

Business as Usual?

A Replication of Andrews, Pritchett, and Woolcock (*Building State Capability: Evidence, Analysis, Action*, 2017)

Ryan H. Murphy *

Journal of Comments and Replications in Economics, Volume 5, 2026-8, DOI:[10.18718/81781.59](https://doi.org/10.18718/81781.59)

JEL: O43; P00

Keywords: State capacity; State capabilities, Institutional development; The Big Stuck; Worldwide Governance Indicators; Quality of government; International Country Risk Guide

Data Availability: The Excel to reproduce the results of this replication can be downloaded at JCRE's data archive (DOI: [10.15456/j1.2026143.1952233993](https://doi.org/10.15456/j1.2026143.1952233993)).

Please Cite As: Murphy, R. H. (2026). Business as Usual? A Replication of Andrews, Pritchett, and Woolcock (*Building State Capability: Evidence, Analysis, Action*, 2017). *Journal of Comments and Replications in Economics*, Vol. 5 (2026-8). DOI: [10.18718/81781.59](https://doi.org/10.18718/81781.59)

Abstract

In *Building State Capability*, Matt Andrews, Lant Pritchett, and Michael Woolcock argue that there has been little institutional development in recent decades (“The Big Stuck”). A subsequent blog post by Pritchett (2020) shores up its empirical foundations, with some emphasis on results derived from the PRS Group’s International Country Risk Guide (ICRG). We first show that the start year (1996) for these analyses is not robust to alternative start dates and appears to be driven by idiosyncratic temporal variation of the ICRG panel. We then show that an index of state quality derived from the Varieties of Democracy (V-Dem) project yields the opposite conclusion to that implied by the ICRG data. We do not conclude that the perspective articulated in *Building State Capability* is necessarily incorrect, but that the existing data on institutional quality are mixed and ambiguous, rather than clear evidence of stagnation.

*Southern Methodist University. rhmurphy@smu.edu

1 Introduction

Andrews et al. (2017; hereinafter APW) argue that developing countries have struggled in the first two decades of the 21st century because development agencies and economists ask them to do too much, too fast. Instead of making steady gains over time, these countries have largely stagnated. They call this stagnation “The Big Stuck.” Part of the empirical basis for their argument is that the quality of state institutions did in fact stagnate over this period. In support, APW (2017: 14-21) show, using data from *Worldwide Governance Indicators* (WGI), the *Quality of Government* (QOG) index, and the *Fragile States Index* (FSI), that very few of the 102 historically developing countries were on a trajectory from 1996-2012 to reach even the lower bound of state quality in the developed world, i.e., the level of institutional quality of countries such as Spain or Czechia. This paper questions the empirical basis for “The Big Stuck.”

An emphasis on making incremental, tangible gains may be a laudable and sound development strategy. That may well be true regardless of our results regarding APW’s empirical findings. However, we identify two interrelated issues with the data supporting their conclusion about the trajectory of state institutional quality. Together, these two issues constitute this paper’s contribution. First, we argue the time-series properties of the *Quality of Government* index (and the *International Country Risk Guide* [ICRG] data underlying it) are not credible when countries are aggregated. Even if the ICRG data are taken at face value, the evidence on the trajectory of state institutional quality is more ambiguous than demonstrative of stagnation. Second, the title phrase “Business as Usual” refers to a condition defined in APW’s Table 1.1, which characterizes only a few developing countries that are progressing rather than stagnating in their institutional development. While we do not believe that it is “Business as Usual” for most developing countries, it is “Business as Usual” for far more countries than are found in APW (or implied by Pritchett [2020]).

Pritchett (2020) updates APW’s analysis, showing that little has changed through 2018, while addressing potential issues with the results published in 2017.¹ Since APW is the formal work published by an academic press, we focus primarily on APW, but that analysis will inform our later analysis of Pritchett (2020), which is perhaps a more definitive statement of the claim. A key difference between the 2017 book and 2020 post is that Pritchett (2020) addresses the concern that the global mean of the *Worldwide Governance Indicators* is normalized to zero in each year, meaning that, in the aggregate, countries cannot show improvement. He shows that QoG, which lacks this constraint, does not show improvement but decline. In other words, the built-in stagnation of WGI may have provided an overly rosy picture in APW, according to Pritchett (2020). “The Big Stuck” is too optimistic, if anything.

