The GDP per capita estimates in Broadberry (2015) are based on historical national accounting data that is constructed from documents such as "government accounts, customs accounts, poll tax returns, parish registers, city records, trading company records, hospital and educational establishment records, manorial accounts, probate inventories, farm accounts, tithe files and other records of religious institutions." (Broadberry, 2015, 5). Broadberry lists the data sources for each country in the main text.[1] As with the Maddison data, we exclude cases for years prior to 1500 from our model.

A.7 Measuring Worth (2019)

Measuring Worth (2019) is an online resource for historical GDP, per capita GDP, and population data for the United States (1790–2014), the United Kingdom (1801–2015), Australia (1828–2015), and Spain (1850–2015).

The data on the United States' GDP are constructed by Louis Johnston and Samuel H. Williamson and rely on figures from the U.S. Bureau of Economic Analysis and a number of economic historians. The GDP original data are presented in millions of 2009 dollars, which we transform into constant 2010 U.S. dollars using the U.S. Bureau of Labor Statistics Consumer Price Index. Data on the GDP of the United Kingdom are derived from the Bank of England's three centuries dataset and originally expressed in millions of constant 2013 pounds.[2] We transform the data into constant 2010 U.S. dollars using a Composite Price Index from the U.K. Office for National Statistics and historic Dollar–Pound exchange rates from Measuring Worth (2019). GDP data for Australia are constructed by Diane Hutchinson and Florian Ploeckl and come from the Australian Bureau of Census and Statistics and a number of economic historians. The data are expressed in constant 2010 U.S. dollars. Data on Spain are constructed by Leandro Prados de la Escosura and originally expressed in millions of constant 2010 Euros, which we convert to constant 2010 US dollars using the official market exchange rate.

[1] Pages 6 and 7 contain the underlying source material for Britain, the Netherlands, Italy, and Spain; page 8 contains the data for China, Japan, and India.

[2] Measuring Worth (2019) presents two alternative series to account for the independence of Ireland in 1921. We include the "Historic" series in our model that measures GDP and population for the historical component units of the United Kingdom at any point in time.

Deng (2004)

Deng (2004) re-examines pre-modern population censuses of Chinese bureaucratic records to construct a revised series of population. Rather than attempting to create a smooth growth curve of population, the author makes only minor adjustments to the official Chinese figures. He argues that even in pre-modern times, Chinese enumerators had the technical skills necessary to construct accurate series and harsh bureaucratic punishments meant that there was little fraud in censuses.

B Dataset Coverage

In this section, we discuss coverage of the datasets. In the section that follows we present several graphical representations of uncertainty and describe how uncertainty increases as a case moves away from a country-year case with an observed dataset value. Recall that a country enters the dataset in 1500 A.D., or the first year in which at least one of the datasets records a value for at least one of the included variables (e.g., England 1500-2015 A.D. or Ghana 1820-2015 A.D.). Table 1 shows the count and proportion of coverage by dataset type, variable type, and time period. Table 2 displays the total number of country-year units with at least one observed value from the Maddison dataset or any other dataset. Table 3 displays the total number of country-year units with at least one observed value from the PWT dataset or any other dataset. There is always at least one observed value for the first year a county enters the dataset. Missing values are estimated using the latent variable model we describe in the main article. Figures 1-3 display the proportion of coverage for the variables included in our measurement model.

B.1 Dataset coverage tables

period	type	proportion	count	total units	sources
1500-1799	GDP	0.00	10	13910	all
1500-1799	GDPpc	0.13	1835	13910	all
1500-1799	Population	0.01	182	13910	all
1500-1799	any	0.14	1962	13910	all
1500-1799	GDPpc	0.13	1810	13910	Maddison
1500-1799	Population	0.01	138	13910	Maddison
1500-1799	GDP	0.00	10	13910	non-Maddison
1500-1799	GDPpc	0.00	47	13910	non-Maddison
1500-1799	Population	0.00	44	13910	non-Maddison
1800-1949	GDP	0.10	1476	14330	all
1800-1949	GDPpc	0.33	4704	14330	all
1800-1949	Population	0.52	7454	14330	all
1800-1949	any	0.54	7802	14330	all
1800-1949	GDPpc	0.32	4626	14330	Maddison
1800-1949	Population	0.36	5162	14330	Maddison
1800-1949	GDP	0.10	1476	14330	non-Maddison
1800-1949	GDPpc	0.10	1468	14330	non-Maddison
1800-1949	Population	0.41	5945	14330	non-Maddison
1950-present	GDP	0.77	10049	13104	all
1950-present	GDPpc	0.87	11442	13104	all
1950-present	Population	0.98	12801	13104	all
1950-present	any	0.98	12801	13104	all
1950-present	GDPpc	0.79	10374	13104	Maddison
1950-present	Population	0.84	10974	13104	Maddison
1950-present	GDP	0.77	10049	13104	non-Maddison
1950-present	GDPpc	0.67	8769	13104	non-Maddison
1950-present	Population	0.91	11864	13104	non-Maddison
Full	GDP	0.28	11535	41344	all
Full	GDPpc	0.43	17981	41344	all
Full	Population	0.49	20437	41344	all
Full	any	0.55	22565	41344	all
Full	GDPpc	0.41	16810	41344	Maddison
Full	Population	0.39	16274	41344	Maddison
Full	GDP	0.28	11535	41344	non-Maddison
Full	GDPpc	0.25	10284	41344	non-Maddison
Full	Population	0.43	17853	41344	non-Maddison

Table 1: Total number of country years and proportion of country years with at least one observed value from any dataset, the Maddison Project dataset, or any dataset but the Maddison Project dataset for three time periods and the full time period covered. In the most recent period, nearly all country-year units (98%) are covered by at least one dataset (any variable). For the 1800-1949 period, a majority of country-year units (54%) are covered by at least one dataset (any variable). For the early historic period, 1500-1799, the proportion of covered country-year units (14%) is much smaller. This is why the uncertainty around the estimated ranges that are derived from our model are much larger for these earlier historic periods relative to the more recent periods.

Maddison	Variable Name	1500-1799	1800-1949	1950-2017
no	Broadberry_ppp_bc	15	4	0
no	WorldBank_con_bc	0	0	1035
no	WorldBank_ppp_bc	0	0	629
no	MW_con_bc	10	0	0
no	BroadberryKlein_ppp_bc	0	49	33
no	Bairoch_ppp_bc	0	30	4
yes	Broadberry_ppp_bc	22	10	0
yes	WorldBank_con_bc	0	0	7144
yes	WorldBank_ppp_bc	0	0	4045
yes	MW_con_bc	0	521	263
yes	BroadberryKlein_ppp_bc	0	1000	1487
yes	Bairoch_ppp_bc	0	233	92
yes	Maddison2018_ppp_bc	1810	4624	10374
yes	Maddison2018_ppp_bt	1605	4558	10374

Table 2: Total number of country years with coverage for each GDP per capita variable. The table is split into 3 sections. In the top section, we see coverage for variables when the Maddison Project does not cover a particular country-year unit. In the middle section, we see coverage for variables when the Maddison Project does cover a particular country-year unit. The lower section, we see coverage for the Maddison Project variables, which represents the total coverage of these variable for each time period. Total coverage for any of the other specific variables is the sum of the value in the upper and middle panel. For example, the Broadberry_ppp_bc variable covers a total of 37 country-year units in the period 1500-1799. For 22 of these units (middle panel), the Maddison Project also provides an observed value. For the other 15 of these units (upperpanel), the Maddison Project does not provide a value.

All sources

Period	Indicator	Proportion covered	Count covered	Total units
1950-present	GDP	0.77	10049	13104
1950-present	GDPpc	0.87	11442	13104
1950-present	Population	0.98	12801	13104

PWT only

Period	Indicator	Proportion covered	Count covered	Total units
1950-present	GDP	0.68	8958	13104
1950-present	Population	0.68	8958	13104

Non-PWT sources only

Period	Indicator	Proportion covered	Count covered	Total units
1950-present	GDP	0.67	8738	13104
1950-present	GDPpc	0.87	11442	13104
1950-present	Population	0.98	12801	13104

Table 3: Country years with coverage by time period for the Penn World Tables compared to data from any other source. The number of country-year units covered by the Penn World Tables for GDP per capita is 68% overall, and 87% when GDP per capita is covered by at least one dataset (any variable).

B.2 Proportion of country-year units with missingness for the Maddison Project variables

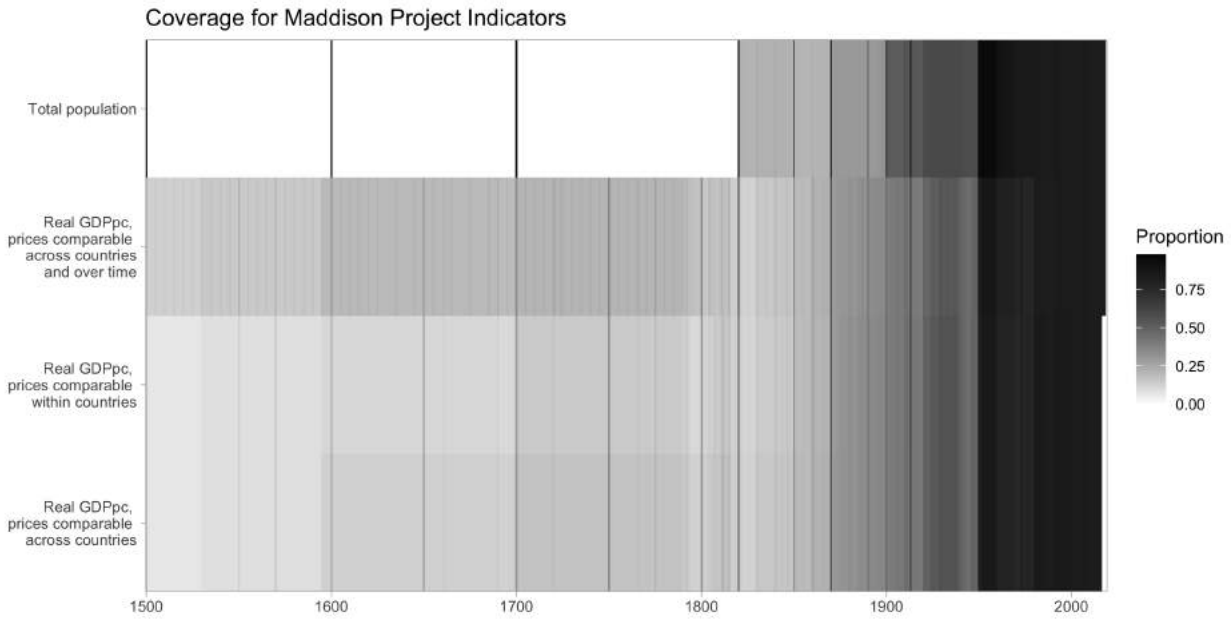

Figure 1: Proportion of country-year units in our sample covered by the Maddison Project.

B.3 Proportion of country-year units with missingness for any variable

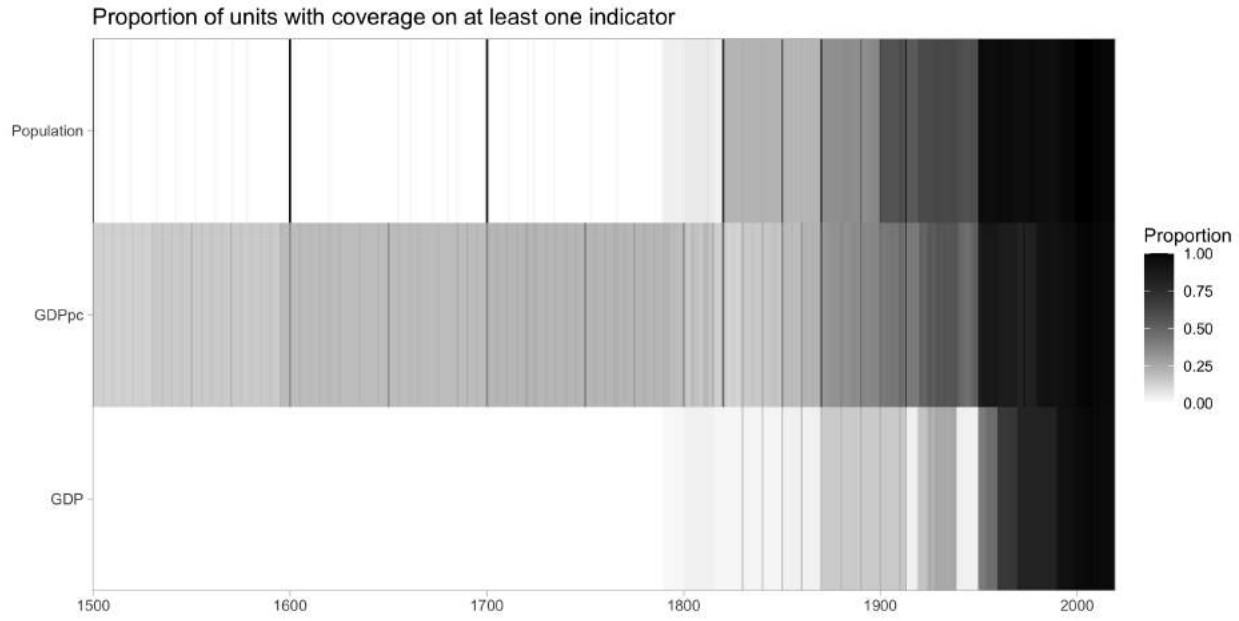

Figure 2: Proportion of country-year units in our sample covered by at least one dataset for each variable type (population, GDP per capita, GDP).

B.4 Proportion of country-year units with missingness for each variable

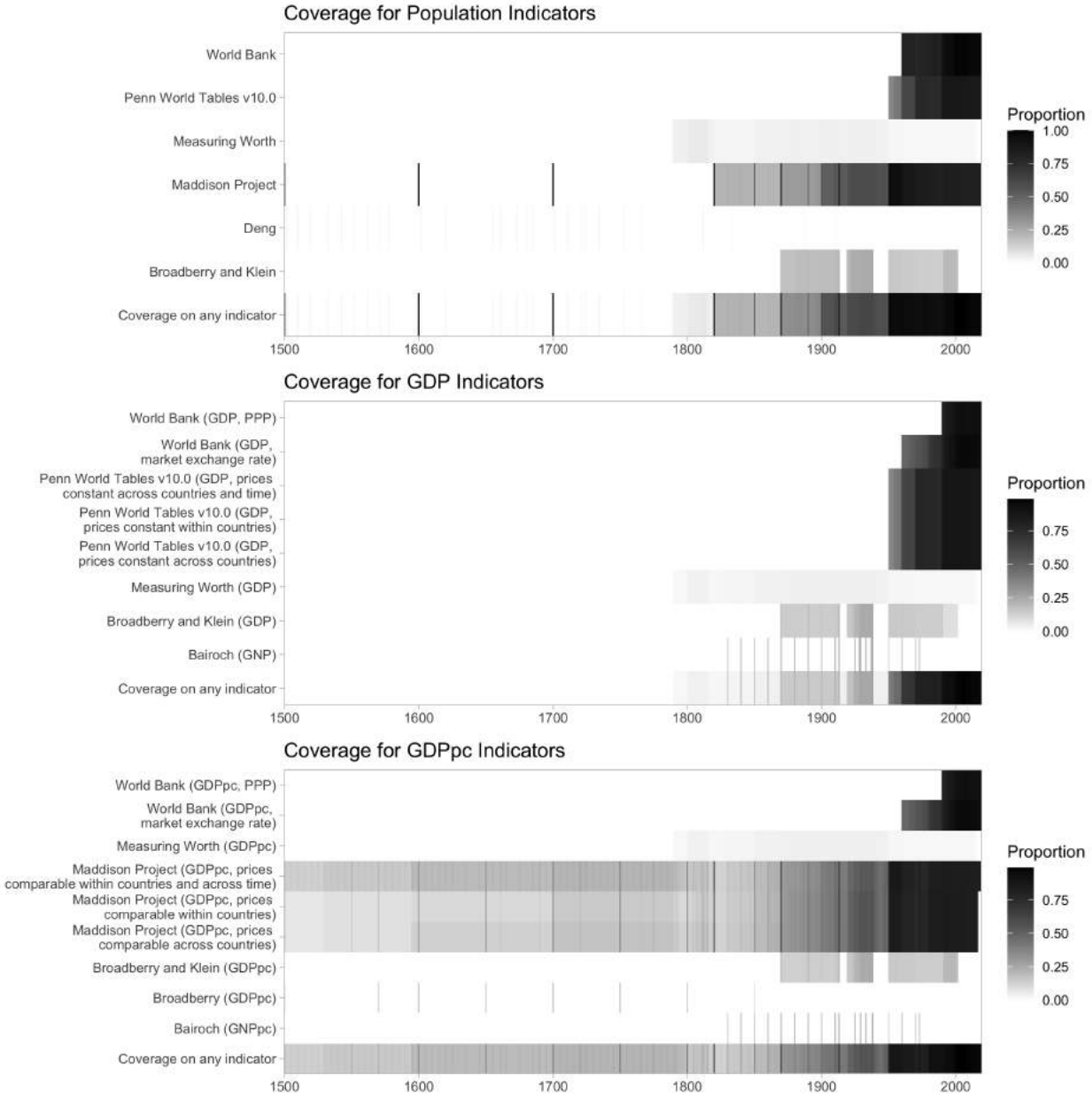

Figure 3: Proportion of country-year units in our sample covered by each dataset for each variable type (population, GDP per capita, GDP).

C Posterior prediction intervals and observed variables

If our measurement model does a good job of estimating GDP per capita, then it must be able to not only generate reasonable estimates for easy cases, but all countries in the sample, which include hard cases. Estimating GDP for hard cases is challenging as these countries often lacked bureaucratic capacity and their officials had strong incentive to lie about the data. As a result, existing dataset values of GDP per capita for these hard cases often vary widely between different datasets. We demonstrate that the observed dataset values fall within the range of our model based estimates, which provides strong evidence of convergent validity and increases our confidence that the measurement model is generating valid estimates even for hard cases

The first figure below displays the distribution of standard deviations based on each country-year posterior prediction interval. This shows how much more certain our model based estimates are for cases with an observed dataset value compared to country-year cases without an observed dataset value. The rest of the figures below display posterior prediction intervals (grey lines) with examples of both historically prominent cases and hard cases to showcase the ability of the measurement to cover the observed dataset values and bridge the areas of missing coverage when those dataset values (black points) do not exist for the United States, Netherlands, France, Spain, Italy, Russia, Sweden, Albania, India, Colombia, Brazil, Democratic Republic of Congo/Zaire, Uganda, Iran, Afghanistan, North Korea, Pakistan, Kosovo, East Timor, and Eritrea. We also show these distributions for one of the Maddison data project variables (GDP per capita in PPP international dollars) for all countries in the years 1500, 1600, 1700, and 1800. Finally, we discuss coverage of our estimates with tabular and visual data for three of the GDP per capita variables in 1990.

All of these visual and tabular presentations showcase the coverage of the estimated intervals for each variable and the location of the observed value for each variable. Each of these graphs demonstrate that the amount of variation of the latent standard deviation is greatest for units without information — i.e., observed dataset values. Thus, the further away a unit is from a unit with an observed dataset value, the greater the level of uncertainty that is captured by the estimate of the standard deviation. Importantly, the level of uncertainty is a country-year unit parameter that we can use in the estimation of any summary statistic that uses the country-year estimates. We provide additional details about how to incorporate uncertainty in the main article and below.

C.1 Summary of posterior prediction intervals for cases with and without an observed dataset value

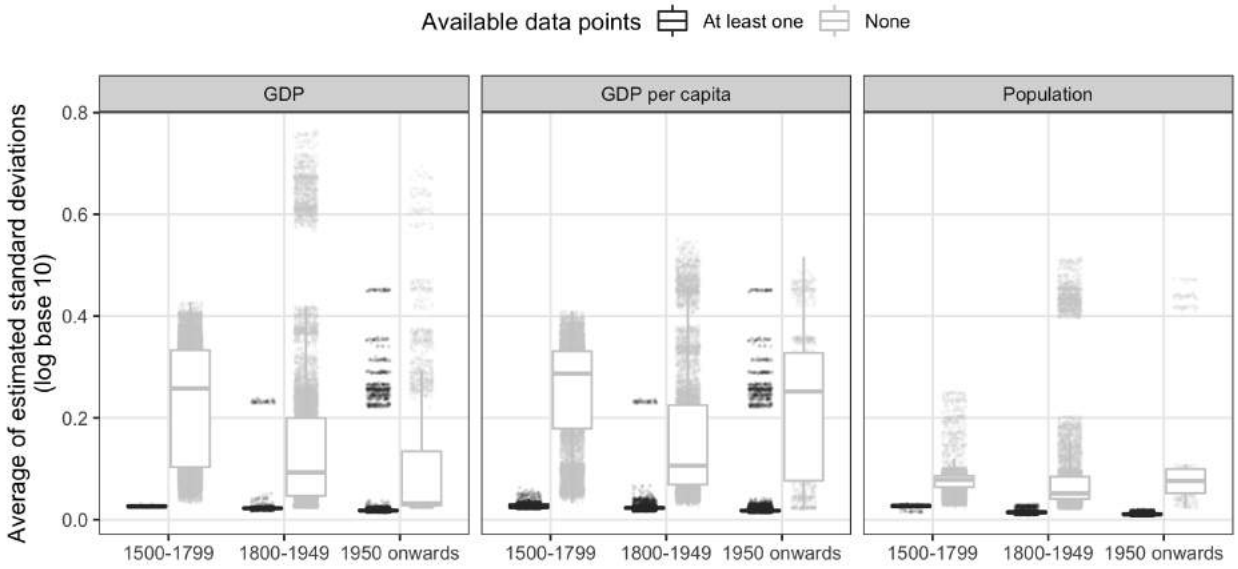

Figure 4: Plot of the distribution of country-year standard deviation estimates for those cases with at least one observed variable (black) and those cases with no observed values (grey). Our latent variable model estimates a range of possible values. The range is characterized by a normal density function, so the standard deviation for each country-year unit is a useful summary of the size of this range. The range of the country-year distribution is much larger when there is no observed value present and when that country-year is far away in time from another observed value. For example, for a country with an observed population value in 1500 and then 1600, the country-year unit with the largest uncertainty range (largest standard deviation) would be 1550. Most of the new estimates in the historic period have a much larger range than the estimates since 1950. This is why using the uncertainty estimates characterized by the standard deviation is critical when describing correlations and other patterns from these using our new estimates of population, GDP per capita, and GDP. We see the results clearly for a set of example countries in the next section of this appendix.

C.2 United States

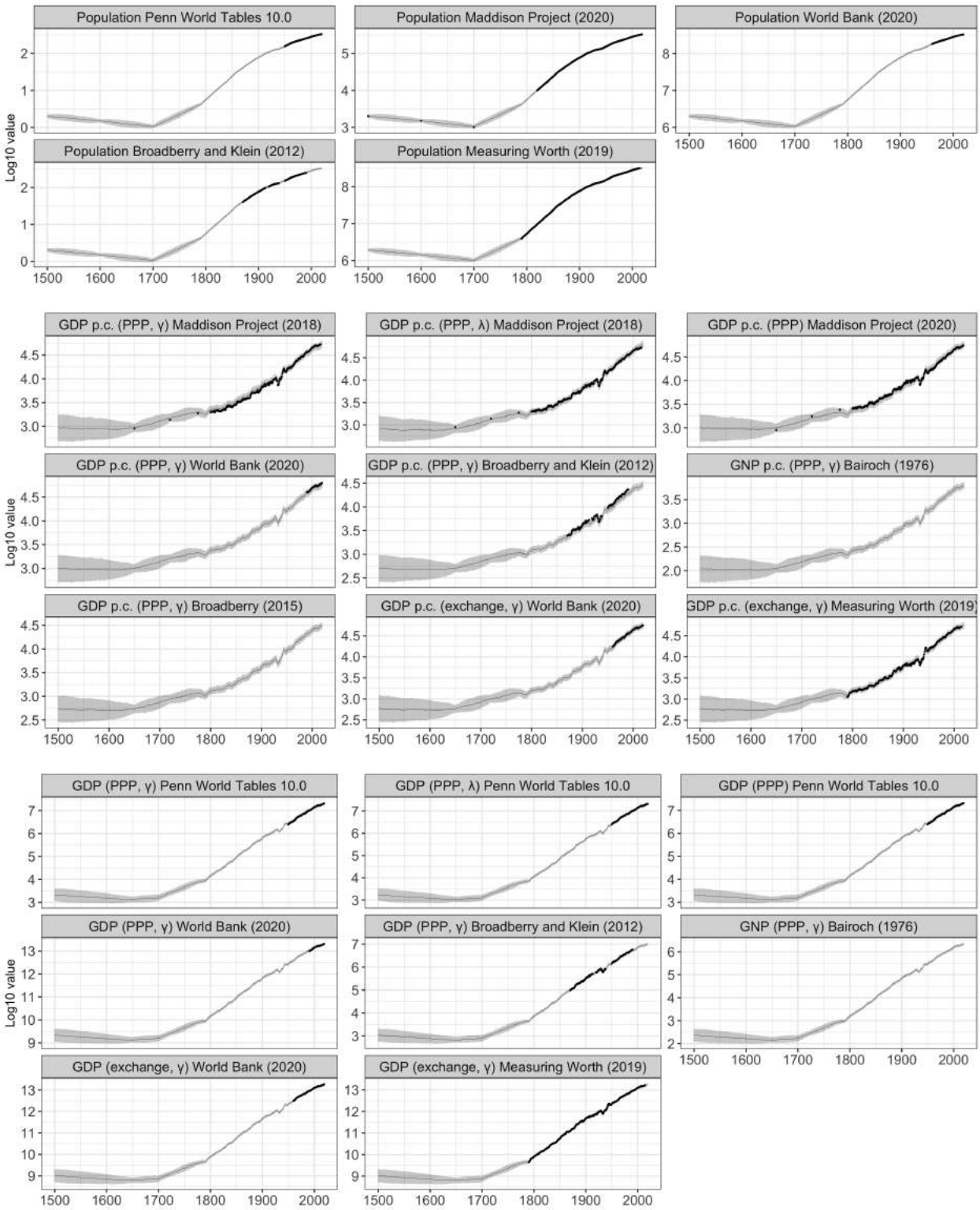

Figure 5: Posterior prediction intervals (grey lines) with ± 1 standard deviation confidence bands and observed variables (black points) for the United States.