Section 2 describes how APW constructed Table 1.1 and assembled the underlying data, and presents an initial replication attempt. Section 3 focuses on the *International Country Risk Guide* and shows that ICRG’s negative trajectory, which drives the results in APW and Pritchett (2020) appears to be driven by a mid-1990s “hump” in the time series of an otherwise stagnant panel dataset; choosing alternative start years outside that window yields different results. Section 4 revisits Pritchett’s (2020) exercise using a measure of state capacity developed in O’Reilly and Murphy (2022) derived from *Varieties of Democracy*, and reports far more optimistic results. We

¹The origin of this project was the observation of an issue in APW that, unbeknownst at the time, was fully addressed in Pritchett (2020).

identify a dozen countries where there is complete disagreement between the trajectories implied by O'Reilly and Murphy (2022) and those implied by ICRG, and provide the precise *Varieties of Democracy* data indicating improvements in governance quality among those dozen countries. Section 5 concludes.

2 Data and Replication

This section describes APW's underlying data sources, the methodology of APW's analysis, and an attempted replication of that analysis. The data and analysis in APW, using *Worldwide Governance Indicators*, the *Quality of Government* index, and the *Fragile States Index*, begin in 1996. APW very likely chose 1996 as the start year in APW because WGI begins in 1996. The present build of WGI is far more complicated, but only a subset of the underlying sources and inputs extend back to 1996. To apply WGI to their research question, APW select three indicators from WGI, "Rule of Law," "Government Effectiveness," and "Control of Corruption," following the conceptual and empirical work in Drumm (2015). These are combined into an aggregate state capabilities variable by taking a simple average.

The *Quality of Government* index is based on data from the PRS Group's *International Country Risk Guide* and is calculated by taking the average of three ICRG indicators: "Corruption," "Law and Order," and "Bureaucratic Quality." These data pre-date the WGI series and are available back to 1984 for many countries. The original purpose of the data was to provide businesses with expert guidance on investment risk. The data are relevant for the WGI and APW's purposes because these three dimensions of government performance are important for investment decisions. Finally, the *Fragile States Index* measure (the quality of public services) used in APW is a single piece of a broader index of state stability and quality.

We will describe in detail how APW apply these three data sources. APW use these data to analyze a broad set of developing countries and categorize their level of institutional development and the trajectory of their institutions. They can then determine which countries are at least on pace to develop good institutions by the end of the 21st century. Countries are placed in a 4x4 matrix based on their current level of institutional development and their trajectory, and countries that are on pace are marked as such.

The authors took each of the three indices, QOG, WGI, and FSI, and rescaled each so that the sample minimum equals 0 and the sample maximum equals 10. APW use a sample of 102 historically developing countries. An average of WGI, QOG, and FSI is used to score countries in 2012, the most recent year in their analysis. The 2012 score is used to assess present institutional quality. Because FSI data are not available before 2006, APW calculate the medium-run trajectory of each country using only WGI and QOG from 1996 to 2012. To calculate the trajectory, they calculate country-level trend lines for each of the two variables using regressions of the score on time. Then, for each country, they average the two coefficients. That average is used to assess the trajectory of institutional quality. It is not stated in-text, but when QOG becomes available for several Eastern European countries in the late 1990s, the trajectory is constructed from the more limited time series beginning when QOG becomes available, rather than omitting QOG for those countries.

Their Table 1.1 (2017: 20) summarizes their results in the aforementioned 4x4 matrix. Their table is reproduced below as Table 1. Along the vertical axis, they classify the quality of state

institutions using the combined FSI/QOG/WGI in 2012 into one of four categories: “Strong” (greater than 6.5), “Moderate” (between 4 and 6.5), “Weak” (between 2.5 and 4), or “Very Weak” (less than 2.5). Along the horizontal axis, they classify countries’ trajectories using the regression coefficients (interpreted as points per year), using four categories: “Rapid Negative” (decline of at least 0.05 points per year in magnitude), “Negative” (decline by less than 0.05 points per year in absolute value), “Positive” (increase by less than 0.05 points per year), and “Rapid Positive” (increase by at least 0.05 points per year). The 6.5 cutoff is meant to represent the lower bound of “strong state institutions” roughly the level of Spain or Czechia.

Of the 102 countries, only 13 were on a trajectory to reach a score of at least 6.5 by 2100. Reaching that threshold by 2100 defines “Business as Usual.” The 13 countries were Chile, Singapore, South Korea, Qatar, the United Arab Emirates, Indonesia, Colombia, Turkey, Algeria, Albania, Saudi Arabia, Uruguay, and Croatia. The recent trajectory of all other countries was stagnation or decline (though an additional three, Niger, Guinea-Bissau, and Liberia, were classified as “Rapid Positive” and appear to be excluded because they are so far away from the threshold; Iraq was also on pace but moving more slowly). In APW’s Table 1.1 and Table 1 below, the “Business as Usual” countries are in bold. Section IV reproduces the methodology from APW using another data source. That data source does not include nine of the 102 countries in the analysis because of data availability constraints. These nine countries are highlighted in blue.