C.3 Netherlands

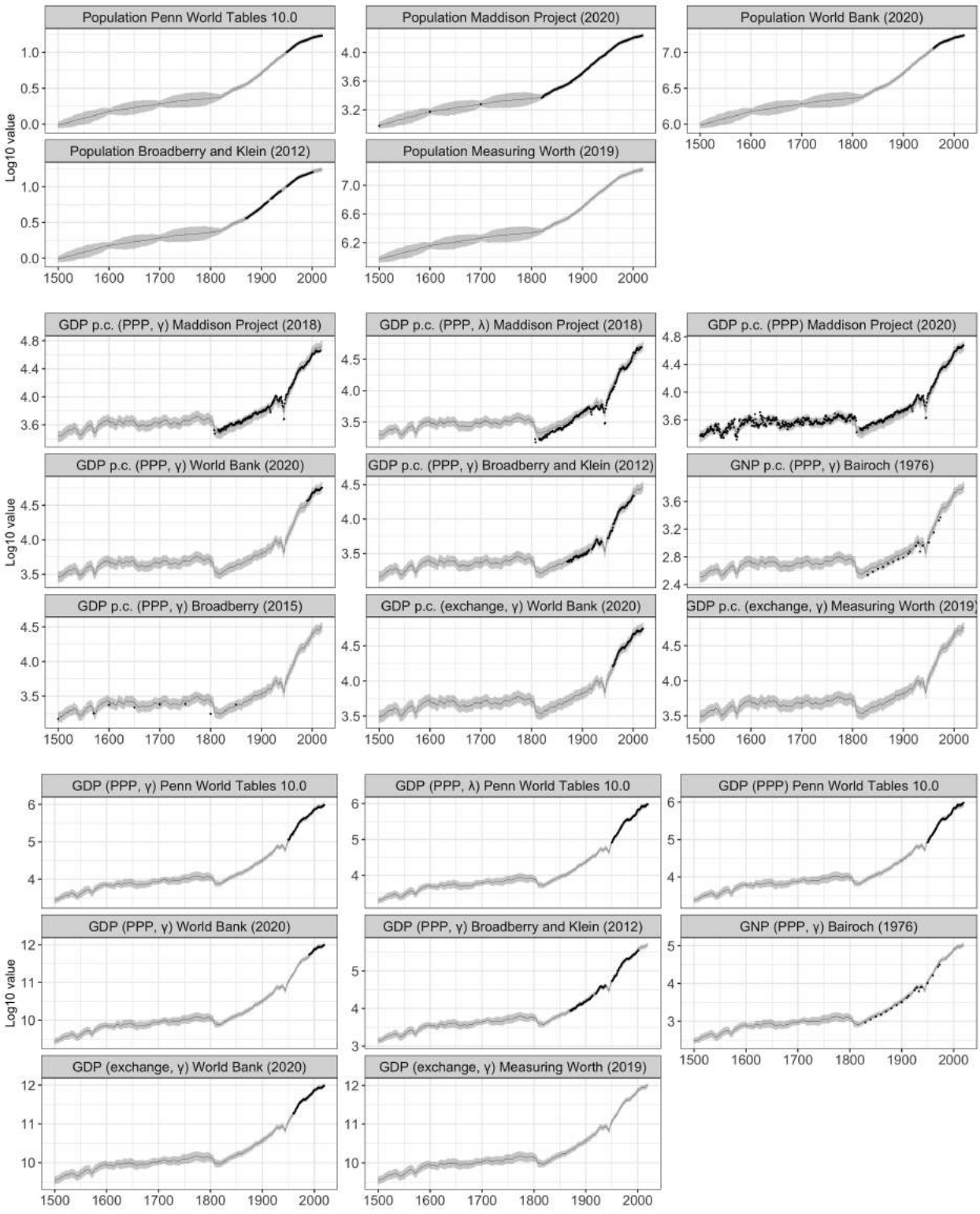

Figure 6: Posterior prediction intervals (grey lines) with ± 1 standard deviation confidence bands and observed variables (black points) for the Netherlands.

C.4 France

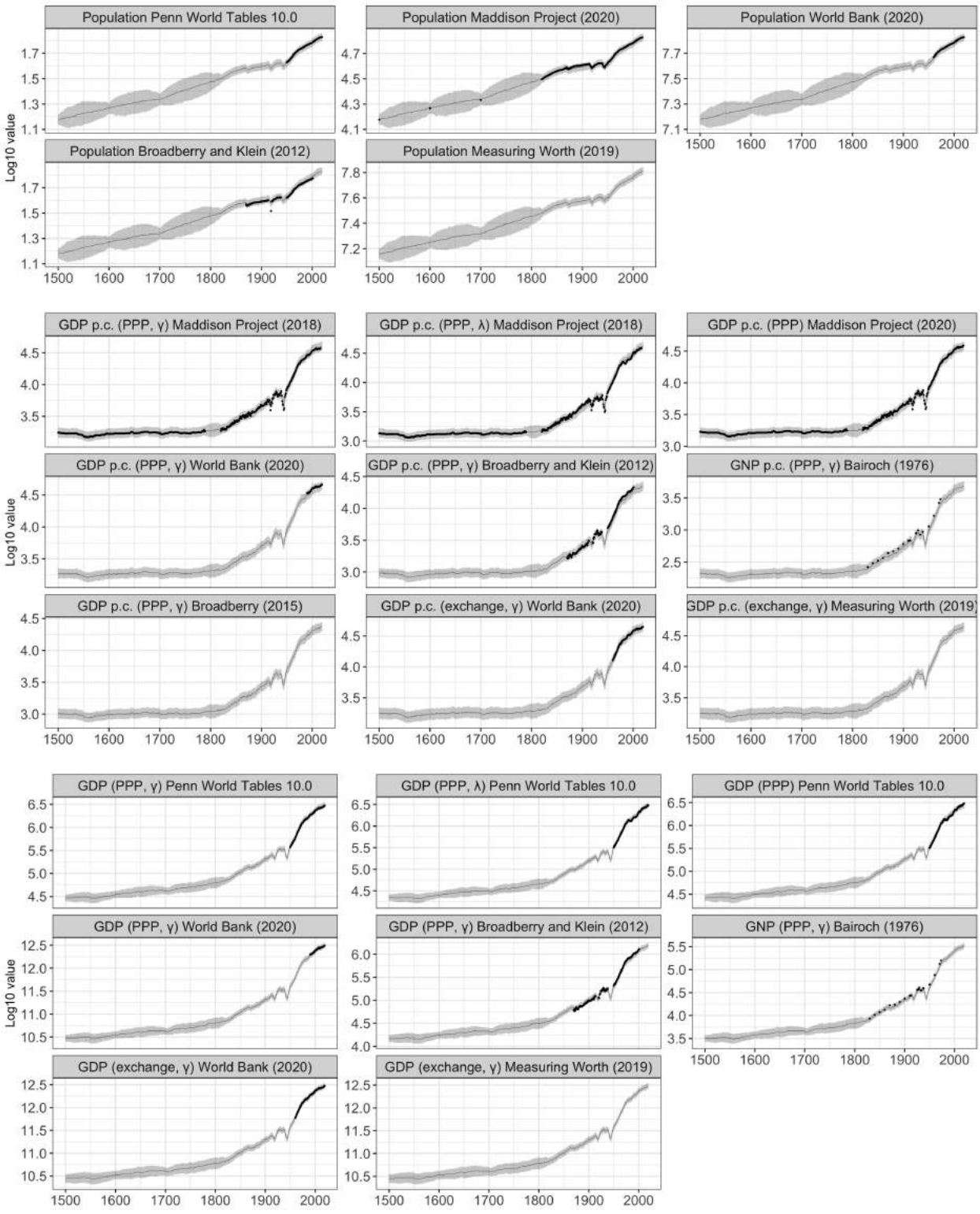

Figure 7: Posterior prediction intervals (grey lines) with ± 1 standard deviation confidence bands and observed variables (black points) for France.

C.5 Spain

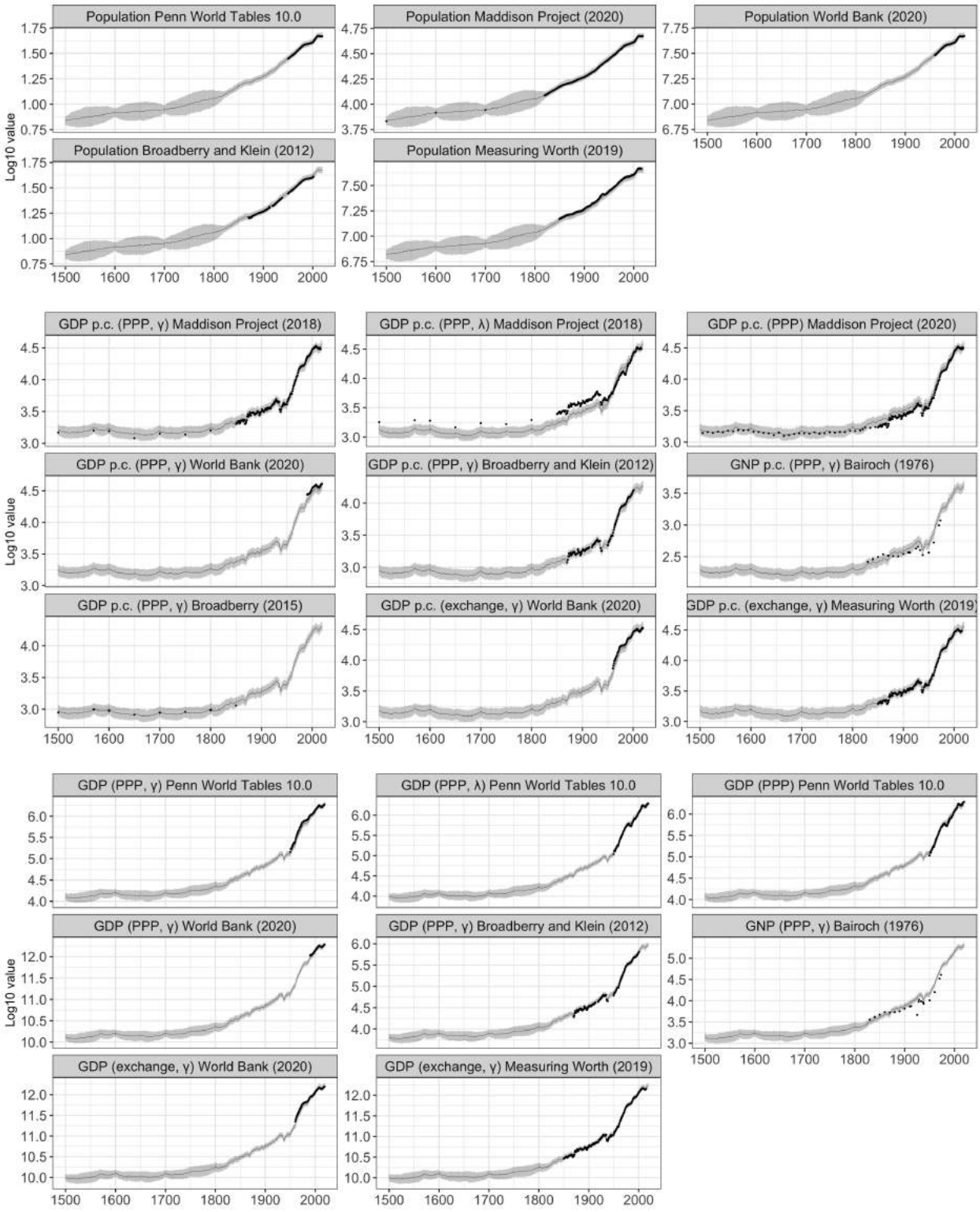

Figure 8: Posterior prediction intervals (grey lines) with ± 1 standard deviation confidence bands and observed variables (black points) for Spain.

C.6 Italy

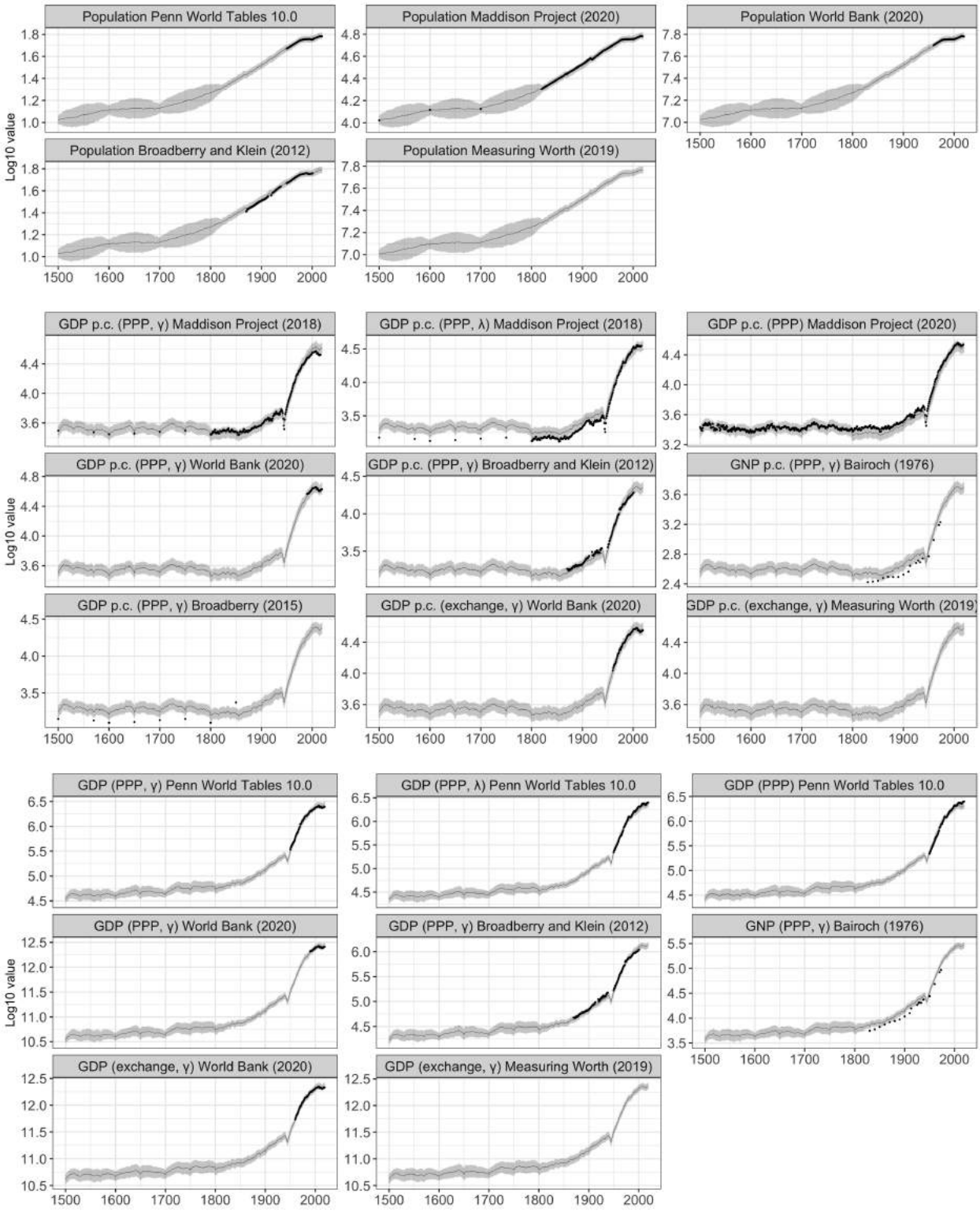

Figure 9: Posterior prediction intervals (grey lines) with ± 1 standard deviation confidence bands and observed variables (black points) for Italy.

C.7 Russia/Soviet Union

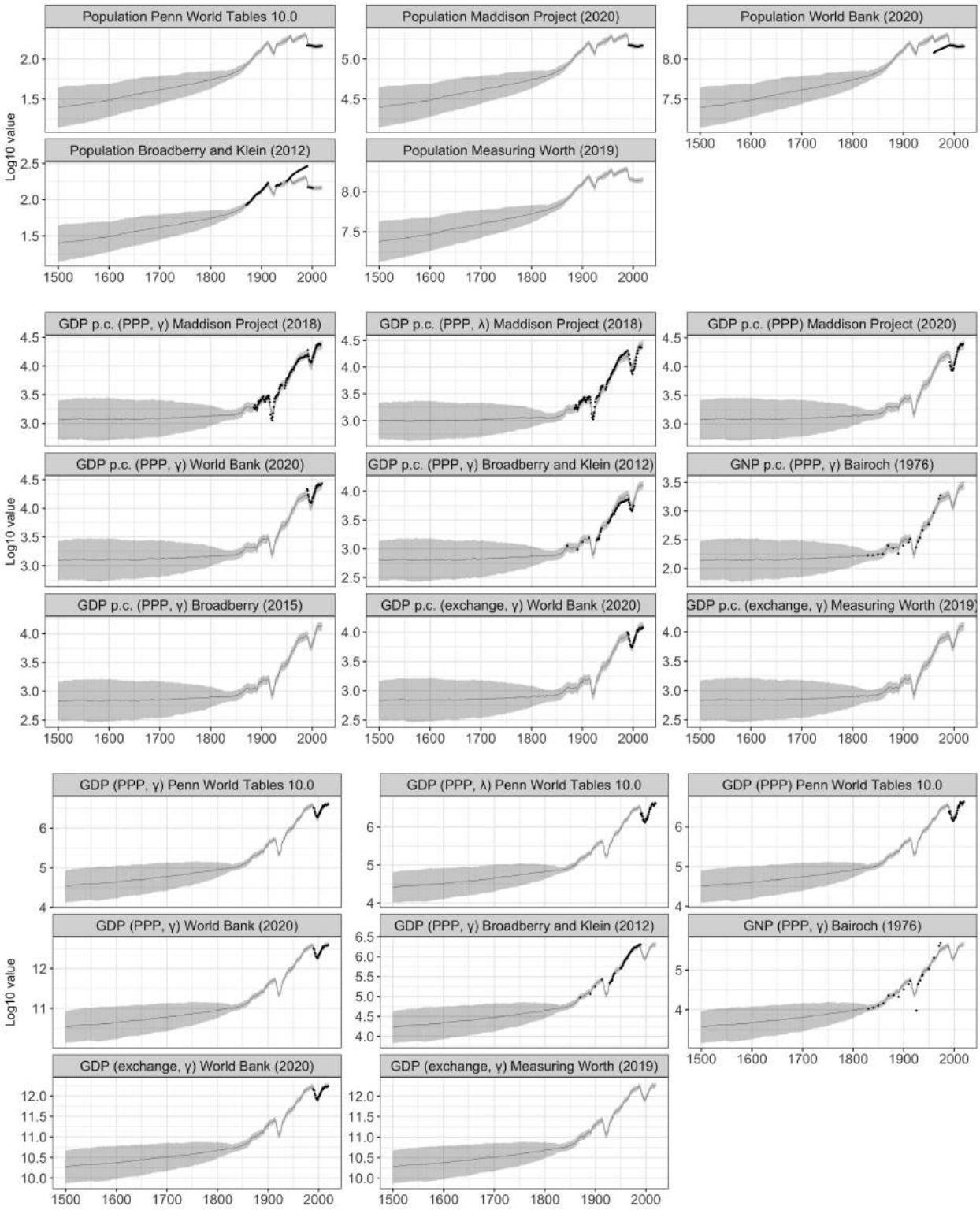

Figure 10: Posterior prediction intervals (grey lines) with ± 1 standard deviation confidence bands and observed variables (black points) for Russia.

C.8 Sweden

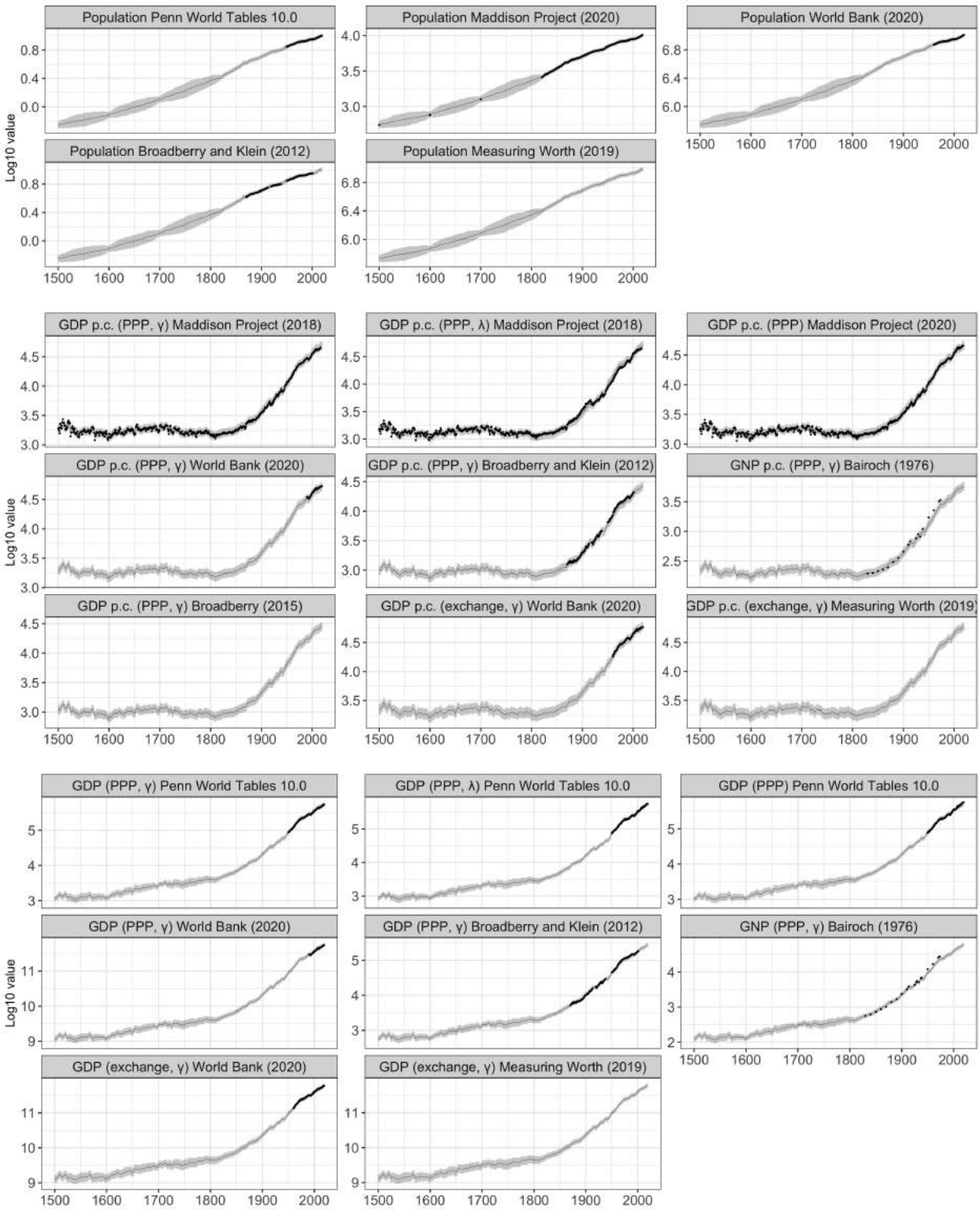

Figure 11: Posterior prediction intervals (grey lines) with ± 1 standard deviation confidence bands and observed variables (black points) for Sweden.

C.9 Albania

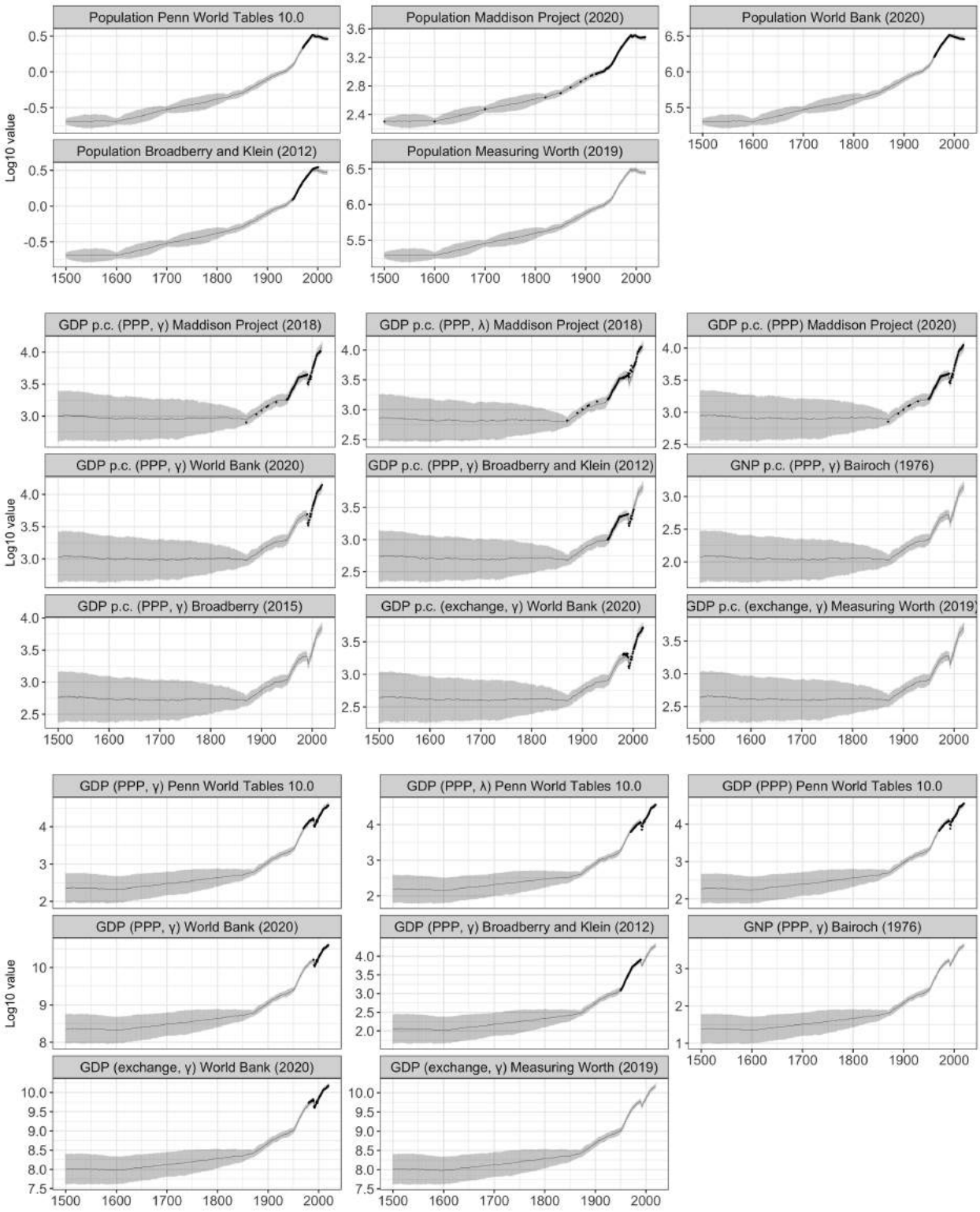

Figure 12: Posterior prediction intervals (grey lines) with ± 1 standard deviation confidence bands and observed variables (black points) for Albania.

C.10 India

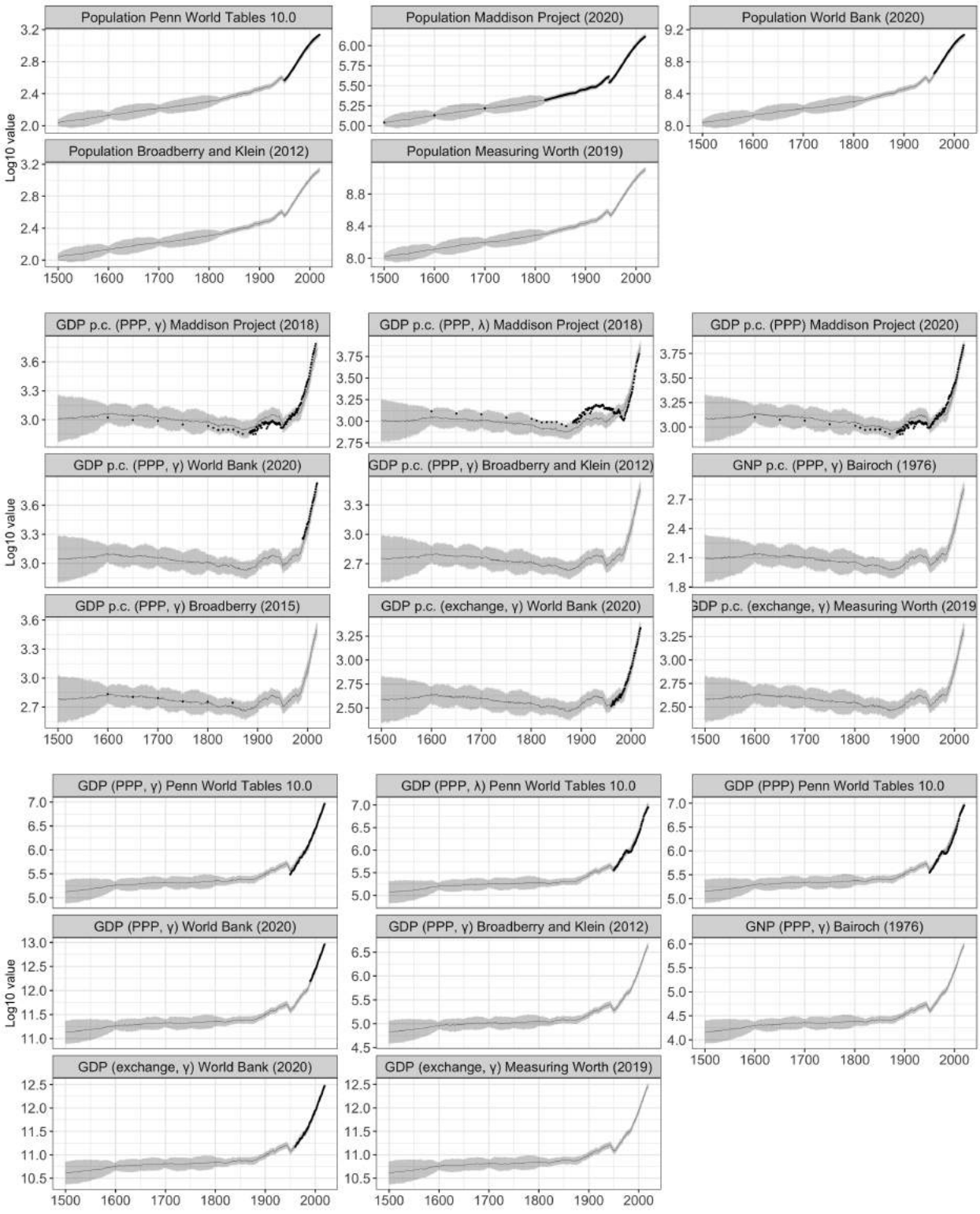

Figure 13: Posterior prediction intervals (grey lines) with ± 1 standard deviation confidence bands and observed variables (black points) for India.

C.11 Colombia

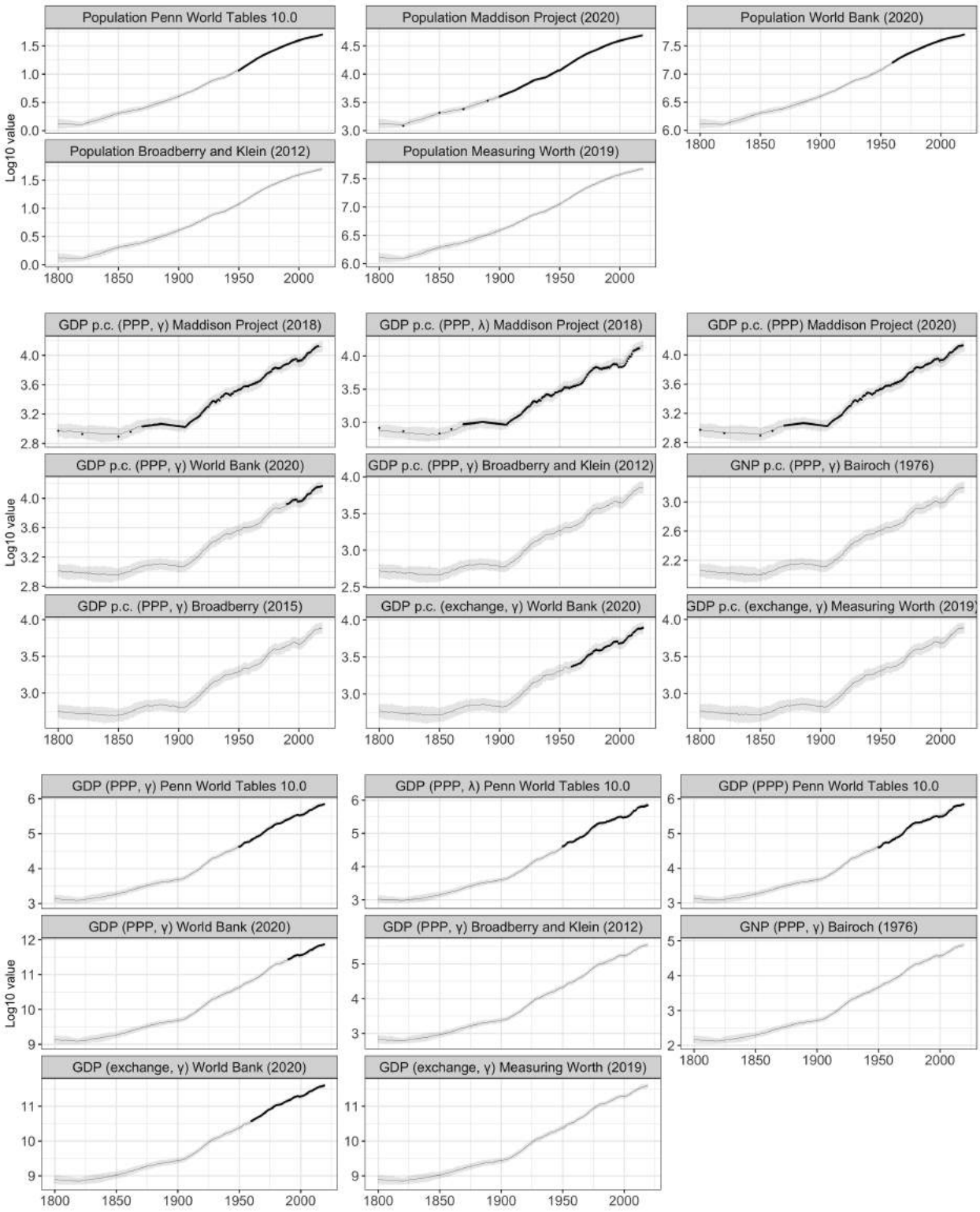

Figure 14: Posterior prediction intervals (grey lines) with ± 1 standard deviation confidence bands and observed variables (black points) for Colombia.

C.12 Brazil

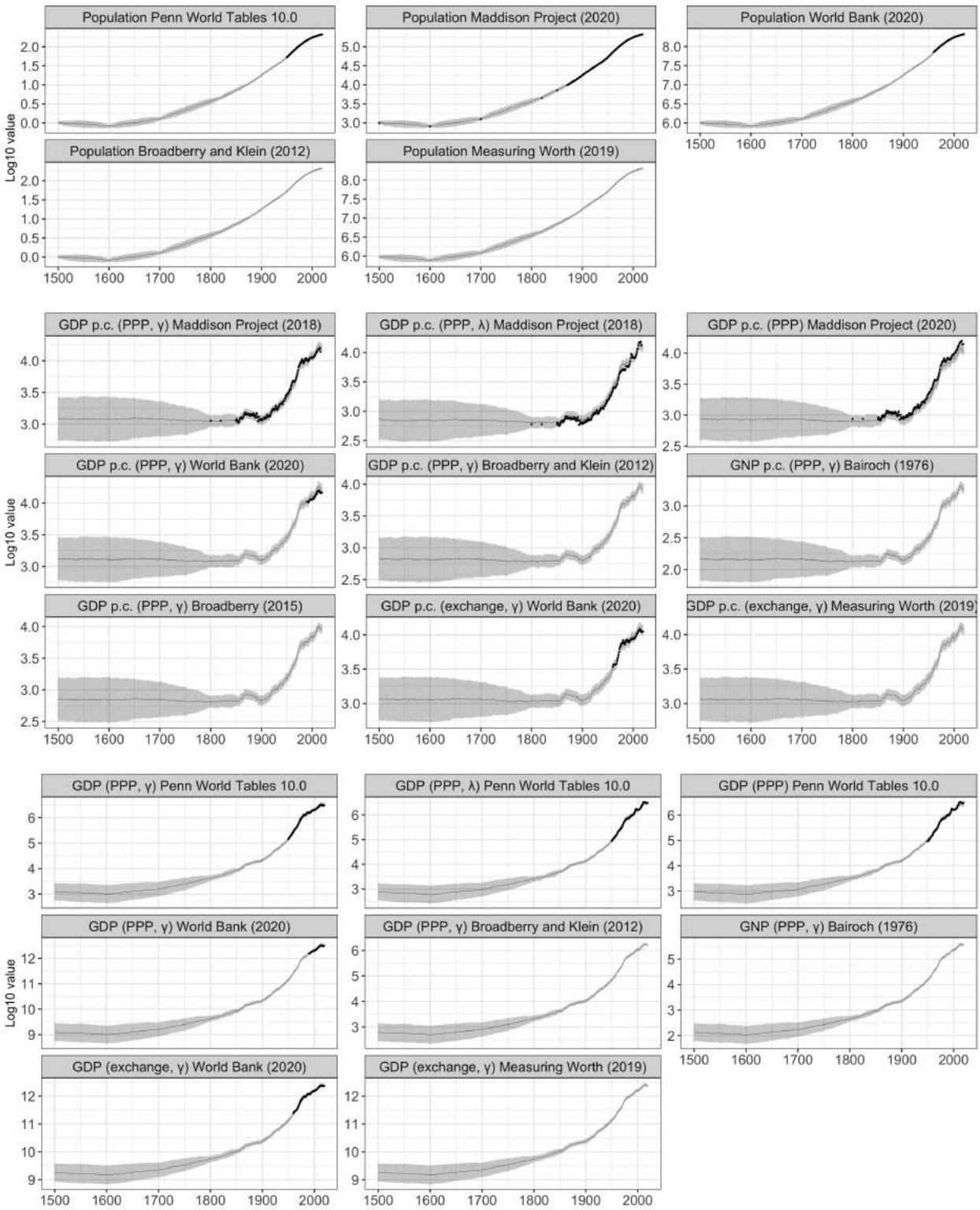

Figure 15: Posterior prediction intervals (grey lines) with ± 1 standard deviation confidence bands and observed variables (black points) for Brazil.

C.13 Democratic Republic of Congo/Zaire

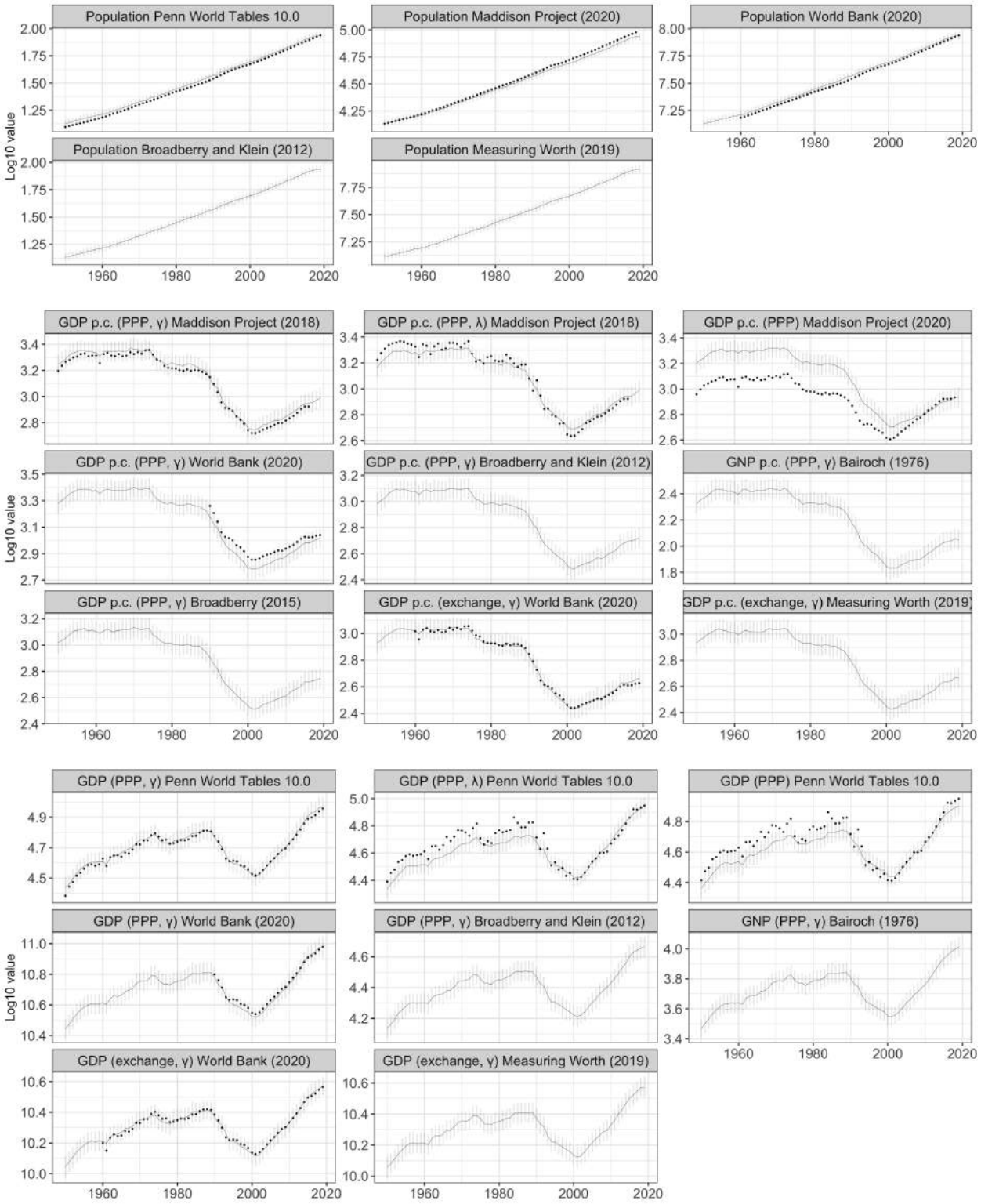

Figure 16: Posterior prediction intervals (grey lines) with ± 1 standard deviation confidence bands and observed variables (black points) for Democratic Republic of Congo/Zaire.

C.14 Uganda

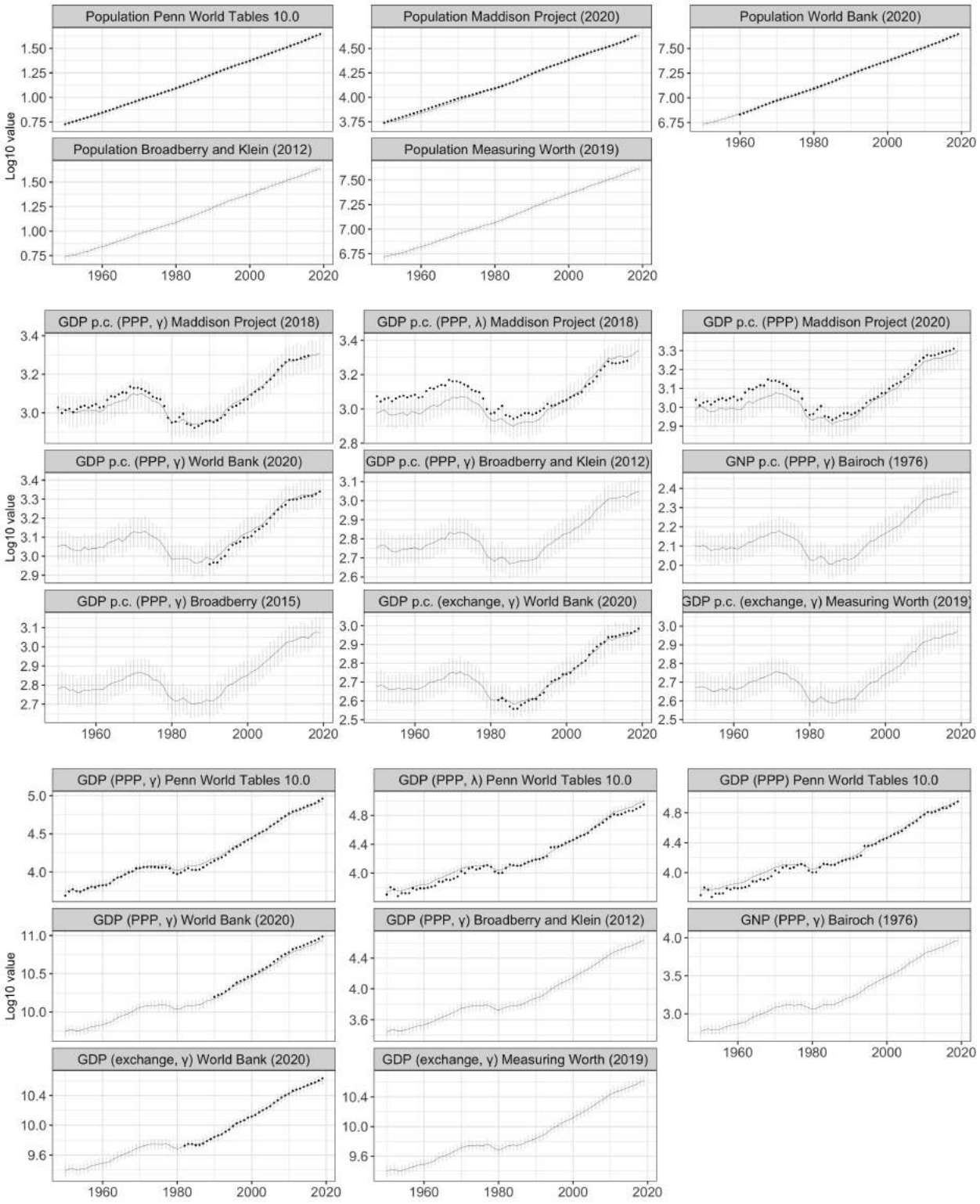

Figure 17: Posterior prediction intervals (grey lines) with ± 1 standard deviation confidence bands and observed variables (black points) for Uganda.

C.15 Iran

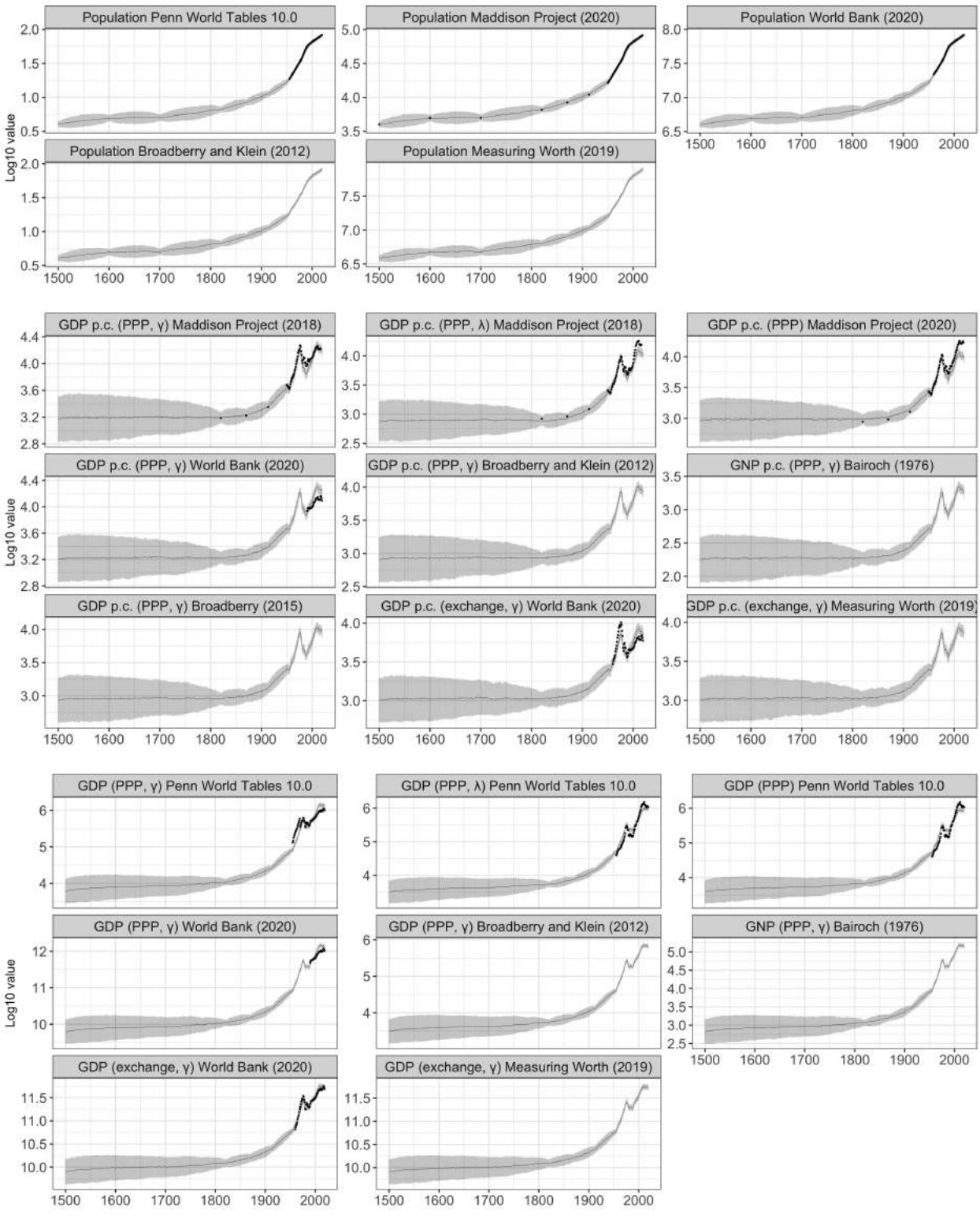

Figure 18: Posterior prediction intervals (grey lines) with ± 1 standard deviation confidence bands and observed variables (black points) for Iran.

C.16 Afghanistan

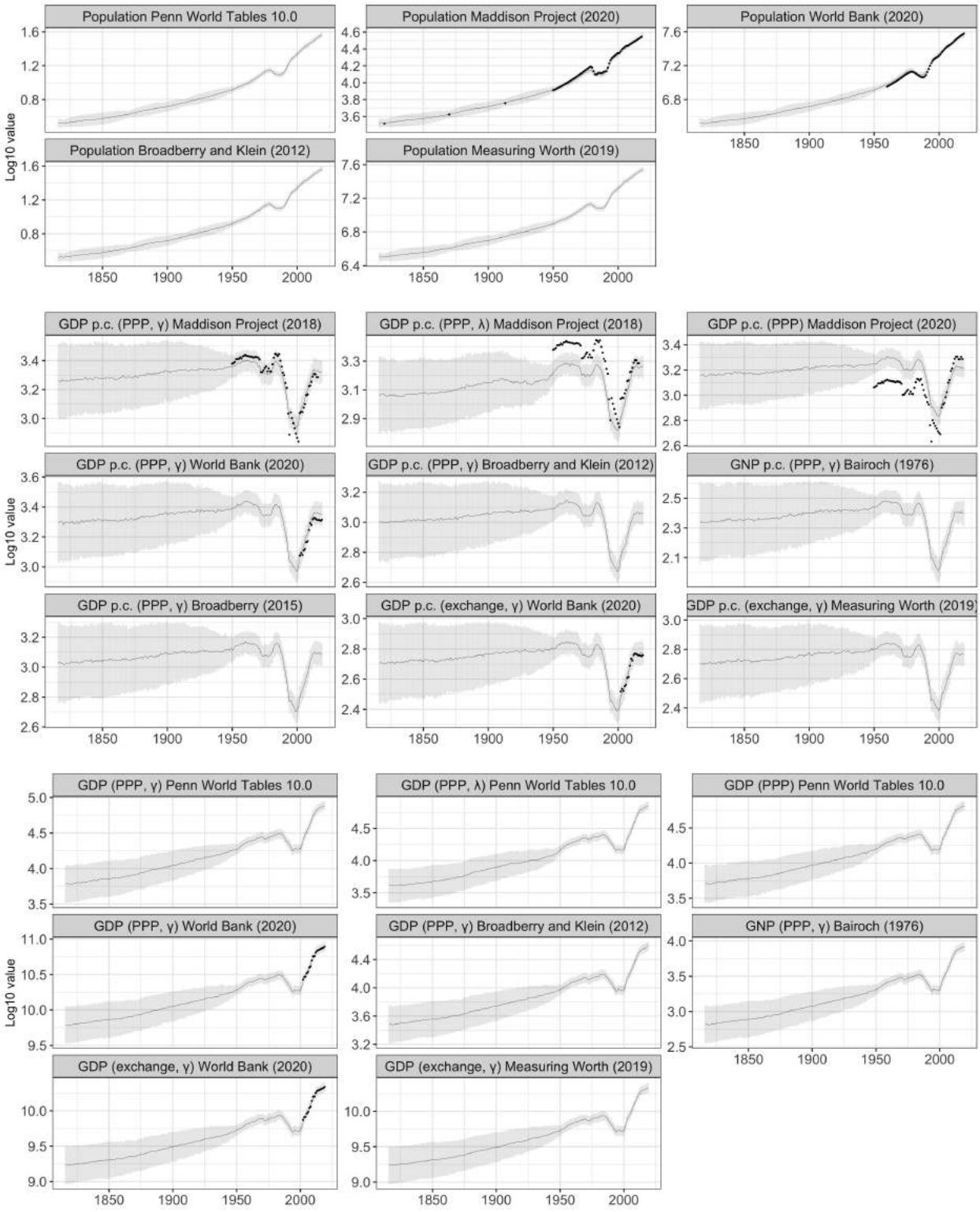

Figure 19: Posterior prediction intervals (grey lines) with ± 1 standard deviation confidence bands and observed variables (black points) for Afghanistan.

C.17 North Korea

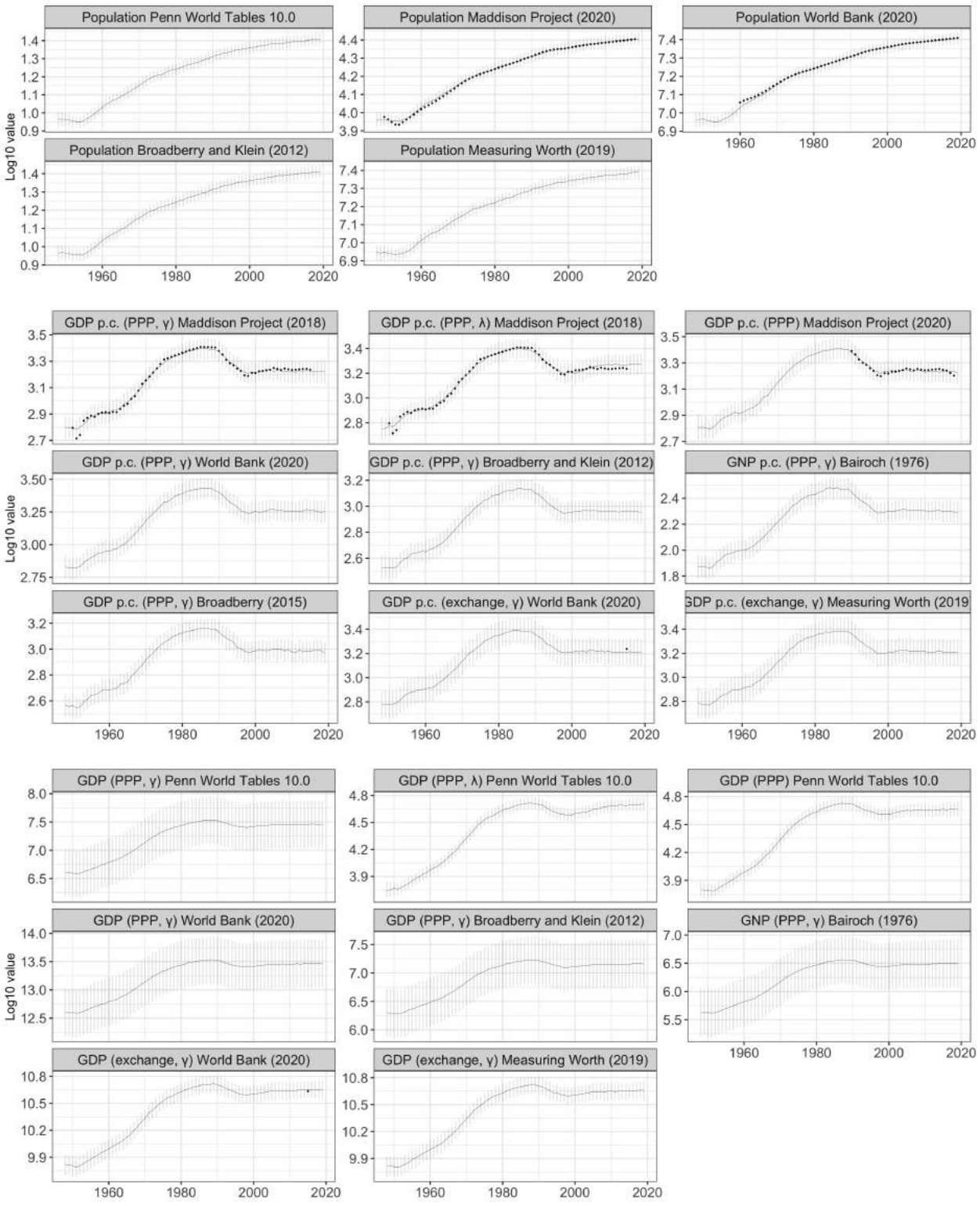

Figure 20: Posterior prediction intervals (grey lines) with ± 1 standard deviation confidence bands and observed variables (black points) for North Korea.

C.18 Pakistan

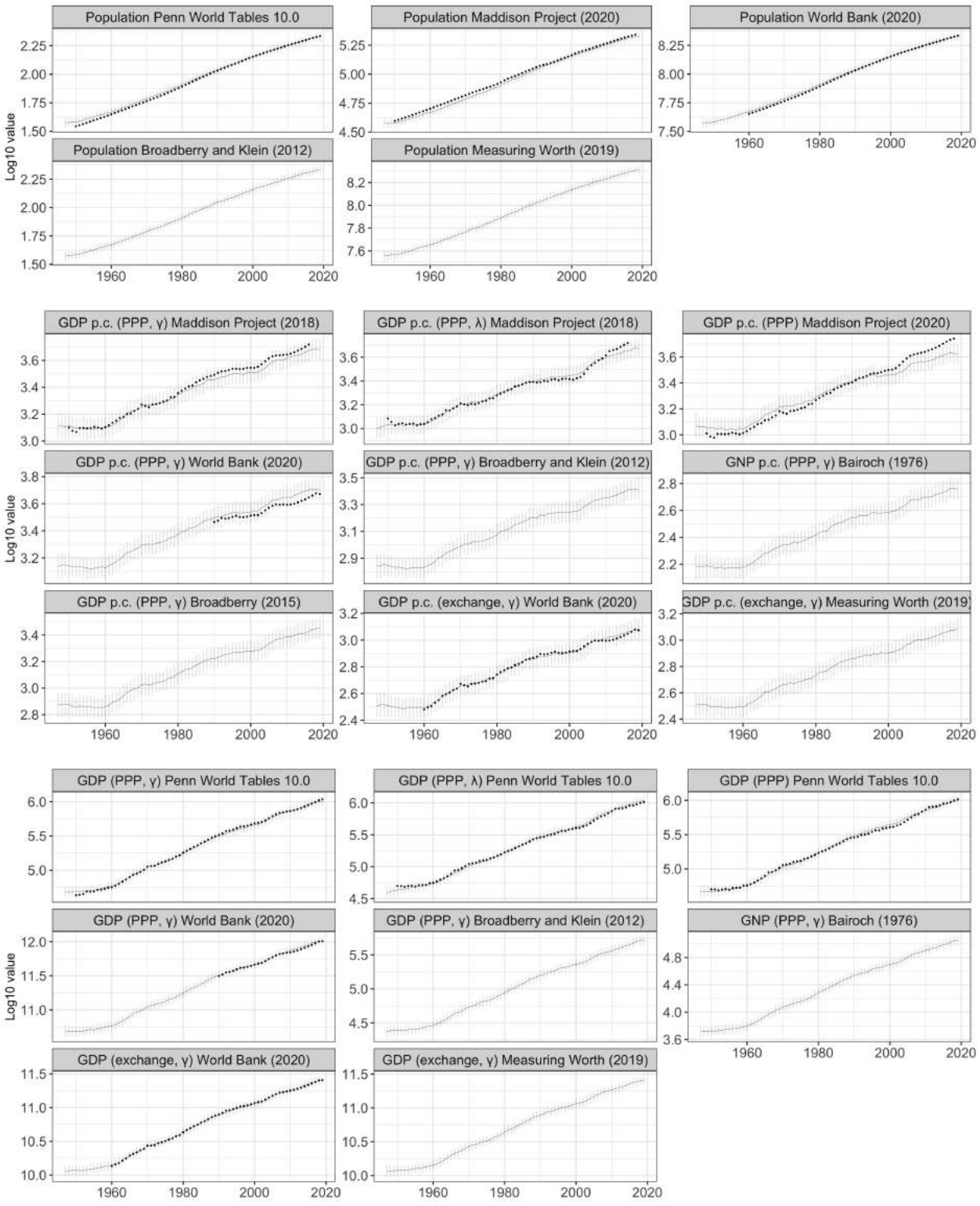

Figure 21: Posterior prediction intervals (grey lines) with ± 1 standard deviation confidence bands and observed variables (black points) for Pakistan.

C.19 Kosovo

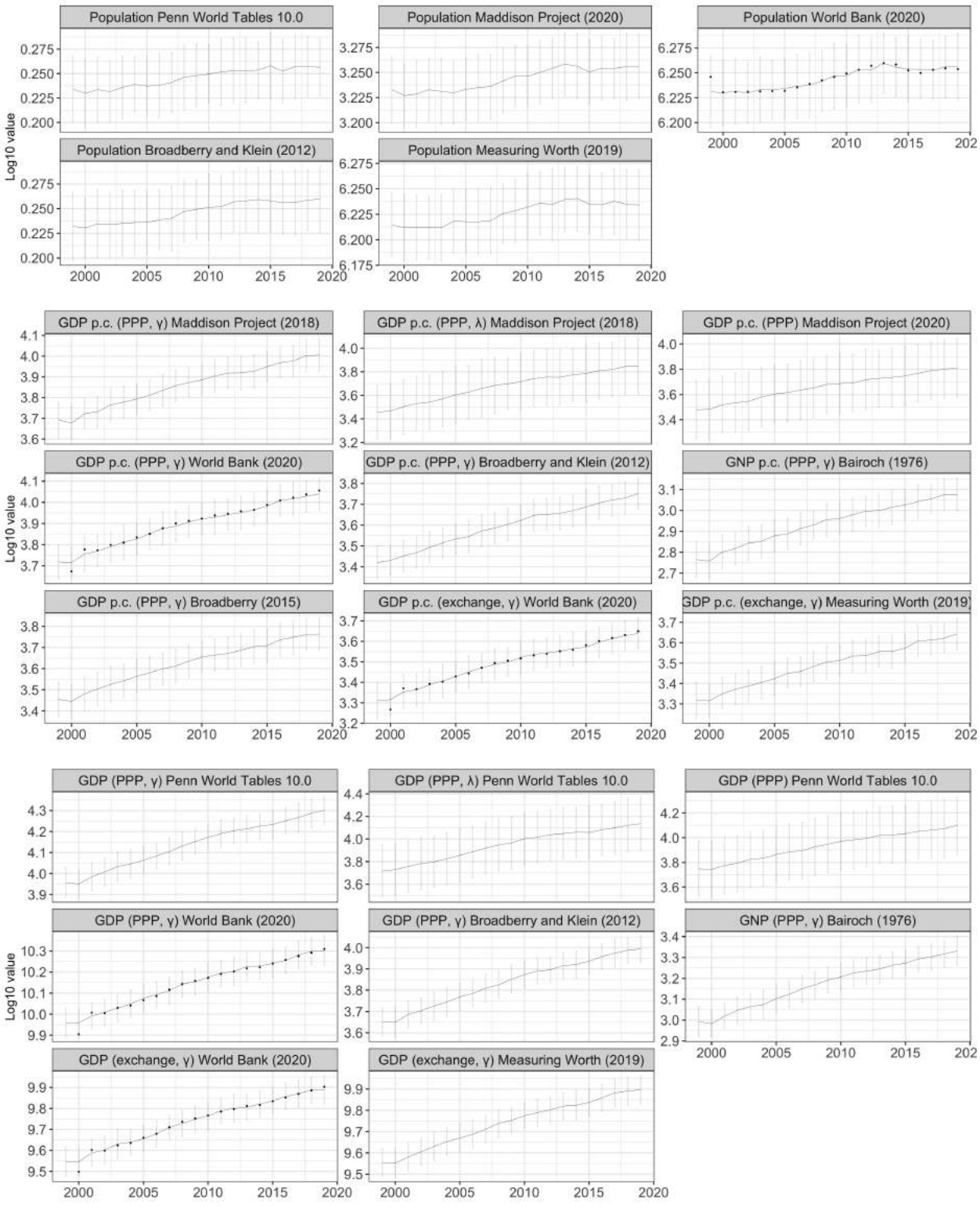

Figure 22: Posterior prediction intervals (grey lines) with ± 1 standard deviation confidence bands and observed variables (black points) for Kosovo.

C.20 East Timor

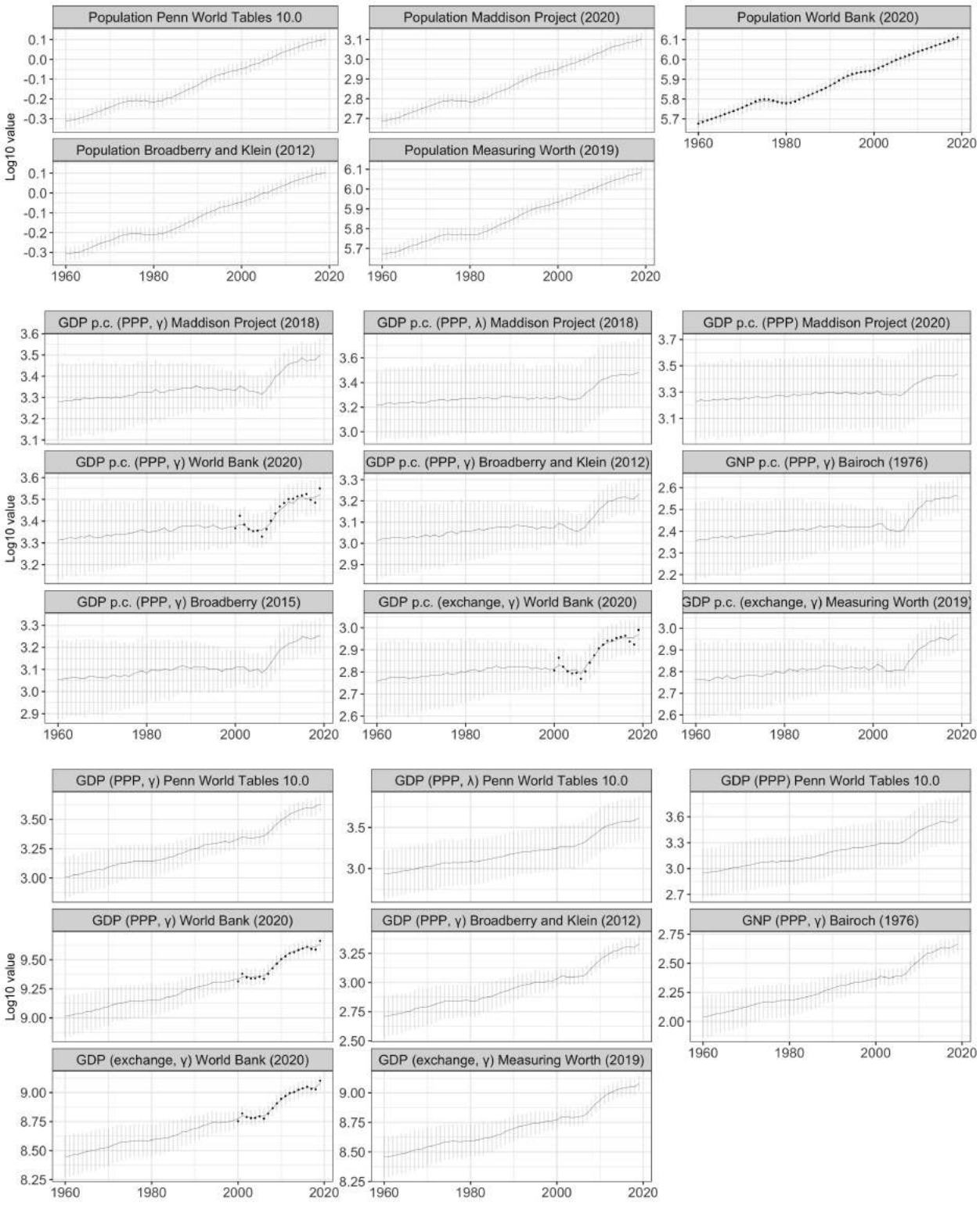

Figure 23: Posterior prediction intervals (grey lines) with ± 1 standard deviation confidence bands and observed variables (black points) for East Timor.

C.21 Eritrea

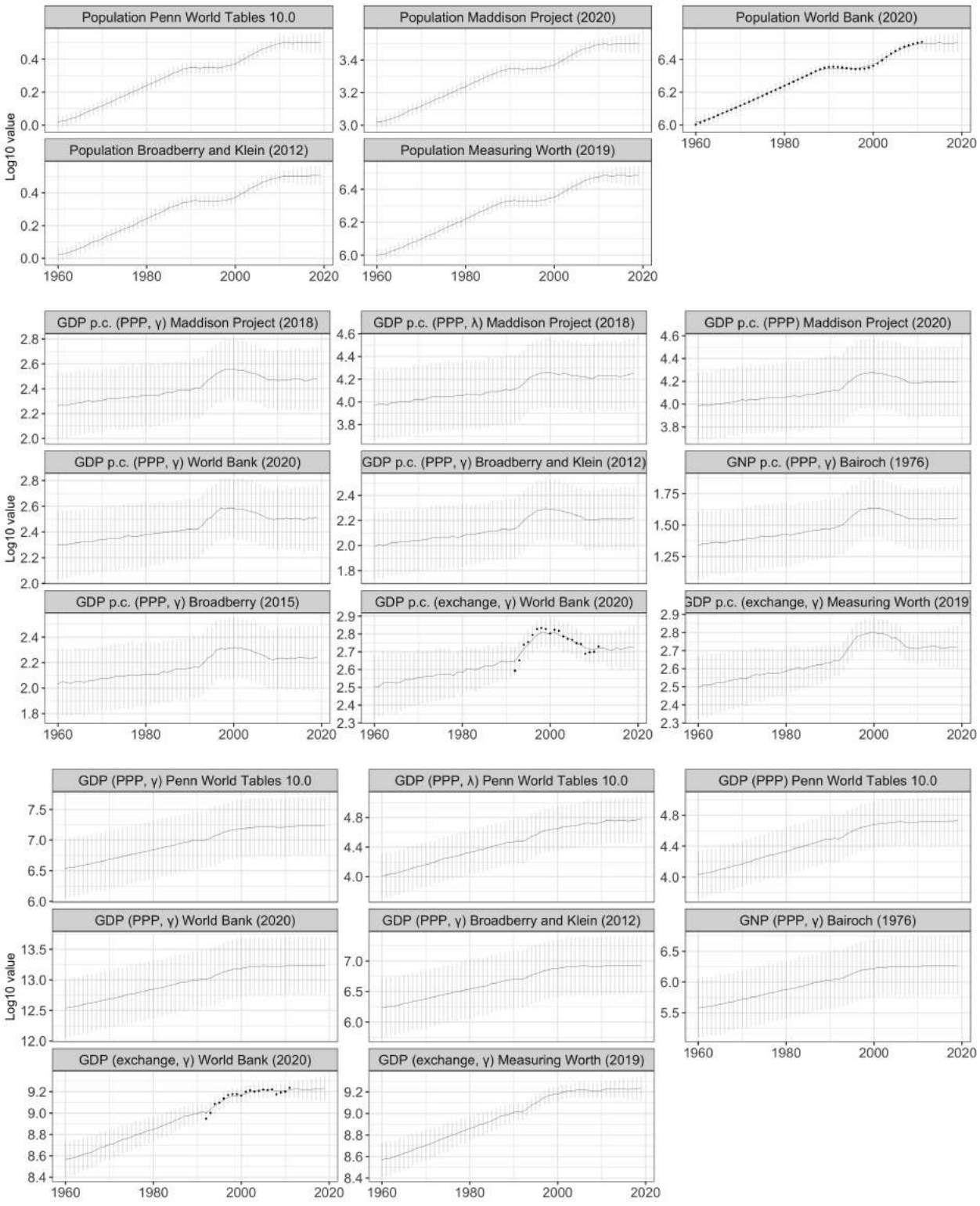

Figure 24: Posterior prediction intervals (grey lines) with ± 1 standard deviation confidence bands and observed variables (black points) for Eritrea.

C.22 Cross-sectional comparisons 1500

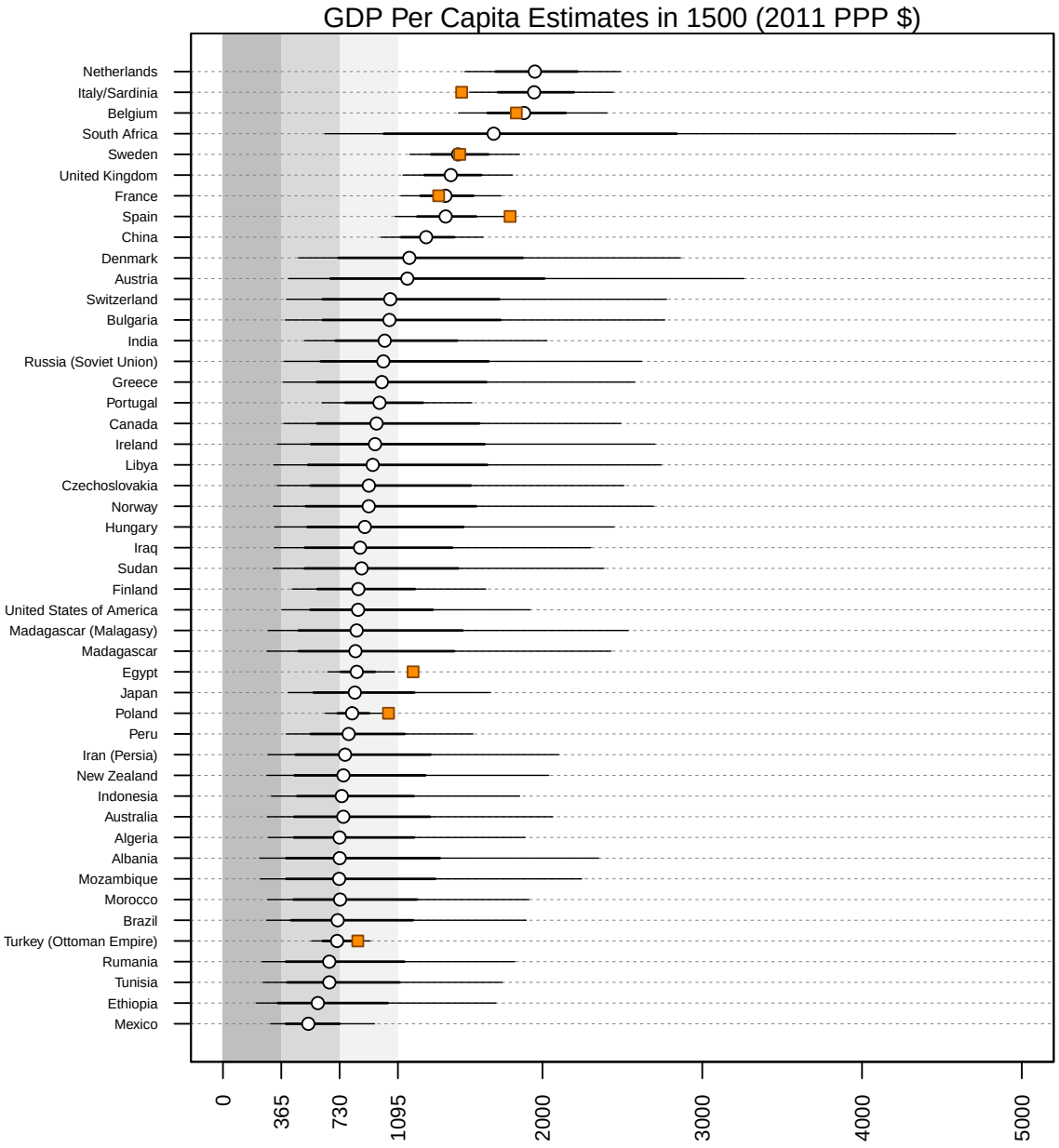

Figure 25: Posterior prediction intervals (lines) with ± 1 and ± 2 standard deviation confidence bands and posterior mean (circles) for GDP per capita for all available countries in 1500. Observed dataset values from the Maddison Project are in orange (for countries covered by Maddison) (Bolt and van Zanden, 2020). The shaded regions represent $1, $2, and $3 dollars a day, the minimum level of income necessary for everyday subsistence. (Anders, Fariss and Markowitz, 2020). Note that these values are in 2011 PPP international dollars.

C.23 Cross-sectional comparisons 1600

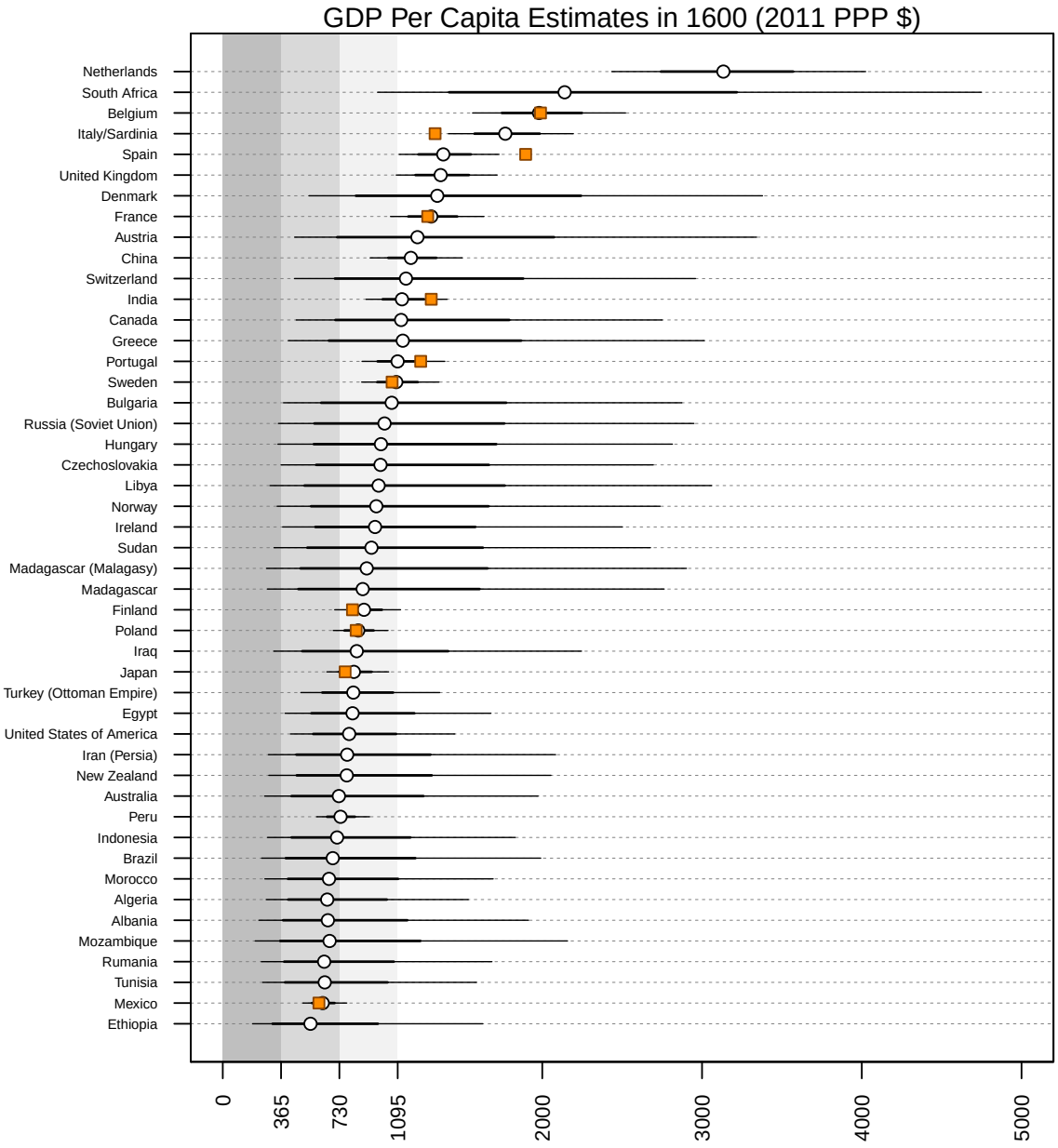

Figure 26: Posterior prediction intervals (lines) with ± 1 and ± 2 standard deviation confidence bands and posterior mean (circles) for GDP per capita for all available countries in 1600. Observed dataset values from the Maddison Project are in orange (for countries covered by Maddison) (Bolt and van Zanden, 2020). The shaded regions represent $1, $2, and $3 dollars a day, the minimum level of income necessary for everyday subsistence. (Anders, Fariss and Markowitz, 2020). Note that these values are in 2011 PPP international dollars.

C.24 Cross-sectional comparisons 1700

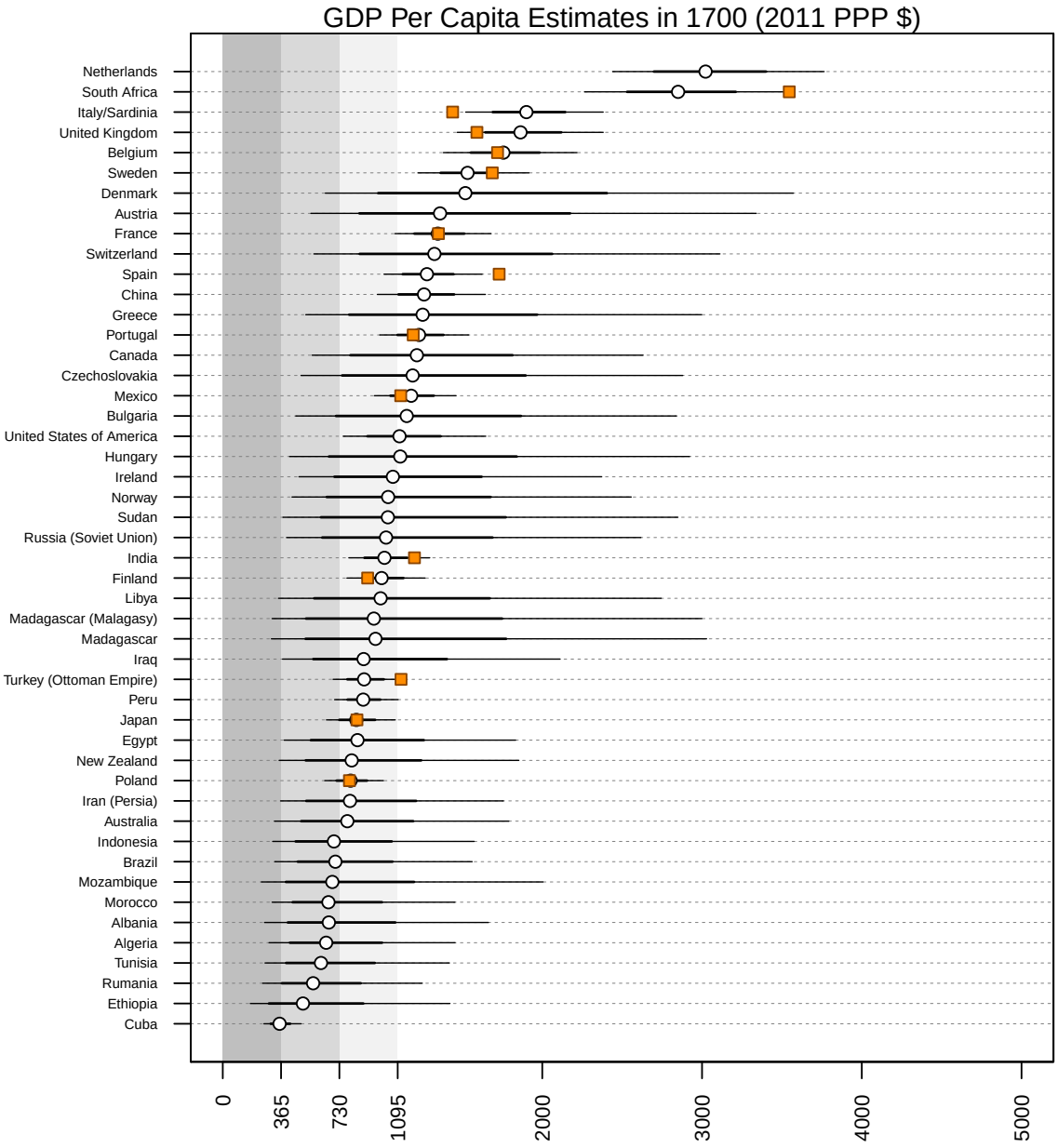

Figure 27: Posterior prediction intervals (lines) with ± 1 and ± 2 standard deviation confidence bands and posterior mean (circles) for GDP per capita for all available countries in 1700. Observed dataset values from the Maddison Project are in orange (for countries covered by Maddison) (Bolt and van Zanden, 2020). The shaded regions represent $1, $2, and $3 dollars a day, the minimum level of income necessary for everyday subsistence. (Anders, Fariss and Markowitz, 2020). Note that these values are in 2011 PPP international dollars.

C.25 Cross-sectional comparisons 1800

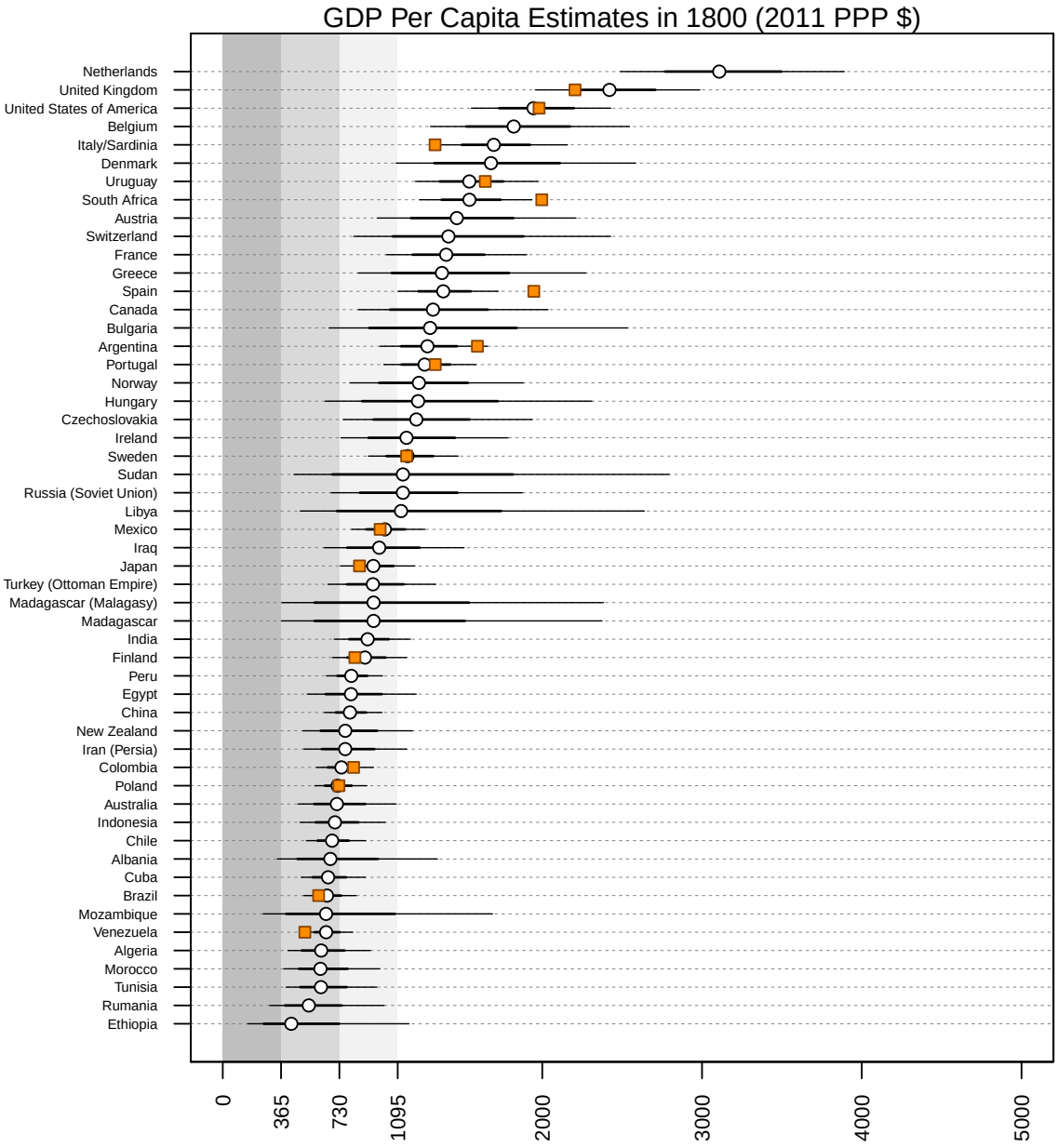

Figure 28: Posterior prediction intervals (lines) with ± 1 and ± 2 standard deviation confidence bands and posterior mean (circles) for GDP per capita for all available countries in 1800. Observed dataset values from the Maddison Project are in orange (for countries covered by Maddison) (Bolt and van Zanden, 2020). The shaded regions represent $1, $2, and $3 dollars a day, the minimum level of income necessary for everyday subsistence. (Anders, Fariss and Markowitz, 2020). Note that these values are in 2011 PPP international dollars.

C.26 Country-year variable examples from 1990

Zooming in on a few examples, Table 4 and Figure 29 display several example countries taken from 1990. We selected 1990 because this is the earliest year in which we have World Bank estimates for Constant International Dollar values and we wanted to show how our measure stacks up against both of the Maddison estimates and those of the World Bank for the countries where both have coverage. Recall that we have already demonstrated that in historical cases in which we only have estimates from the Maddison data project our estimates will be no different than the values from the Maddison Data Project. Therefore, we wanted to show how our estimates would stack up against cases in which we have estimates other than Maddison. For this set of hard cases that the reviewers are interested in, the World Bank is the only dataset other than those from the Maddison project that generates estimates across most of these cases using a unit of value similar enough to be easily comparable to those used by Maddison (i.e., constant PPP international dollars). However, in the interest of being transparent, we have provided a graphical illustration of the degree to which all values from existing datasets fall within the predicted range of our historical estimates for all of these hard cases over time, as well as for a dozen other countries in the Section above.

The values in the **original value** column come from one of three GDP per capita datasets respectively for each country in 1990 all in terms of PPP (purchasing power parity 2011 international dollars). The **mean estimate** is the average estimate from the distribution of country-year estimates produced by our measurement model. The $\pm$ **1 sd** is the 68% highest probability density of the distribution of country-year estimates. The **unit Z-score**'s value represents the positions of the observed variable's value relative to the center of the the posterior predicted interval for each of the country-year unit and for every dataset variable. A value of 0 indicates that the observed variable's value falls directly at the center of the posterior predicted interval. Z-score units above and below 0 represent standard deviation differences from the center value of the posterior predicted interval.[3] As the table reveals, observed dataset values from both the Maddison data project and the World Bank reside close to the center of the range of estimates that our model generates.

[3]As described in the main article, the country-year-variable Z-score values take the form: $z_{itj} = \frac{y_{itj} - E(\tilde{y}_{itj})}{\sigma_{\tilde{y}_{itj}}}$, where y_{itj} is the observed value for the country-year-variable, $E(\tilde{y}_{itj})$ is the expected value or mean for the posterior predicted interval, and $\sigma_{\tilde{y}_{itj}}$ is the standard deviation for the posterior predicted interval.

To provide some illustration as to why it is substantively meaningful that the observed values fall within one standard deviation of our estimates, it is helpful to consider what that would mean for a real-world case. For this illustration we use the case of Vietnam in Figure 29 and also in Table 4. Both of Maddison's estimates and those from the World Bank fall within the ranges predicted by our model and are actually relatively close to the point estimates as well.

	country	original	our estimate	± 1 sd	unit Z-score
1	Afganistan	1992.00	1797.34	[1470.2, 2124.48]	-0.66
2	Afganistan	1992.00	1356.62	[1123.96, 1589.28]	-2.37
3	Afganistan	963.00	1413.81	[1143.78, 1683.84]	1.82
4	Afganistan		1895.74	[1568.27, 2223.21]	
5	Albania	4454.00	4179.19	[3414.81, 4943.57]	-0.44
6	Albania	4099.00	3597.48	[2997.9, 4197.07]	-0.87
7	Albania	3983.00	3651.44	[3021.73, 4281.15]	-0.60
8	Albania	4927.88	4493.55	[3729.57, 5257.53]	-0.63
9	Democratic Republic of Congo	1407.00	1378.25	[1139.99, 1616.51]	-0.20
10	Democratic Republic of Congo	1197.00	1242.80	[1037.85, 1447.76]	0.15
11	Democratic Republic of Congo	813.00	1271.43	[1076.19, 1466.67]	2.90
12	Democratic Republic of Congo	1819.72	1521.13	[1254.65, 1787.6]	-1.13
13	Iran	9816.00	7664.60	[6351.12, 8978.07]	-1.54
14	Iran	5129.00	4568.11	[3737.84, 5398.39]	-0.74
15	Iran	5620.00	4626.39	[3860.35, 5392.43]	-1.27
16	Iran	8649.58	8190.16	[6784.27, 9596.05]	-0.40
17	North Korea	2389.00	2398.44	[1985.21, 2811.67]	-0.06
18	North Korea	2389.00	2346.03	[1943.67, 2748.38]	-0.19
19	North Korea	2455.82	2424.72	[2043.38, 2806.05]	-0.16
20	North Korea		2540.42	[2086.16, 2994.68]	
21	Pakistan	3108.00	2906.35	[2444.49, 3368.2]	-0.50
22	Pakistan	2472.00	2540.95	[2106.16, 2975.74]	0.08
23	Pakistan	2531.00	2647.78	[2191.25, 3104.3]	0.18
24	Pakistan	2915.90	3187.32	[2645.57, 3729.08]	0.44
25	Uganda	908.00	920.12	[761.87, 1078.37]	-0.01
26	Uganda	946.00	859.19	[701.68, 1016.69]	-0.62
27	Uganda	932.00	870.03	[714.97, 1025.1]	-0.48
28	Uganda	907.76	986.05	[807.26, 1164.84]	0.37
29	Vietnam	1512.00	1536.12	[1272.09, 1800.14]	0.01
30	Vietnam	1200.00	1345.97	[1128.58, 1563.37]	0.63
31	Vietnam	1634.00	1434.47	[1174.7, 1694.24]	-0.81
32	Vietnam	1673.25	1673.72	[1383.15, 1964.29]	-0.08

Table 4: 1990 dataset values for several examples countries. The dataset values are from three distinct GDP per capita indicators, with the three two from the Maddison data project and the third from the World Bank. Each dataset presents the original values in $2011 international dollars, which are adjusted for purchasing power parity (PPP) to make cross-sectional comparisons possible.

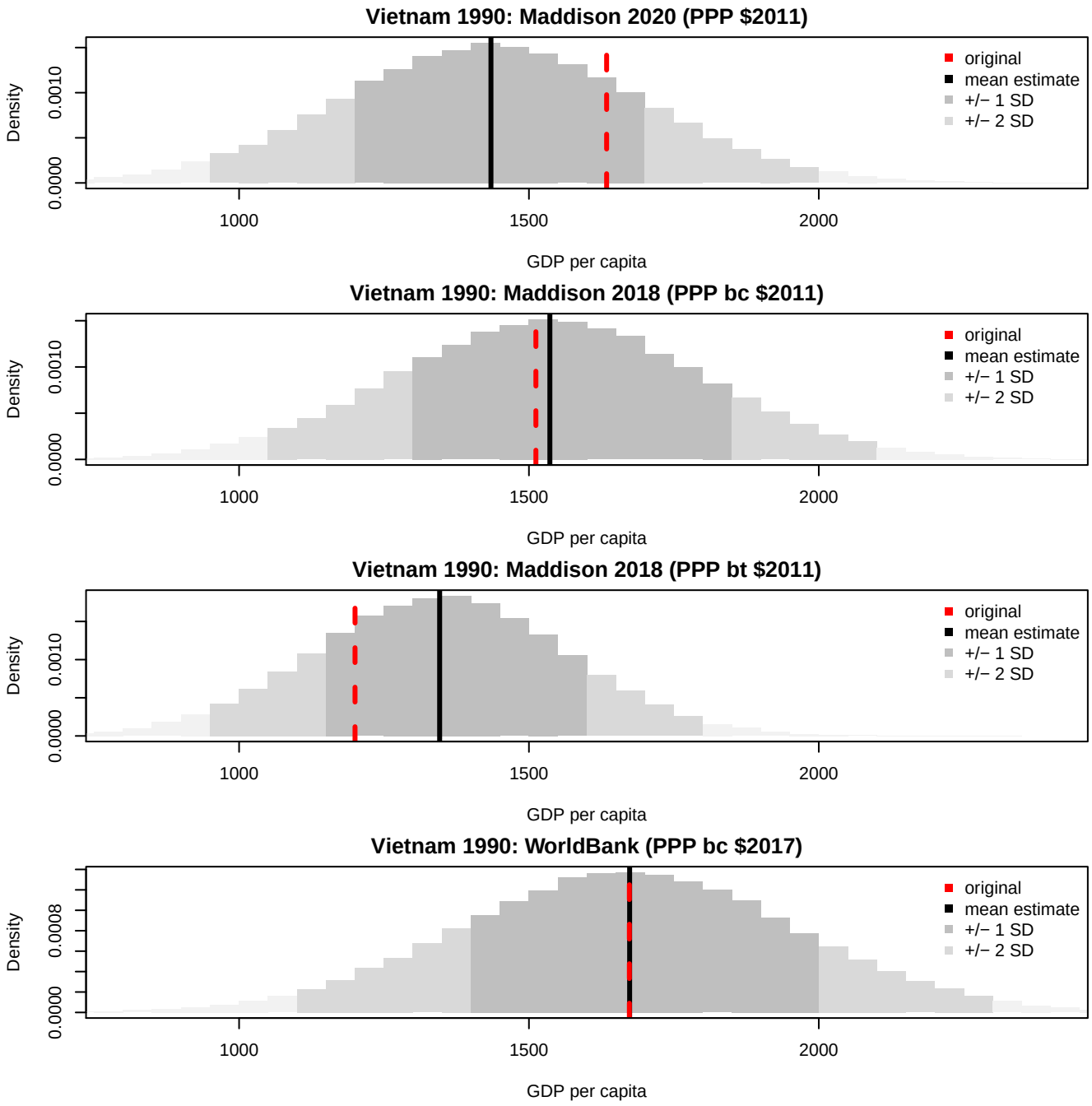

Figure 29: Vietnam 1990 dataset values plotted along with the range of estimates for these three country-year cases derived from the measurement model. The dataset values are from four distinct GDP per capita indicators, the first three from the Maddison data project and the third from the World Bank. Each dataset presents the original values in $2011 international dollars, which are adjusted for purchasing power parity (PPP) to make cross-sectional comparisons possible.

D New Results from Expanded Measurement Coverage

Here, we build on correlation results presented in the main article. The figures below plot Spearman rank-order correlation coefficients for the relationship between economic development and measures of health and wellbeing (i.e., child mortality, literacy rates, life expectancy), measures of democracy (i.e., V-DEM Polyarchy and Polity IV), measures of repression, measures of power projection capabilities (i.e., capital ships and naval tonnage), and measures of conflict (i.e., MIDs and COW wars, combat fatalities), for one-year, ten-year, and 50-year rolling temporal windows. Each series of temporal correlations are found in the upper row of each figure. The lower row of each figure (except for measures of conflict) graphs the distribution of Spearman rank-order correlation between the latent GDP per capita variable and a binary variable that indicates whether or not the other variable is observed for the country-year unit. This second set of graphs captures whether wealth is correlated with the measurement coverage of the specific variable. In summary, we show that the relationship between wealth and the different variables we consider changes through the transitional period of the industrial revolution. Many of the relationships are much more volatile in the 19th century but stabilize in the early to mid 20th century.

To estimate the correlations we employ a simple procedure to incorporate uncertainty from the country-year distributions for the GDP per capita variable. We provide additional discussion about incorporating uncertainty into regression analysis below. Specifically, we measure economic development by taking $m = 100$ draws from the posterior distribution of our latent GDP per capita variable and then correlating these draws with another observed variable. This procedure allows us to incorporate uncertainty from the distribution of GDP per capita when estimating the correlation coefficient for each time-period (with 95%-intervals generated from the distribution of 100 correlation coefficients). We can also take draws from the country-year distribution of the second variable for those datasets that are also measured with uncertainty (e.g., the VDEM Polyarchy or VDEM killing variable).

D.1 Measures of Health and Wellbeing

D.1.1 Correlations between GDP per capita and Life Expectancy, 1700 – Present

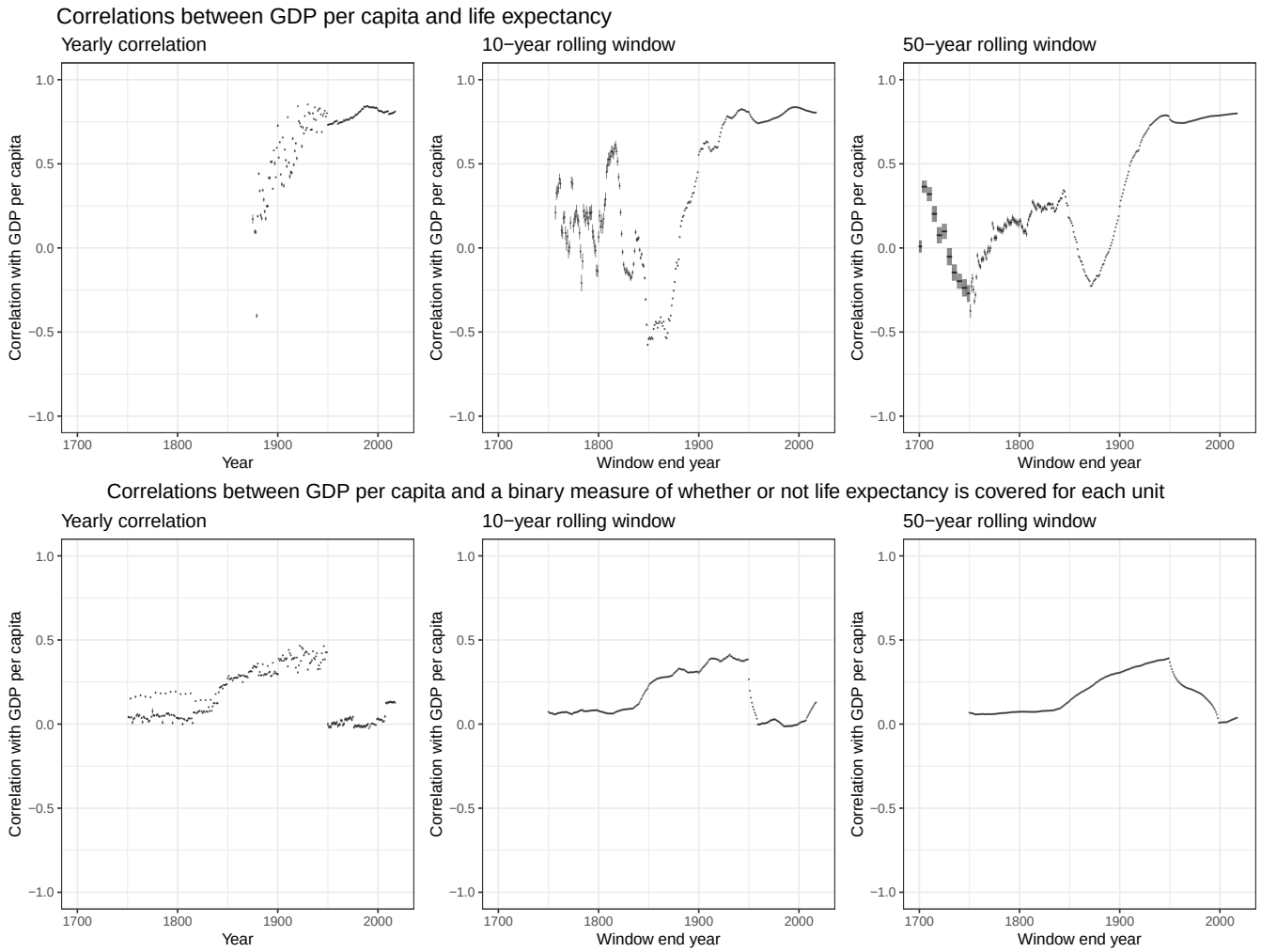

Figure 30: Distribution of Spearman rank-order correlation between the latent GDP per capita variable and a measure of life expectancy. Life expectancy is measured using data from Max Roser and Ritchie (2013). Correlations are calculated for each 1-year period (left column), 10-year period (middle column), and 50-year period (right column). The second row of panels is the distribution of Spearman rank-order correlation between the latent GDP per capita variable and a variable measuring if the life expectancy variable is observed for the country-year unit. Note that we only estimate a correlation coefficient for year-periods with at least 10 cases, which means that some of the 1-year correlations begin later than the 10-year or 50-year correlations. The correlation between life expectancy and GDP per capita has dramatically increased since governments first began tracking this variable. The build up of the industrial revolution appears to temporarily weaken this relationship for several decades in the mid to late 19th century.

D.1.2 Correlations between GDP per capita and Child Mortality, 1800 – Present

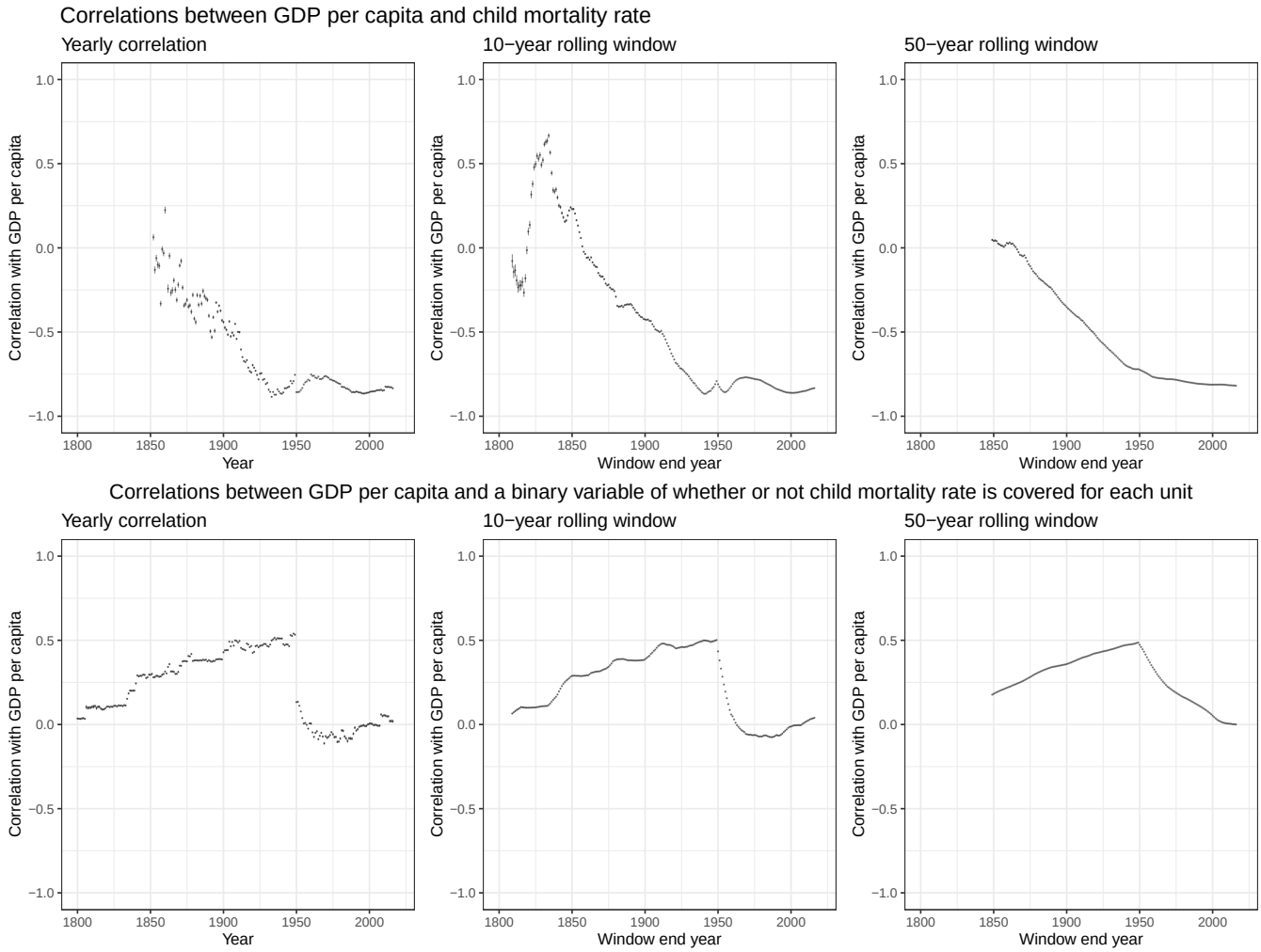

Figure 31: Distribution of Spearman rank-order correlation between the latent GDP per capita variable and a measure of child mortality. Child mortality is measured using data from Max Roser and Dadonaite (2013). Correlations are calculated for each 1-year period (left column), 10-year period (middle column), and 50-year period (right column). The year listed on the x-axis is the start year for each 10-year period (middle column), and 50-year period (right column). The number of country-year units with an observed value for the child mortality variable is quite low in the early years of the dataset. This is why the 1-year period and even the 10-year period correlations fluctuate in the early period of the time-series. Once there is sufficient observed data starting in the mid-1800s, the correlations based on a single year of data or years of data begin to stabilize. The second row of panel is the distribution of Spearman rank-order correlation between the latent GDP per capita variable and a variable measuring if the child mortality variable is observed for the country-year unit. Note that we only estimate a correlation coefficient for year-periods with at least 10 cases, which means that some of the 1-year correlations begin later than the 10-year or 50-year correlations. As with life expectancy, the negative correlation between child mortality and GDP per capita (more wealth, less death) has dramatically decreased since governments first started to track this variable. Again though, the build up of the industrial revolution appears to weaken this relationship for several decades in the mid 19th century.

D.1.3 Correlations between GDP per capita and Literacy Rates, 1500 – Present

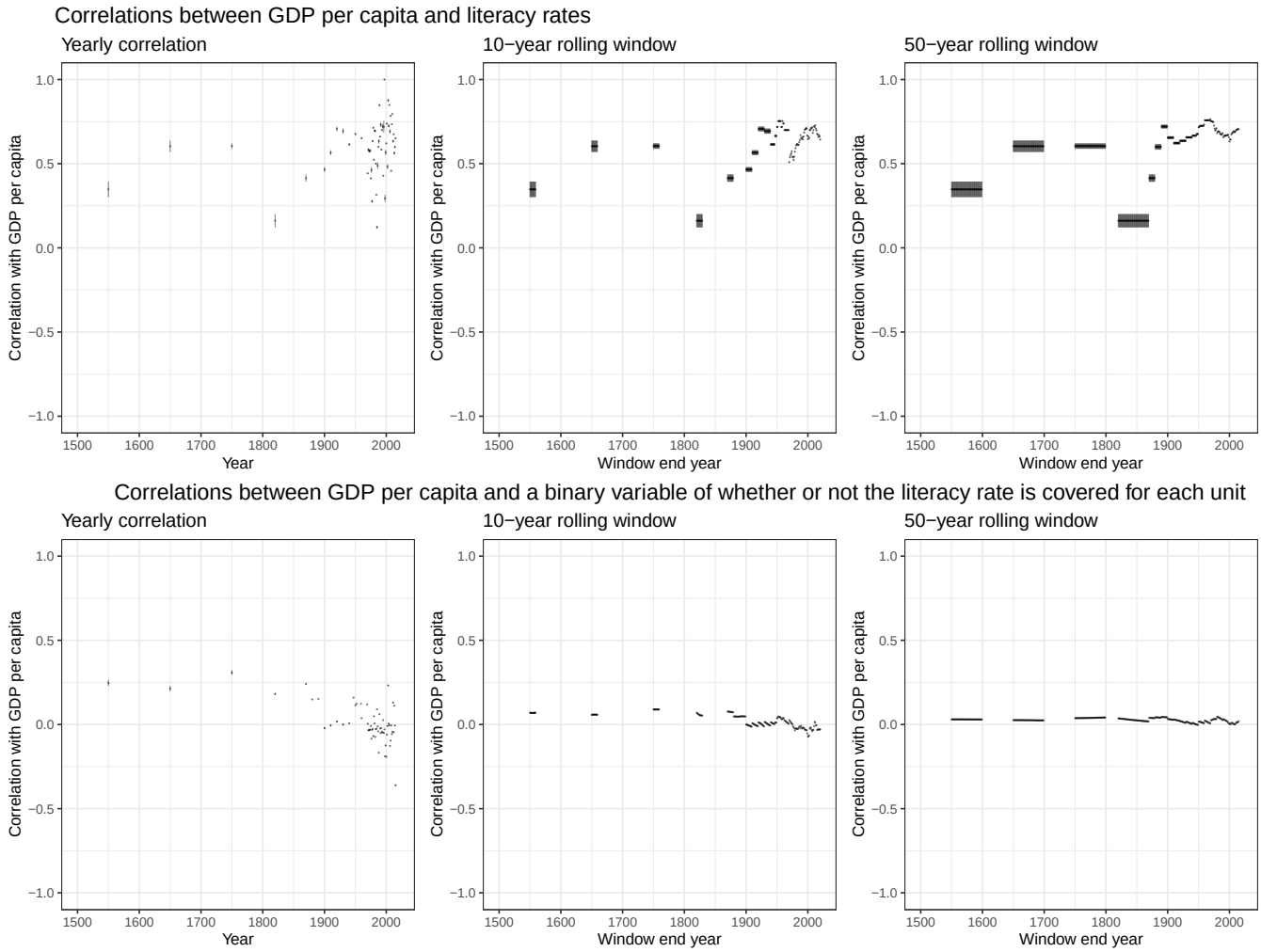

Figure 32: Distribution of Spearman rank-order correlation between the latent GDP per capita variable and a measure of literacy. Literacy rates are measured using data from Roser and Ortiz-Ospina (2016). Correlations are calculated for each 1-year period (left column), 10-year period (middle column), and 50-year period (right column). The second row of panels is the distribution of Spearman rank-order correlation between the latent GDP per capita variable and a variable measuring if the literacy variable is observed for the country-year unit. Note that we only estimate a correlation coefficient for year-periods with at least 10 cases, which means that some of the 1-year correlations begin later than the 10-year or 50-year correlations.

D.2 Measures of Democracy

D.2.1 Correlations between GDP per capita and Democracy (VDEM Polyarchy) 1790 – Present

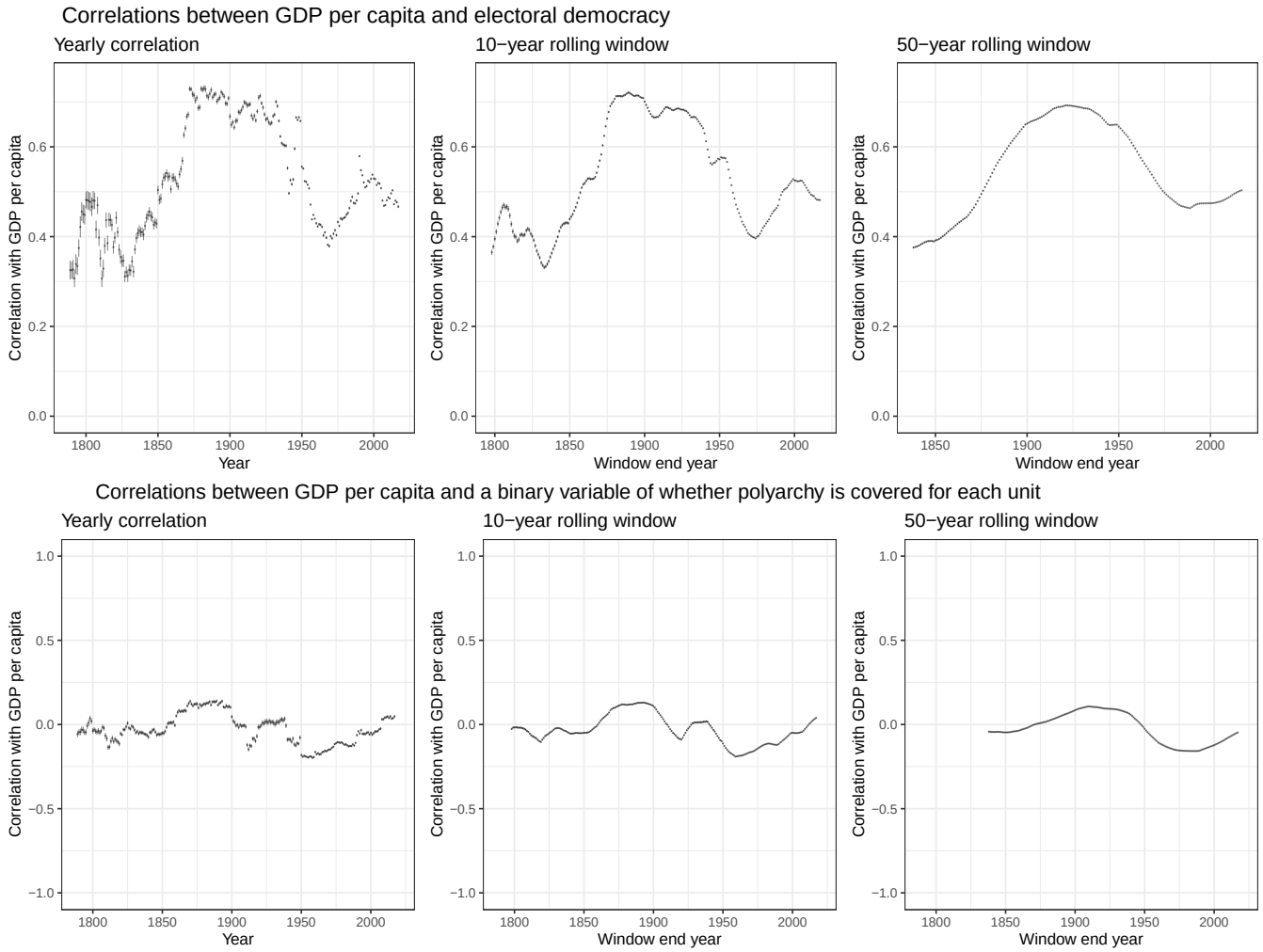

Figure 33: Distribution of Spearman rank-order correlation between the latent GDP per capita variable and a measure of democracy. Democracy is measured by taking $m = 100$ draws from the posterior distribution of a country's latent level of electoral democracy based on the polyarchy variable from VDEM (Coppedge et al., 2019). Correlations are calculated for each 1-year period (left column), 10-year period (middle column), and 50-year period (right column). The second row of panels is the distribution of Spearman rank-order correlation between the latent GDP per capita variable and a variable measuring if the VDEM variable is observed for the country-year unit. Note that we only estimate a correlation coefficient for year-periods with at least 10 cases.

D.2.2 Correlations between GDP per capita and Democracy (Polity IV) 1800 – Present

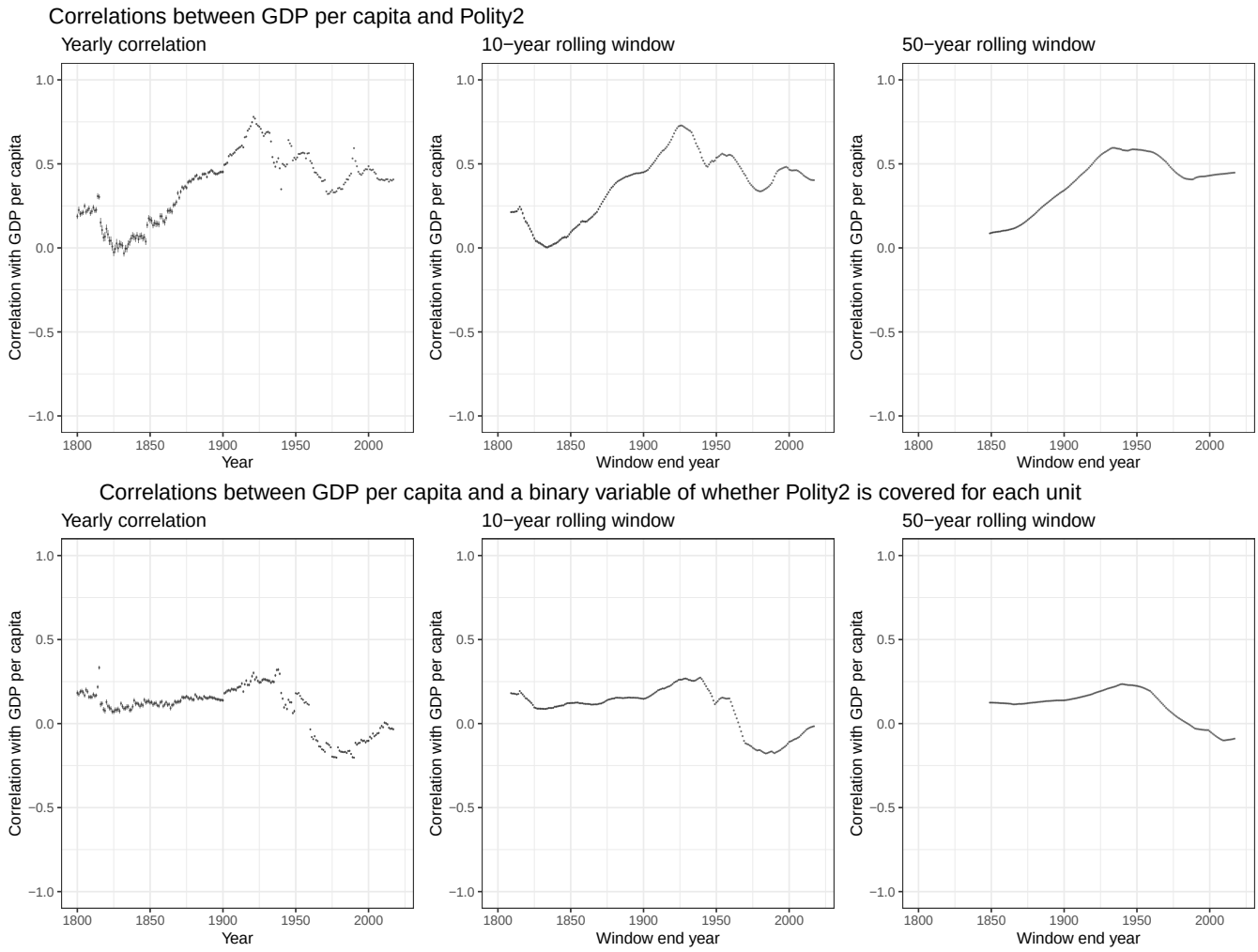

Figure 34: Distribution of Spearman rank-order correlation between the latent GDP per capita variable and a measure of democracy. Democracy is measured using the Polity2 indicator from Marshall, Gurr and Jaggers (2016). Correlations are calculated for each 1-year period (left column), 10-year period (middle column), and 50-year period (right column). The second row of panels is the distribution of Spearman rank-order correlation between the latent GDP per capita variable and a variable measuring if the polity variable is observed for the country-year unit. Note that we only estimate a correlation coefficient for year-periods with at least 10 cases, which means that some of the 1-year correlations begin later than the 10-year or 50-year correlations.

D.3 Measures of Repression

D.3.1 Correlations between GDP per capita and Repression (VDEM Killing) 1790 – Present

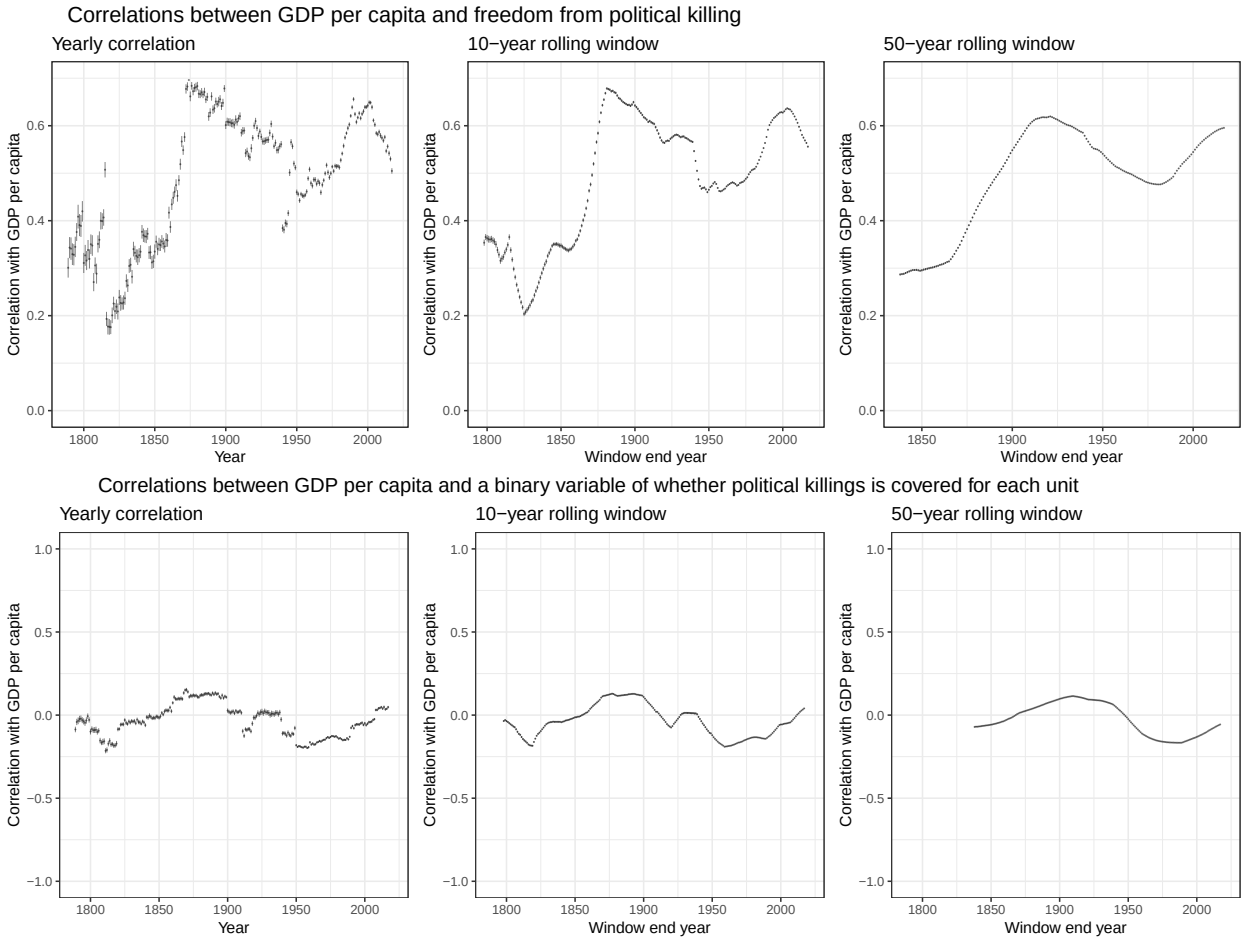

Figure 35: Distribution of Spearman rank-order correlation between the latent GDP per capita variable and a measure of repression. Repression is measured by taking $m = 100$ draws from the posterior distribution of VDEM's freedom from political killing index (Coppedge et al., 2019). Correlations are calculated for each 1-year period (left column), 10-year period (middle column), and 50-year period (right column). The second row of panels is the distribution of Spearman rank-order correlation between the latent GDP per capita variable and a variable measuring if the killing variable is observed for the country-year unit. Note that we only estimate a correlation coefficient for year-periods with at least 10 cases, which means that some of the 1-year correlations begin later than the 10-year or 50-year correlations.

D.3.2 Correlations between GDP per capita and Repression (Latent Human Rights Scores) 1946 – Present

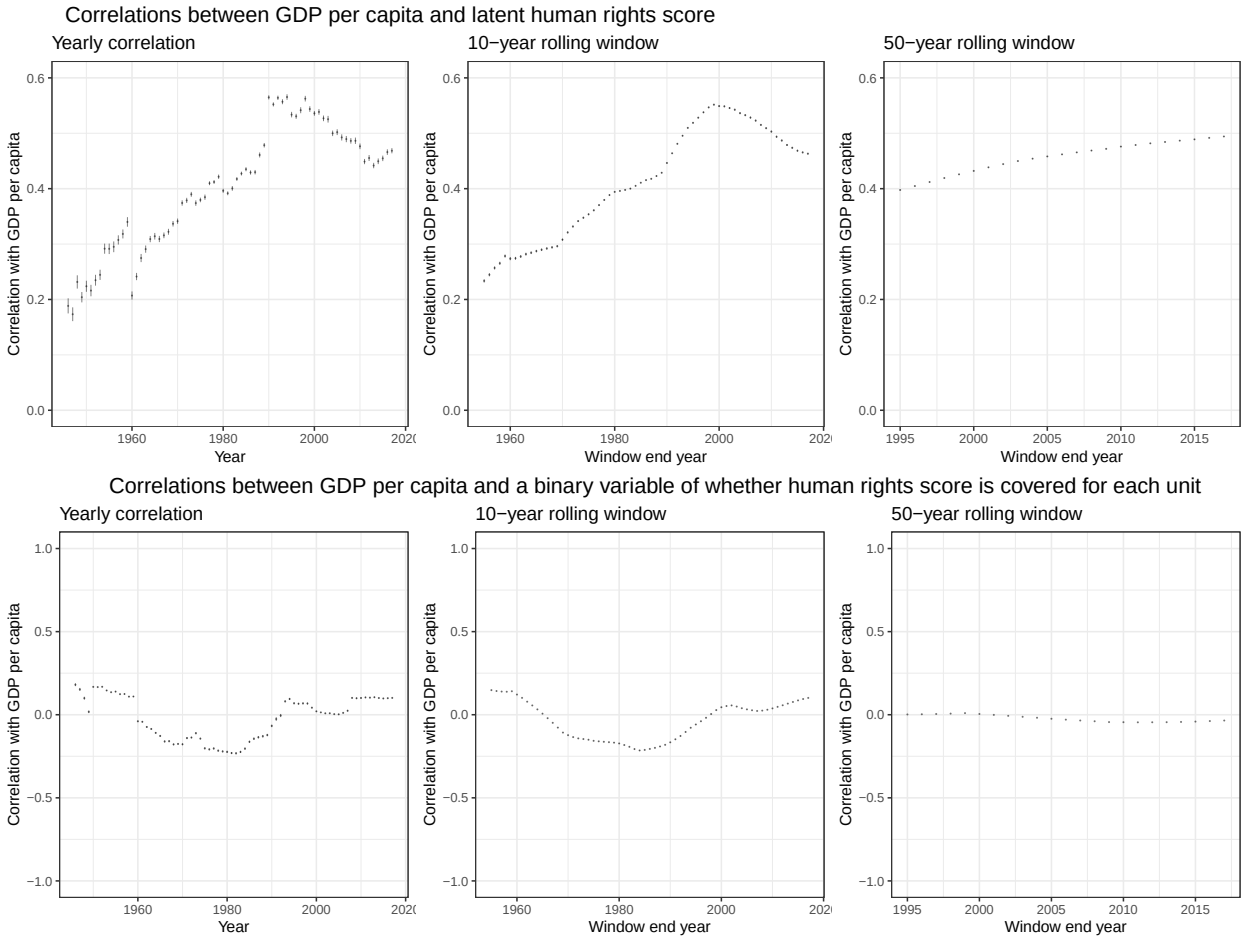

Figure 36: Distribution of Spearman rank-order correlation between the latent GDP per capita variable and a measure of repression. Repression is measured by taking $m = 100$ draws from the posterior distribution of a human rights protection scores latent variable (Fariss, 2014, 2019). Correlations are calculated for each 1-year period (left column), 10-year period (middle column), and 50-year period (right column). The second row of panels is the distribution of Spearman rank-order correlation between the latent GDP per capita variable and a variable measuring if the human rights variable is observed for the country-year unit. Note that we only estimate a correlation coefficient for year-periods with at least 10 cases, which means that some of the 1-year correlations begin later than the 10-year or 50-year correlations.

D.4 Measures of Power Projection

D.4.1 Correlations between GDP per capita and Naval Arming (Proportion of Global Ship Count), 1655 – Present

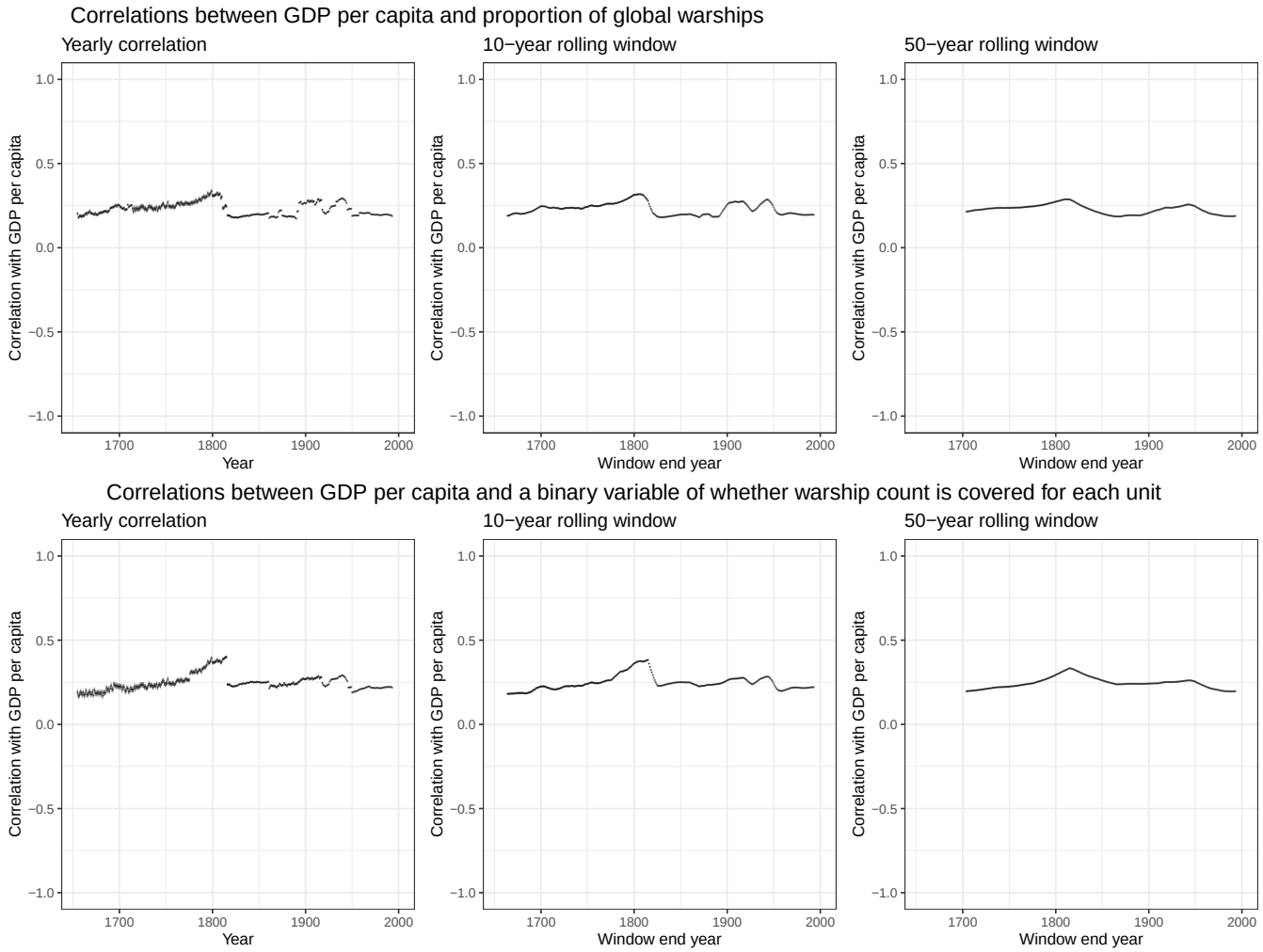

Figure 37: Distribution of Spearman rank-order correlation between the latent GDP per capita variable and a measure of naval arming (number of warships as a proportion of the global total in a given year) (Modelski and Thompson, 1988). Correlations are calculated for each 1-year period (left column), 10-year period (middle column), and 50-year period (right column). The second row of panels is the distribution of Spearman rank-order correlation between the latent GDP per capita variable and a variable measuring if the ship count variable is observed for the country-year unit. Note that we only estimate a correlation coefficient for year-periods with at least 10 cases, which means that some of the 1-year correlations begin later than the 10-year or 50-year correlations.

D.4.2 Correlations between GDP per capita and Naval Arming (Ship Tonnage), 1870 – Present

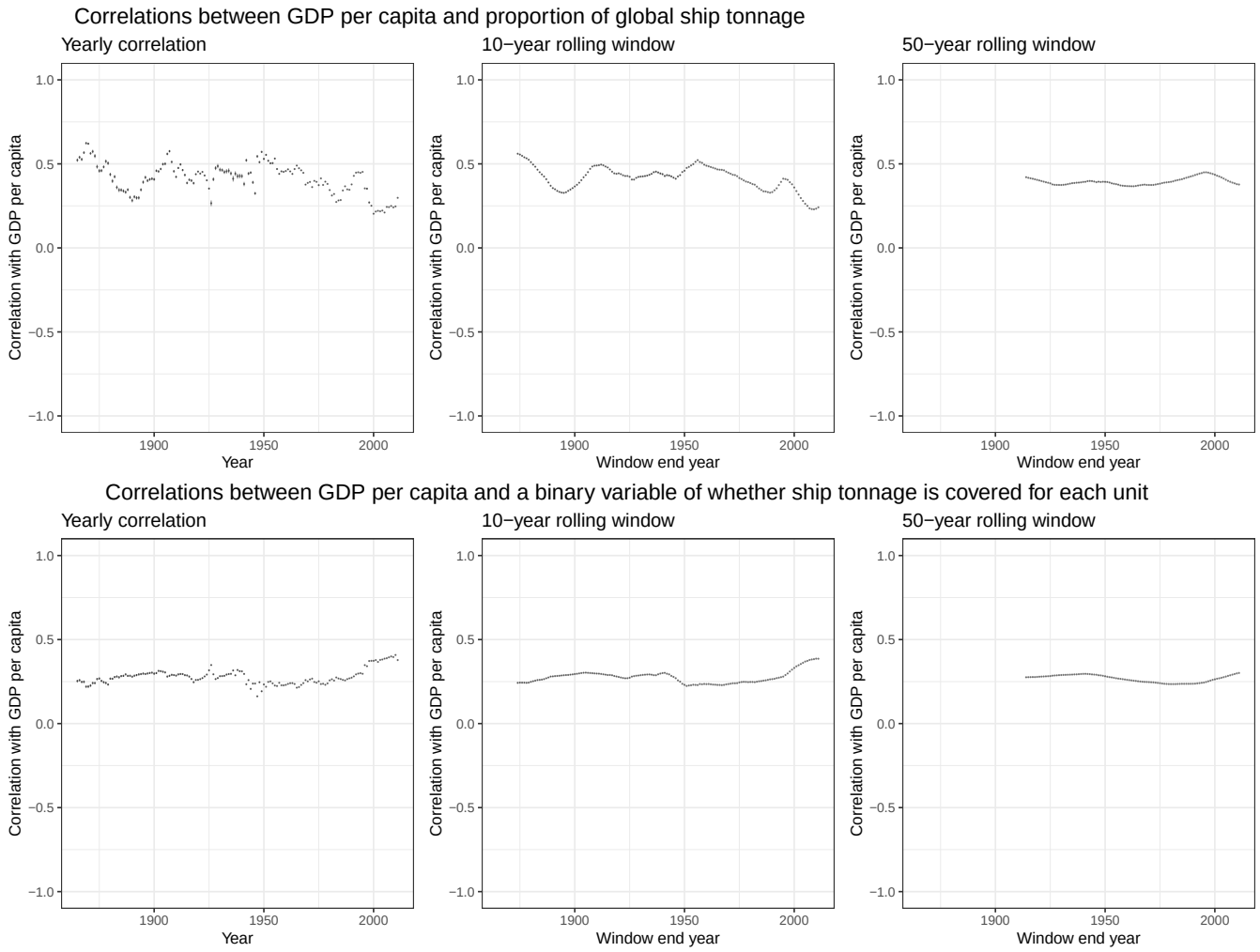

Figure 38: Distribution of Spearman rank-order correlation between the latent GDP per capita variable and a measure of naval tonnage (as a proportion of global tonnage) (Crisher and Souva, 2014). Correlations are calculated for each 1-year period (left column), 10-year period (middle column), and 50-year period (right column). The second row of panels is the distribution of Spearman rank-order correlation between the latent GDP per capita variable and a variable measuring if the naval tonnage variable is observed for the country-year unit. Note that we only estimate a correlation coefficient for year-periods with at least 10 cases, which means that some of the 1-year correlations begin later than the 10-year or 50-year correlations.

D.5 Measures of Conflict

D.5.1 Correlations between GDP per capita and Militarized Interstate Disputes 1816 – Present

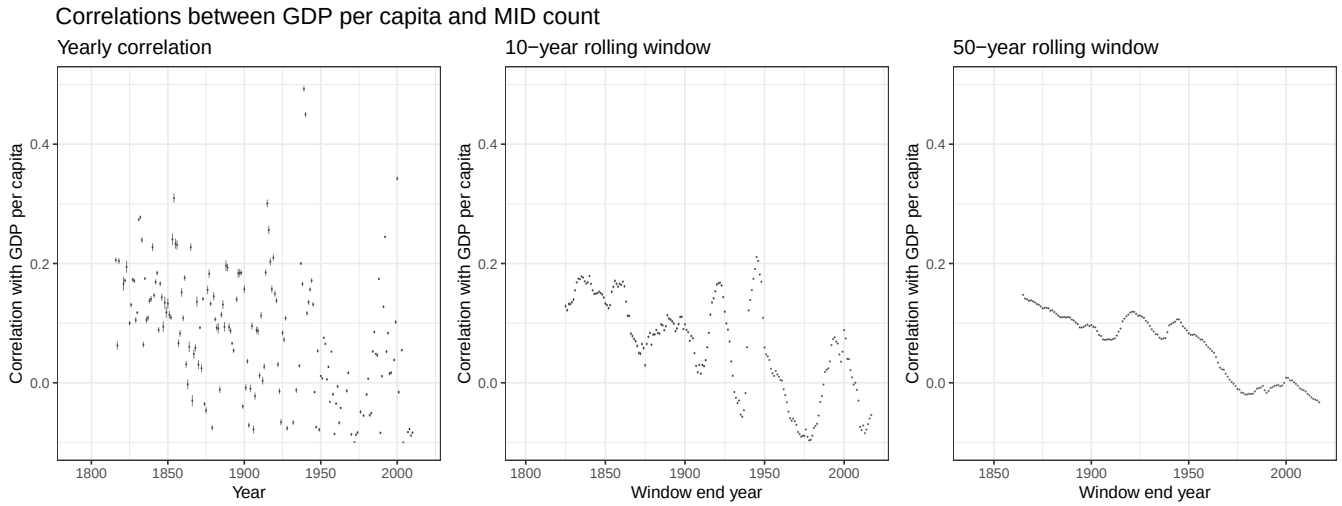

Figure 39: Distribution of Spearman rank-order correlation between the latent GDP per capita variable and a measure of conflict. Conflict is operationalized as the annual count of militarized interstate disputes that a country is involved in (Palmer et al., 2015). Correlations are calculated for each 1-year period (left column), 10-year period (middle column), and 50-year period (right column). Note that we only estimate a correlation coefficient for year-periods with at least 10 cases, which means that some of the 1-year correlations begin later than the 10-year or 50-year correlations. For both MIDs and COW Wars, the correlation has oscillated around 0 over time but generally decreased from a positive correlation towards 0. In early periods, countries with higher levels of GDP per capita were engaged in more conflict (MIDs or Wars) than countries with less GDP per capita. Today, the correlation of 0 suggests that both rich and poor countries are engaging in this type of interstate interaction, though the rate of this type of interaction has also been in decline.

D.5.2 Correlations between GDP per capita and Interstate Wars 1816 – Present

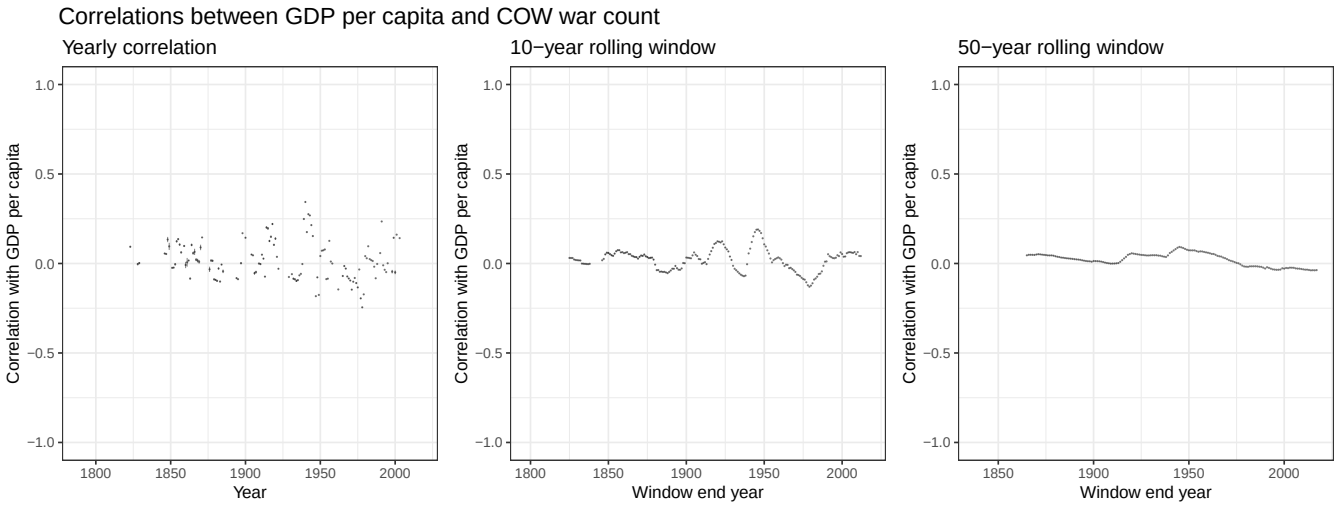

Figure 40: Distribution of Spearman rank-order correlation between the latent GDP per capita variable (top row) and a measure of conflict. Conflict is operationalized as the annual count of interstate wars that a country is involved in from the COW War data v4.0 (Sarkees and Wayman, 2010). Correlations are calculated for each 1-year period (left column), 10-year period (middle column), and 50-year period (right column). Note that we only estimate a correlation coefficient for year-periods with at least 10 cases, which means that some of the 1-year correlations begin later than the 10-year or 50-year correlations. For both MIDs and COW Wars, the correlation has oscillated around 0 over time but generally decreased from a positive correlation towards 0. In early periods, countries with higher levels of GDP per capita were engaged in more conflict (MIDs or Wars) than countries with less GDP per capita. Today, the correlation of 0 suggests that both rich and poor countries are engaging in this type of interstate interaction, though the rate of this type of interaction has also been in decline.

D.5.3 Correlations between GDP per capita and the Total Combat Fatalities 1946 – Present

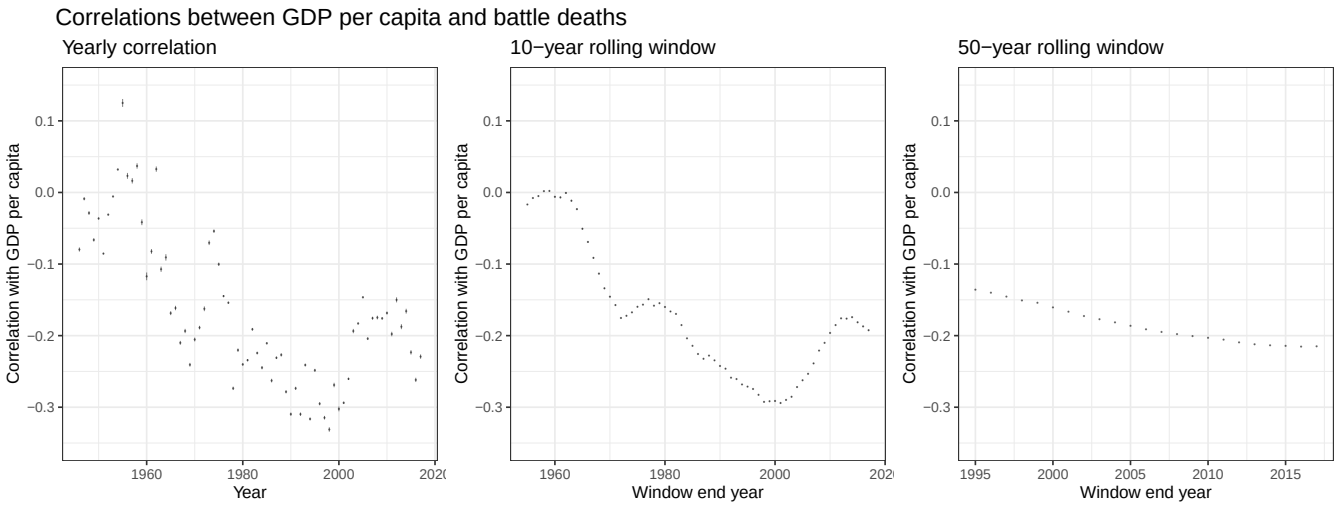

Figure 41: Distribution of Spearman rank-order correlation between the latent GDP per capita variable and a measure of conflict intensity. Conflict intensity is operationalized as the total combat fatalities experienced by a country in a given year, with 1946-1988 data from (Lacina and Gleditsch, 2005) and 1989-2017 data from (Pettersson and Öberg, 2020). Correlations are calculated for each 1-year period (left column), 10-year period (middle column), and 50-year period (right column). Note that we only estimate a correlation coefficient for year-periods with at least 10 cases, which means that some of the 1-year correlations begin later than the 10-year or 50-year correlations. Countries with higher levels of GDP per capita experience lower combat casualty rates than countries with lower levels of GDP per capita. This negative correlation has increased over time.

E Analysis of Measurement Uncertainty

E.1 Regression analysis with uncertainty

As we describe in the main article our latent variable model estimates the level of agreement between indicators for all observed values of the component indicators, which produces country-year estimates of the relative level of uncertainty for each variable. This uncertainty, or measurement error, is a well known problem in standard regression specifications. The new data provided in this article offer quantified estimates of the relative agreement between the component indicators of each variable. This measured uncertainty can easily be incorporated into regression models following guidance described in Bolck, Croon and Hagenaars (2004), Mislevy (1991), Schnakenberg and Fariss (2014), or Fariss (2019).

To incorporate measurement uncertainty in a statistical model, we take $m = 1000$ random draws from the posterior distribution of latent variables, estimate $m = 1000$ regression models, and combine the results (c.f., Schnakenberg and Fariss, 2014). We construct new datasets using random draws from the posterior distribution of the latent variable and then combined using the Rubin (1987) formulas, where the point estimate for each parameter is the mean from the m estimates, the standard error is $\sqrt{\frac{1}{m}\sum_k^m s_k^2 + \left(1 + \frac{1}{m}\right)\sigma_\beta^2}$ where s_k^2 is the standard error from dataset k, and σ_β^2 is the variance in the regression coefficients between datasets. In words, the standard error is the average standard error from each model, plus the variance in the regression coefficients times a correction factor for $m < \infty$.

E.2 Regression analysis of the VDEM Killing variable with uncertainty from GDP per capita (1946 onwards sample)

Here we study the statistical association of two variables — GDP per capita and democracy — with the relative level of political killings (Coppedge et al., 2019) for the same period that the repression data from (Fariss, 2014, 2019) is available: 1946 to 2017. Table 5 replicates the coefficients presented in the main article for a subsample of the post-WWII period from 1946 to 2017 using estimates of freedom from political killings from VDEM (Coppedge et al., 2019), which can be compared to coefficients in Table 6 (reproduced from the main article). The temporal coverage of the data matches the model presented in the main article that uses latent human rights

protection scores from Fariss (2019). For this model, like the models presented in the article, there is substantial attenuation in the size of the coefficient for GDP per capita when the uncertainty is incorporated into the model.

Dependent variable: Draws from freedom from political killings scores (VDEM, Coppedge et al. 2019)

	Point Democracy, Point GDPPC	Draw Democracy, Point GDPPC	Point Democracy, Draw GDPPC	Draw Democracy, Draw GDPPC
GDPPC	0.104 (0.011)	0.106 (0.011)	0.101 (0.011)	0.103 (0.011)
		+ **1.79**	**-2.43**	**-0.68**
Democracy	1.044 (0.055)	1.003 (0.054)	1.047 (0.055)	1.006 (0.054)
		-3.98	+ **0.26**	**-3.73**
LDV	0.716 (0.01)	0.721 (0.009)	0.717 (0.01)	0.722 (0.009)
		+ **0.61**	+ **0.09**	+ **0.70**
Intercept	-0.457 (0.025)	-0.445 (0.025)	-0.454 (0.025)	-0.442 (0.025)
		-2.65	**-0.62**	**-3.29**

Table 5: Comparison of coefficients between point estimates and draws (m = 1000) of democracy and GDP per capita variables 1946–2017. Ordinary Least Squares regression coefficients with standard errors in parentheses. Percentage change in coefficient size compared to baseline model (point estimates for both democracy and GDP per capita) in bold.

Dependent variable: Draws from latent human rights protection scores (Fariss, 2019)

	Point Democracy, Point GDPPC	Draw Democracy, Point GDPPC	Point Democracy, Draw GDPPC	Draw Democracy, Draw GDPPC
GDPPC	0.045 (0.008)	0.045 (0.008)	0.044 (0.008)	0.044 (0.008)
		+ **1.73**	**-2.32**	**-0.62**
Democracy	0.288 (0.033)	0.279 (0.032)	0.289 (0.033)	0.28 (0.032)
		-3.08	+ **0.54**	**-2.55**
LDV	0.891 (0.008)	0.891 (0.008)	0.891 (0.008)	0.892 (0.008)
		+ **0.07**	+ **0.03**	+ **0.10**
Intercept	-0.174 (0.018)	-0.171 (0.018)	-0.173 (0.018)	-0.17 (0.018)
		-1.37	**-0.62**	**-2.01**

Table 6: Comparison of coefficients between point estimates and draws (m = 1000) of democracy and GDP per capita variables 1946–2017. Ordinary Least Squares regression coefficients with standard errors in parentheses. Percentage change in coefficient size compared to baseline model (point estimates for both democracy and GDP per capita) in bold.

E.3 Regression analysis of the Battle Deaths variable with uncertainty from GDP per capita (1946 onwards sample)

Here we study the statistical association of two variables — GDP per capita and democracy — with the number of battle deaths (Lacina and Gleditsch, 2005; Pettersson and Öberg, 2020) from 1946 to 2017. For each version of the model (OLS, Poisson regression, negative binomial regression), there is substantial attenuation in the size of the coefficient for GDP per capita when the uncertainty is incorporated into the model.

Dependent variable: Draws from battle deaths estimates (Pettersson and Öberg, 2020)

	Point Democracy, Point GDPPC	Draw Democracy, Point GDPPC	Point Democracy, Draw GDPPC	Draw Democracy, Draw GDPPC
GDPPC	0.154 (13.543)	0.197 (13.553)	0.16 (13.508)	0.201 (13.517)
		+27.52	**+ 3.51**	**+30.51**
Democracy	10.911 (54.226)	10.601 (53.496)	10.9 (54.358)	10.592 (53.626)
		-2.84	**-0.1**	**-2.92**
sqrt(LDV)	84.423 (27.865)	84.423 (27.864)	84.424 (27.865)	84.424 (27.864)
		0	**0**	**0**
Intercept	-13.857 (44.185)	-13.799 (44.063)	-13.861 (44.116)	-13.803 (43.991)
		-0.42	**+ 0.03**	**-0.39**

Table 7: Comparison of coefficients between point estimates and draws (m = 1000) of democracy and GDP per capita variables 1946–2017. Ordinary Least Squares regression coefficients with standard errors in parentheses. Percentage change in coefficient size compared to baseline model (point estimates for both democracy and GDP per capita) in bold.

Dependent variable: Draws from battle deaths estimates(Pettersson and Öberg, 2020)

	Point Democracy, Point GDPPC	Draw Democracy, Point GDPPC	Point Democracy, Draw GDPPC	Draw Democracy, Draw GDPPC
GDPPC	-0.627 (0.074)	-0.63 (0.074)	-0.577 (0.078)	-0.581 (0.078)
		+ 0.62	**-7.86**	**-7.31**
Democracy	-4.021 (0.861)	-3.948 (0.888)	-4.072 (0.861)	-4 (0.887)
		-1.82	**+ 1.26**	**-0.53**
sqrt(LDV)	0.02 (0.004)	0.02 (0.004)	0.02 (0.004)	0.02 (0.004)
		+ 0.04	**+ 0.31**	**+ 0.35**
Intercept	5.344 (0.215)	5.329 (0.217)	5.308 (0.212)	5.293 (0.215)
		-0.28	**-0.67**	**-0.96**

Table 8: Comparison of coefficients between point estimates and draws (m = 1000) of democracy and GDP per capita variables 1946–2017. Poisson regression coefficients with standard errors in parentheses. Percentage change in coefficient size compared to baseline model (point estimates for both democracy and GDP per capita) in bold.

Dependent variable: Draws from battle deaths estimates(Pettersson and Öberg, 2020)

	Point Democracy, Point GDPPC	Draw Democracy, Point GDPPC	Point Democracy, Draw GDPPC	Draw Democracy, Draw GDPPC
GDPPC	-0.76 (0.264)	-0.761 (0.273)	-0.719 (0.27)	-0.721 (0.279)
		+ **0.21**	**-5.33**	**-5.1**
Democracy	-0.789 (1.255)	-0.8 (1.287)	-0.894 (1.265)	-0.902 (1.294)
		+ **1.37**	**+13.32**	**+14.41**
sqrt(LDV)	0.109 (0.03)	0.109 (0.03)	0.108 (0.03)	0.108 (0.03)
		+ **0.17**	**-0.78**	**-0.62**
Intercept	4.017 (0.519)	4.022 (0.519)	4.001 (0.52)	4.006 (0.521)
		+ **0.13**	**-0.4**	**-0.28**

Table 9: Comparison of coefficients between point estimates and draws (m = 1000) of democracy and GDP per capita variables 1946–2017. Negative Binomial regression coefficients with standard errors in parentheses. Percentage change in coefficient size compared to baseline model (point estimates for both democracy and GDP per capita) in bold.

E.4 Additional Suggestions for Assessing the Influence of Missingness in Data

Finally, the solution to missingness from our measurement model highlights a key benefit of our approach relative to prior work: it provides researchers with an easy way to solve two problems left unaddressed in prior research. The first problem is that most prior scholarship that produced publicly available datasets did not generate estimates for missing values, which both forced researchers to use list-wise deletion and did not provide a principled means to assess whether their results were sensitive to using list-wise deletion. The second problem is that those such as Gleditsch (2002) that did generate estimates for missing values, did not provide an easy way for researchers to assess whether these estimates were biasing their results. We solve both problems by creating a simple version indicator-variable in our dataset that indicates whether the country-year unit was completely missing dataset values or not. Researchers can subset the dataset so that they can analyze statistical relationships between our new estimates and any other variable only for country-year units that have at least one observed dataset value. This allows researchers to easily run their models with and without list-wise deletion and with or without the missing value estimates.

Data are often missing for a reason. Thus, any attempt to estimate missing values based on observed values may introduce measurement error that has the potential to severely bias any estimate of a statistical relationship that relies on those missing values. Given this, we believe that it is critical to rule out the possibility that this source of potential measurement error is large enough to severely bias estimates of a statistical relationship that rely on this data. To do this we suggest a simple solution: running any statistical models using our estimates with and without the units that are completely missing observed dataset values. Critically, this solution actually highlights one of the key benefits of our approach. Specifically, it allows researchers to both rule out the possibility biased estimates of missing values are driving their findings and to assess the sensitivity of their results to list-wise deletion. If researchers find that their results are consistent when running the data with and without estimates of missing values, then they can be more confident that neither issue threatens their results.

F Improving the Validity of Estimates for Missingness

In this section, we explain why our approach is both theoretically and empirically more valid for estimating missing values than prior work. Specifically, we compare our approach to the one adopted by Gleditsch (2002). We first present information about the various choices Gleditsch (2002) adopted to estimating missing values for GDP per capita, and discuss the shortcomings associated with some of these choices. We then explain why how our approach is more theoretically sounds on first principles, and empirically accurate.

F.1 Gleditsch's Approach to Missingness

Gleditsch (2002) also attempts to address missingness for GDP per capita data. Table 10 displays the origin codes from the codebook associated with the original Gleditsch (2002) article, which describes the approach to missingness for several types of cases.

origin code	missingness coding decision	missing
0	- From PWT 6.2 (but see exceptions for population figures in item 3 below)	7687
-1	- From PWT 5.6.	415
-2	- From Maddison Project Database.	742
1	- Imputations for lead/tails based on first/last available value, deflated to current value for gdppc using the GDP deflator from the Bureau of Economic Analysis (see ksgmdw/beagdpdef2000.asc)	161
2	Interpolated value (within series)	0
3	Estimates based on figures from World Bank Global development indicators, using shares to reference countries (See ksgmdw/scaling.asc for details)	622

Table 10: Table of origin codes from Gleditsch (2002) codebook. Additional details are available in the codebook which is available here: `http://ksgleditsch.com/exptradegdp.html`. The beagdpdef2000.asc file and scaling.asc file mentioned in the table and the codebook from Gleditsch (2002) are also available at this website.

For this discussion, we focus on origin code 3 cases, which are cases that are missing data for all years of coverage. For these cases, Gleditsch (2002) calculates a weight between a reference country

and the country with missing GDP data. For these country-year units, a weight is generated and used to scale a country that is missing most or all data relative to another country. The weight is the ratio of the country value in one year based on the world development indicators and the same year for the reference country (e.g., Western Samoa's value divided by Jamaica's value). Note that the first version of the published data from Gleditsch (2002) used values from the CIA World Factbook. In subsequent updates to the dataset, Gleditsch (2002) estimated these weights based on data from the World Development Indicators. The origin 3 country cases that are completely missing data, are just re-weighted versions of the reference country (the reference countries include: Ethiopia, Jamaica, Switzerland, Hungary, Guinea, or Pakistan). In the most recent version of the Gleditsch (2002) data, most of these fully missing state-years are based on Jamaica or Ethiopia.

As an example, we consider how Gleditsch (2002) filled in missing values for East Timor. Gleditsch (2002) sets the missing values for East Timor based on a weighted version of an ideal-type country (for which data was available). The ideal case for East Timor used by Gleditsch (2002) was "a developing socialist economy" (pg. 714). Gleditsch (2002) used Ethiopia as his ideal-type for a "developing socialist economy." Gleditsch used this value to calculate the size of the GDP per-capita and population from the missing case (East Timor) relative to the reference country (Ethiopia), i.e., a proportion based on the relative size of the reference country to the value from the target country (pg. 714). With the proportion from this reference country established, he would then use the reference country's GDP per-capita values, weighted by the proportion. Overall, Gleditsch (2002) used six ideal-type countries (e.g., Jamaica for "small developing country" or Pakistan for "large developing country") to fill in the missing values for all countries with complete missingness.

Clearly, the approach developed by Gleditsch (2002) is not very specific in terms of the theoretical principles used to inform the measurement model, (for illustration, what qualifies a "small or large developing" or "socialist country" is never defined, ex-ante) and will therefore be imprecise in its estimates (small or large developing countries range widely in term of their per-capita GDP depending on how 'small', 'large' and 'developing' are defined). Moreover, while this approach could potentially be used for cross-national comparison, over-time comparison within a single country becomes impossible for any of the 22 countries in which missing data was filled by one of the 6 ideal-type countries (e.g. Ethiopia). Any model that makes over-time comparisons

based on within-country change in GDP or GDP per-capita should drop these 22 countries from their analysis because the within-country variation is identical to the reference case. For illustration, the over-time variation for East Timor will be identical to overtime variation in Ethiopia. However, it is implausible that the two countries' per-capita GDP, GDP or populations would vary in an identical manner over time, given that they are very different countries, with different demographics, on different continents, at different stages of development. To be fair to Gleditsch (2002), some ideal-type style comparisons may be more reasonable than others, but they should all still be dropped if a researcher is interested in making temporal comparisons.

F.2 Comparing Estimates for Missing Cases

We now explain why our approach (i.e., the theoretical principles used to inform our measurement model) is more specific and therefore generates more precise empirical estimates than the method that Gleditsch (2002) employed. Rather than relying on a single ideal-type country from a single year to fill in the missing values for a given country-year, we instead use multiple economic indicators from multiple data sources. To illustrate our method (and why the approach is more valid in comparison to the choices used by Gleditsch (2002)), it is helpful to demonstrate how we would estimate the country-year GDP per-capita values that are missing for East Timor in the Penn World Tables (PWT) dataset (the dataset that Gleditsch (2002) used to estimate his data). Because our approach uses multiple economic indicators from multiple datasets, it is able to utilize information from these indicators regarding East Timor to estimate what its per-capita GDP would be according to the PWT dataset. For example, our model can use per-capita GDP data from the World Bank to estimate what East Timor's per-capita GDP would be according to the PWT dataset. Critically, it does not simply substitute the World Bank data for the missing PWT data, but instead estimates the relationship between the values that are actually observed in each dataset and from this relationship generates an estimate of what East Timor's per-capita GDP would be.

To illustrate why our approach is more accurate in its ability to estimate missing values than the approach adopted by Gleditsch (2002), we showcase two simple correlations. First, we show how precisely aligned our estimates of the Penn World Tables dataset values are with the observed PWT data values that Gleditsch (2002) uses. This in essence gives us a baseline estimate of

how well our model would predict the GDP per-capita values that Gleditsch (2002) uses. If our model is more accurate in its ability to predict the observed data, then this correlation should be higher. The correlation between our predicted estimates and the reported values from PWT that are not missing in the Gleditsch (2002) dataset is 0.989, confirming that our model is very similar and is precisely estimating these values. Now, we turn to the harder case in which the Penn World Tables does not report data for any year for several countries. We generate a correlation between the values produced by Gleditsch (2002) for these cases and observed values from the World Bank, which is a source not used by Gleditsch (2002). If the approach used by Gleditsch (2002) does a good job of calculating the missing values, then our estimates and the estimates from Gleditsch (2002) should be tightly correlated as they were with the actual observations for a country's per-capita GDP. On the other hand, if the model developed by Gleditsch (2002) does not do a good job of estimating the missing values, then our estimates and Gleditsch (2002) estimates should not be tightly correlated. This inference hinges on the assumption that our model is more conceptually grounded in its handling of these cases and is therefore doing a better job estimating the missing values for these hard cases. There is strong evidence for this assumption, as our first correlation demonstrated our model is excellent at predicting the values for the PWT data. The second correlation reveals that the estimates from Gleditsch (2002) for the missing values and our estimates are correlated at only 0.305 in comparison to 0.989. Note that we are comparing our estimates with observed data from the Gleditsch (2002) dataset on the one hand and the World Bank dataset on the other hand to generate these two correlations.

Table 11 summarizes the results we present in the discussion above. In this table, we generate correlation estimates between our estimates of the PWT 9.0 CGDP^e in 2011 US$ variable (Feenstra, Inklaar and Timmer, 2015) and the Gleditsch (2002) variable for each origin code. In sum, our estimates are highly correlated with the values from Gleditsch (2002) when missingness is not an issue. The correlation is much lower when countries are completely missing (origin code 3). In sum, we have demonstrated that our approach to estimating missing values is both theoretically more appropriate and empirically more accurate.

Finally, table 12 displays the count and proportion of country-years that are covered by data from the World Bank but neither the Penn World Tables data nor the Maddison Project data. Our new estimates make use of these additional observed values, in contrast to the data from Gleditsch

(2002) and the Maddison Project, which provide no estimates for these cases.

origin code	correlation
0	0.989
-1	0.883
-2	0.805
1	0.932
2	—
3	0.329

Table 11: Table of correlation between estimated GDP point estimate specifically for the PWT 9.0 $CGDP^e$ in 2011 US$ variable (Feenstra, Inklaar and Timmer, 2015) and the Gleditsch (2002) variable for each origin code. The correlations for the observed data and the data taken from other datasets are each quite close to 1. The only correlation that is low is for origin code 3, which, as described above, are based entirely on a single re-weighted country. Note that the correlation for origin code 2 is not estimated because in the updated version of the Gleditsch (2002), the number of cases that are included for the type of missing case is 0. See the documentation provided by Gleditsch (2002) for additional details.

country_name	Var1	Excluded	Included	Proportion
Monaco	221	67	1	0.01
Liechtenstein	223	67	1	0.01
Andorra	232	24	44	0.65
San Marino	331	68	0	0.00
Kosovo	347	42	16	0.28
Abkhazia	396	10	0	0.00
South Ossetia	397	10	0	0.00
Zanzibar	511	2	0	0.00
Eritrea	531	38	20	0.34
South Sudan	626	2	5	0.71
Yemen, People's Republic of	680	24	0	0.00
Tibet	711	1	0	0.00
East Timor	860	41	17	0.29
Vanuatu	935	21	37	0.64
Kiribati	970	12	46	0.79
Nauru	971	49	9	0.16
Tonga	972	23	35	0.60
Tuvalu	973	32	26	0.45
Marshall Islands	983	23	35	0.60
Palau	986	33	25	0.43
Federated States of Micronesia	987	28	30	0.52
Samoa/Western Samoa	990	24	34	0.59

Table 12: Table displays count and proportion of country-years that are covered by data from the World Bank but neither the Penn World Tables data PWT 9.0 CGDP^e (Feenstra, Inklaar and Timmer, 2015) nor the Maddison Project data. These cases are inferred by a reference country in the Gleditsch (2002) data. Our measurement model instead incorporates all available data for each country to generates country-year distributions of estimates for each of the observed variables from the World Bank.

G Unique Country Coverage by Variable Type

When does a country enter our dataset? A country enters our dataset beginning in the earliest year in which it is covered by any of the datasets we draw on for our project. We use the Gleditsch and Ward (1999) revised list of independent states as the base set of units. Conversion issues between country identifiers in alternative data sources are solved using a script from Graham and Tucker (2019). For example, we generate estimates for England from 1500 to 2018 A.D. and for Ghana from 1820 to 2018 A.D. because the Maddison Project dataset covers England beginning in 1500 and Ghana beginning in 1820.[4]

If a country does not have any coverage across datasets (i.e, it has no observed values for any variable) it does not enter our dataset. Our estimates cover every country that exists in the international system at or after 1816, according to the Gleditsch and Ward (1999) revised list of independent states. Based on the coverage in the available economic and population datasets, we provide coverage for GDP for 197 unique country cases, GDP per capita for 199 unique country cases, and population for 217 unique country cases.[5]

Please note that we do not plan to provide estimates of GDP or GDP per capita for the 18 cases for which we only have observed dataset values for population. In an ongoing and related project, we are currently gathering additional data about countries that existed prior to the 1816 start date of the Gleditsch and Ward (1999) revised list of independent states. The table below indicates when whether we cover each variable for each unique country case.

[4]Details on the underlying source materials for each component measure and coding decisions are available in this document above.

[5]We are missing dataset values for only 8 countries from the Gleditsch and Ward dataset: United Provinces of Central America, Great Colombia, Papal States, Abkhazia, South Ossetia, Transvaal, Orange Free State, and Tibet.

	gwno_code	country_name	pop	gdp	gdppc
1	2	United States of America	1	1	1
2	20	Canada	1	1	1
3	31	Bahamas	1	1	1
4	40	Cuba	1	1	1
5	41	Haiti	1	1	1
6	42	Dominican Republic	1	1	1
7	51	Jamaica	1	1	1
8	52	Trinidad and Tobago	1	1	1
9	53	Barbados	1	1	1
10	54	Dominica	1	1	1
11	55	Grenada	1	1	1
12	56	Saint Lucia	1	1	1
13	57	Saint Vincent and the Grenadines	1	1	1
14	58	Antigua & Barbuda	1	1	1
15	60	Saint Kittsand Nevis	1	1	1
16	70	Mexico	1	1	1
17	80	Belize	1	1	1
18	90	Guatemala	1	1	1
19	91	Honduras	1	1	1
20	92	El Salvador	1	1	1
21	93	Nicaragua	1	1	1
22	94	Costa Rica	1	1	1
23	95	Panama	1	1	1
24	100	Colombia	1	1	1
25	101	Venezuela	1	1	1
26	110	Guyana	1	1	1
27	115	Surinam	1	1	1
28	130	Ecuador	1	1	1
29	135	Peru	1	1	1
30	140	Brazil	1	1	1
31	145	Bolivia	1	1	1
32	150	Paraguay	1	1	1
33	155	Chile	1	1	1
34	160	Argentina	1	1	1
35	165	Uruguay	1	1	1
36	200	United Kingdom	1	1	1
37	205	Ireland	1	1	1
38	210	Netherlands	1	1	1
39	211	Belgium	1	1	1
40	212	Luxembourg	1	1	1
41	220	France	1	1	1
42	221	Monaco	1	1	1
43	223	Liechtenstein	1	1	1
44	225	Switzerland	1	1	1
45	230	Spain	1	1	1

	gwno_code	country_name	pop	gdp	gdppc
46	232	Andorra	1	1	1
47	235	Portugal	1	1	1
48	240	Hanover	1		
49	245	Bavaria	1		
50	255	Germany (Prussia)	1	1	1
51	260	German Federal Republic	1	1	1
52	265	German Democratic Republic	1	1	1
53	267	Baden	1		
54	269	Saxony	1		
55	271	Wurttemberg	1		
56	273	Hesse-Kassel (Electoral)	1		
57	275	Hesse-Darmstadt (Ducal)	1		
58	280	Mecklenburg-Schwerin	1		
59	290	Poland	1	1	1
60	300	Austria-Hungary	1	1	1
61	305	Austria	1	1	1
62	310	Hungary	1	1	1
63	315	Czechoslovakia	1	1	1
64	316	Czech Republic	1	1	1
65	317	Slovakia	1	1	1
66	325	Italy/Sardinia	1	1	1
67	329	Two Sicilies	1		
68	331	San Marino	1		
69	332	Modena	1		
70	335	Parma	1		
71	337	Tuscany	1		
72	338	Malta	1	1	1
73	339	Albania	1	1	1
74	340	Serbia	1	1	1
75	341	Montenegro	1	1	1
76	343	Macedonia (Former Yugoslav Republic of)	1	1	1
77	344	Croatia	1	1	1
78	345	Yugoslavia	1	1	1
79	346	Bosnia-Herzegovina	1	1	1
80	347	Kosovo	1	1	1
81	349	Slovenia	1	1	1
82	350	Greece	1	1	1
83	352	Cyprus	1	1	1
84	355	Bulgaria	1	1	1
85	359	Moldova	1	1	1
86	360	Rumania	1	1	1
87	365	Russia (Soviet Union)	1	1	1
88	366	Estonia	1	1	1
89	367	Latvia	1	1	1
90	368	Lithuania	1	1	1

	gwno_code	country_name	pop	gdp	gdppc
91	369	Ukraine	1	1	1
92	370	Belarus (Byelorussia)	1	1	1
93	371	Armenia	1	1	1
94	372	Georgia	1	1	1
95	373	Azerbaijan	1	1	1
96	375	Finland	1	1	1
97	380	Sweden	1	1	1
98	385	Norway	1	1	1
99	390	Denmark	1	1	1
100	395	Iceland	1	1	1
101	402	Cape Verde	1	1	1
102	403	São Tomé and Príncipe	1	1	1
103	404	Guinea-Bissau	1	1	1
104	411	Equatorial Guinea	1	1	1
105	420	Gambia	1	1	1
106	432	Mali	1	1	1
107	433	Senegal	1	1	1
108	434	Benin	1	1	1
109	435	Mauritania	1	1	1
110	436	Niger	1	1	1
111	437	Cote D<cd>Ivoire	1	1	1
112	438	Guinea	1	1	1
113	439	Burkina Faso (Upper Volta)	1	1	1
114	450	Liberia	1	1	1
115	451	Sierra Leone	1	1	1
116	452	Ghana	1	1	1
117	461	Togo	1	1	1
118	471	Cameroon	1	1	1
119	475	Nigeria	1	1	1
120	481	Gabon	1	1	1
121	482	Central African Republic	1	1	1
122	483	Chad	1	1	1
123	484	Congo	1	1	1
124	490	Congo, Democratic Republic of (Zaire)	1	1	1
125	500	Uganda	1	1	1
126	501	Kenya	1	1	1
127	510	Tanzania/Tanganyika	1	1	1
128	511	Zanzibar	1		
129	516	Burundi	1	1	1
130	517	Rwanda	1	1	1
131	520	Somalia	1		
132	522	Djibouti	1	1	1
133	530	Ethiopia	1	1	1
134	531	Eritrea	1	1	1
135	540	Angola	1	1	1

	gwno_code	country_name	pop	gdp	gdppc
136	541	Mozambique	1	1	1
137	551	Zambia	1	1	1
138	552	Zimbabwe (Rhodesia)	1	1	1
139	553	Malawi	1	1	1
140	560	South Africa	1	1	1
141	565	Namibia	1	1	1
142	570	Lesotho	1	1	1
143	571	Botswana	1	1	1
144	572	Swaziland	1	1	1
145	580	Madagascar (Malagasy)	1	1	1
146	581	Comoros	1	1	1
147	590	Mauritius	1	1	1
148	591	Seychelles	1	1	1
149	600	Morocco	1	1	1
150	615	Algeria	1	1	1
151	616	Tunisia	1	1	1
152	620	Libya	1	1	1
153	625	Sudan	1	1	1
154	626	South Sudan	1	1	1
155	630	Iran (Persia)	1	1	1
156	640	Turkey (Ottoman Empire)	1	1	1
157	645	Iraq	1	1	1
158	651	Egypt	1	1	1
159	652	Syria	1	1	1
160	660	Lebanon	1	1	1
161	663	Jordan	1	1	1
162	666	Israel	1	1	1
163	670	Saudi Arabia	1	1	1
164	678	Yemen (Arab Republic of Yemen)	1	1	1
165	680	Yemen, People's Republic of	1		
166	690	Kuwait	1	1	1
167	692	Bahrain	1	1	1
168	694	Qatar	1	1	1
169	696	United Arab Emirates	1	1	1
170	698	Oman	1	1	1
171	700	Afghanistan	1	1	1
172	701	Turkmenistan	1	1	1
173	702	Tajikistan	1	1	1
174	703	Kyrgyz Republic	1	1	1
175	704	Uzbekistan	1	1	1
176	705	Kazakhstan	1	1	1
177	710	China	1	1	1
178	712	Mongolia	1	1	1
179	713	Taiwan	1	1	1
180	730	Korea	1		

	gwno_code	country_name	pop	gdp	gdppc
181	731	Korea, People's Republic of	1		1
182	732	Korea, Republic of	1	1	1
183	740	Japan	1	1	1
184	750	India	1	1	1
185	760	Bhutan	1	1	1
186	770	Pakistan	1	1	1
187	771	Bangladesh	1	1	1
188	775	Myanmar (Burma)	1	1	1
189	780	Sri Lanka (Ceylon)	1	1	1
190	781	Maldives	1	1	1
191	790	Nepal	1	1	1
192	800	Thailand	1	1	1
193	811	Cambodia (Kampuchea)	1	1	1
194	812	Laos	1	1	1
195	815	Vietnam (Annam/Cochin China/Tonkin)	1		1
196	816	Vietnam, Democratic Republic of	1	1	1
197	817	Vietnam, Republic of	1		
198	820	Malaysia	1	1	1
199	830	Singapore	1	1	1
200	835	Brunei	1	1	1
201	840	Philippines	1	1	1
202	850	Indonesia	1	1	1
203	860	East Timor	1	1	1
204	900	Australia	1	1	1
205	910	Papua New Guinea	1	1	1
206	920	New Zealand	1	1	1
207	935	Vanuatu	1	1	1
208	940	Solomon Islands	1	1	1
209	950	Fiji	1	1	1
210	970	Kiribati	1	1	1
211	971	Nauru	1	1	1
212	972	Tonga	1	1	1
213	973	Tuvalu	1	1	1
214	983	Marshall Islands	1	1	1
215	986	Palau	1	1	1
216	987	FederatedStates of Micronesia	1	1	1
217	990	Samoa/Western Samoa	1	1	1

References

Anders, Therese, Christopher J. Fariss and Jonathan N. Markowitz. 2020. "Bread before guns or butter: Introducing Surplus Domestic Product (SDP)." *International Studies Quarterly* 64(2):392–405.

Bairoch, Paul. 1976. "Europe's gross national product, 1800-1975." *Journal of European Economic History* 5(2):273–340.

Bolck, Annabel, Marcel Croon and Jacques Hagenaars. 2004. "Estimating Latent Structure Models with Categorical Variables: One-Step Versus Three-Step Estimators." *Political Analysis* 12(1):3–27.

Bolt, Jutta and Jan Luiten van Zanden. 2014. "The Maddison Project: Collaborative research on historical national accounts." *The Economic History Review* 67(3):627–651.

Bolt, Jutta and Jan Luiten van Zanden. 2020. "Maddison style estimates of the evolution of the world economy. A new 2020 update." *Maddison-Project Working Paper WP-15, University of Groningen, Groningen, The Netherlands* .

Bolt, Jutta, Robert Inklaar, Herman de Jong and Jan Luiten van Zanden. 2018. Rebasing 'Maddison': new income comparisons and the shape of long-run economic development. GGDC Research Memorandum 174 University of Groningen. Groningen growth and development centre.

Broadberry, Stephen. 2015. "Accounting for the Great Divergence." Online. https://www.nuffield.ox.ac.uk/users/Broadberry/AccountingGreatDivergence6.pdf, accessed 23 November 2016.

Broadberry, Stephen and Alexander Klein. 2012. "Aggregate and per capita GDP in Europe, 1870–2000: Continental, regional and national data with changing boundaries." *Scandinavian Economic History Review* 60(1):79–107.

Coppedge, Michael, John Gerring, Carl Henrik Knutsen, Staffan I. Lindberg, Jan Teorell, David Altman, Michael Bernhard, M. Steven Fish, Adam Glynn, Allen Hicken, Anna Lührmann, Kyle L. Marquardt, Kelly McMann, Pamela Paxton, Daniel Pemstein, Brigitte Seim, Rachel Sigman, Svend-Erik Skaaning, Jeffrey Staton, Steven Wilson, Agnes Cornell, Lisa Gastaldi, Haakon Gjerløw, Nina Ilchenko, Joshua Krusell, Laura Maxwell, Valeriya Mechkova, Juraj Medzihorsky, Josefine Pernes, Johannes von Römer, Natalia Stepanova, Aksel Sundström, Eitan Tzelgov, Yi ting Wang, Tore Wig and Daniel Ziblatt. 2019. "V-Dem [Country-Year/Country-Date] Dataset v9" Varieties of Democracy (V-Dem) Project." https://doi.org/10.23696/vdemcy19.

Crisher, Brian and Mark Souva. 2014. "Power at Sea: A Naval Power Data Set, 1865-2011." *International Interactions* 40(4):602–629.

Deng, Kent G. 2004. "Unveiling China's True Population Statistics for the Pre-Modern Era with Official Census Data." *Population Review* 43(2):32–69.

Fariss, Christopher J. 2014. "Respect for Human Rights has Improved Over Time: Modeling the Changing Standard of Accountability in Human Rights Documents." *American Political Science Review* 108(2):297–318.

Fariss, Christopher J. 2019. "Yes, Human Rights Practices Are Improving Over Time." *American Political Science Review* 113(3):868–881.

Feenstra, Robert C., Robert Inklaar and Marcel P. Timmer. 2015. "The Next Generation of the Penn World Table." *American Economic Review* 105(10):3150–82.

Gleditsch, Kristian S. and Michael D. Ward. 1999. "A revised list of independent states since the congress of Vienna." *International Interactions* 25(4):393–413.

Gleditsch, Kristian Skrede. 2002. "Expanded Trade and GDP Data." *Journal of Conflict Resolution* 46(5):712–724.

Graham, Benjamin A.T. and Jacob R. Tucker. 2019. "The International Political Economy Data Resource." *Review of International Organizations* 14(1):149–161.

Lacina, Bethany and Nils Petter Gleditsch. 2005. "Monitoring Trends in Global Combat: A New Dataset of Battle Deaths." *European Journal of Population* 21(2-3):145–166.

Maddison, Angus. 1995. *Monitoring the World Economy, 1820-1992*. Development Centre Studies Paris: OECD.

Maddison, Angus. 2001. *The World Economy. A Millenial Perspective*. Development Centre Studies Paris: OECD.

Maddison, Angus. 2003. *The World Economy. Historial Statistics*. Development Centre Studies Paris: OECD.

Maddison, Angus. 2010. "Statistics on World Population, GDP and Per Capita GDP, 1-2008 AD." Online. http://www.ggdc.net/maddison/oriindex.htm, accessed 20 August 2016.

Marshall, Monty G., Ted Robert Gurr and Keith Jaggers. 2016. Polity IV Project. Political Regime Charateristics and Transitions, 1800-2015. Dataset users' manual Center for Systemic Peace Vienna, VA: .

Max Roser, Esteban Ortiz-Ospina and Hannah Ritchie. 2013. "Life Expectancy." *Our World in Data* . https://ourworldindata.org/life-expectancy.

Max Roser, Hannah Ritchie and Bernadeta Dadonaite. 2013. "Child and Infant Mortality." *Our World in Data* . https://ourworldindata.org/child-mortality.

Measuring Worth. 2019. Online. https://www.measuringworth.com/glossary/index.php, accessed 1 February 2019.

Mislevy, Robert. 1991. "Randomization-based inference about latent variables from complex samples." *Psychometrika* 56(2):177–196.

Mitchell, Brian R. 2003. *International Historical Statistics: Europe, 1750 2000*. 5 ed. New York: Palgrave Macmillan.

Modelski, George and William R. Thompson. 1988. *Sea Power in Global Politics, 1494-1993*. Seattle, WA: University of Washington Press.

Palmer, Glenn, Vito D'Orazio, Michael R. Kenwick and Matthew Lane. 2015. "The MID4 Data Set, 2002-2010: Procedures, Coding Rules, and Description." *Conflict Management and Peace Science* 32(2):222–242.

Pettersson, Therese and Magnus Öberg. 2020. "Organized violence, 1989-2019." *Journal of Peace Research* 57(4):597–613.

Roser, Max and Esteban Ortiz-Ospina. 2016. "Literacy." *Our World in Data* . https://ourworldindata.org/literacy.

Rubin, Donald B. 1987. *Multiple Imputation for Nonresponse in Surveys*. New York et al.: John Wiley & Sons.

Sarkees, Meredith Reid and Frank Wayman. 2010. *Resort to War: 1816 - 2007*. Washington DC: CQ Press.

Schnakenberg, Keith E. and Christopher J. Fariss. 2014. "Dynamic Patterns of Human Rights Practices." *Political Science Research and Methods* 2(1):1–31.

World Bank. N.d.