We attempt to replicate their results, using descriptions found in APW and the closest online data we could identify, before continuing. Table 2 presents the results of our attempt to match their methodology as closely as possible, though some ambiguity remains.² For QOG and FSI, we use the versions currently made available online; for WGI, we use the 2014 vintage, since that appears to be the version APW use.³ We are able to reproduce the results for 94 of the 102 countries. The successfully replicated 94 countries are highlighted in green. The seven countries that were off by one cell in one dimension are highlighted in yellow. The one country that was off by one cell in both directions, Algeria (DZA), is left unhighlighted. Countries that are calculated to be on pace to reach strong institutional quality by 2100 are shown in bold.

Among the countries that did not replicate, Israel is classified as having a higher level of institutions (one category higher than in APW), whereas Bahrain, Angola, and Guinea are classified one level lower. We also find that the trajectories of Uganda, Ukraine, and Somalia are one category worse than in APW. Algeria differs by one category in both dimensions and is by far the most difficult discrepancy as a result of a minor methodological difference. All other countries match APW’s classifications. The only substantively important difference is Algeria, and if the results here are correct, the revision would only reinforce APW’s argument (since the revised classification is more pessimistic in outlook), rather than detract from it.

²They write (2017: 15) “In order to compare these three separate data sources (QOG, FSI, WGI) we rescale each of them to a zero to 10 range by assigning the lowest recorded country/year observation (this was typically Somalia) and the highest recorded country/year observation as 10 (this was typically Singapore).” There is some ambiguity because “typically” seems to imply that this was done several different times, meaning that the min and max were reestablished for each individual year. If there is instead a single global min/max, there is also ambiguity concerning whether it is the global max within the sample of 102 countries, for all countries but just for the years in question, or the global max for all countries and all years. We chose the last of these as closest to the text.

³They use 2013 data for WGI, whereas for FSI and QOG they use 2012, saying that these are the most recent data available. The 2014 vintage of WGI is the first reporting the year 2013.

We also fully reproduce the list of countries where “Business as Usual” (on pace for strong state capabilities by 2100) for all countries identified by APW except Algeria. As above, these are shown in bold. We do not find any additional countries for which “Business as Usual” applies using APW’s methodology as we understand it. We attribute the remaining differences to very minor details concerning issues such as how the minima or maxima were implemented. Seemingly minor differences in how the data are assembled can yield much larger changes than those shown below. These minor potential differences are not the concerns addressed in this paper, and the various possible permutations down the garden of forking paths do not appear to be driving APW’s result. When we set up the data as APW do, we also find evidence for “The Big Stuck.”

3 Problems in the ICRG Panel

Two data sources are used to calculate countries’ medium-run trajectories: QOG and WGI. A substantive theoretical issue with using WGI is that by construction the global mean in each indicator is set to zero in each year. Practically and empirically, it is less obviously problematic. The issue of whether WGI’s underlying data have a trend is a question that has been studied multiple times by WGI’s authors, and they conclude that no correction is warranted (Kaufmann et al. 2005; Kaufmann and Kraay 2024). In a footnote on page 19, APW perform checks on the averages of QOG, WGI, and FSI, and conclude that there is no reason to correct for underlying trends when they perform their own analysis.

We are not seeking to challenge the broader point about overall trends in WGI. We are looking instead at particular indicators underlying WGI in the section that follows. We can make the narrower points relevant to this comment on APW without delving into the much larger question about the *Worldwide Governance Indicators* research programme. We find that when we focus on consistent sets of countries across time, from 1996 to more recent periods, specific issues quickly arise.

Although APW use QOG/ICRG and WGI as two separate line items, the ICRG indicators also partly underlie WGI. Among the WGI underlying sources for which we can make consistent comparisons, QOG/ICRG is by far the most pessimistic, and there is a reasonably strong case that its pessimism is a statistical artifact, reflects an unreported change in methodology, or is otherwise somehow unrelated to underlying institutional quality.

That said, the weight WGI assigns to ICRG data is limited. ICRG “Corruption” contributes to WGI “Control of Corruption,” ICRG “Law and Order” contributes to WGI “Rule of Law,” and ICRG “Bureaucratic Quality” contributes to WGI “Government Effectiveness.” Among the “Representative Sources” (i.e., those covering many countries), the other variables feeding into the three WGI indicators for 1996 were: *The Economist* Intelligence Unit, the World Economic Forum’s *Global Competitiveness Report*, *Varieties of Democracy* (at least in the current vintage of WGI⁴), and the S&P Global Country Risk Service. Other “Representative Sources,” including the *Institutional Profiles Database*, the *Gallup World Poll*, and *Trafficking in Persons* did not yet exist in 1996, so they can be ignored for the purposes of 1996.

⁴It had not been part of the 2014 vintage of WGI that was used by APW, but we are using it as another point of comparison for ICRG.

Suppose we take the WGI data and restrict them as follows: 1) we consider only the 102 countries analyzed by APW; 2) for each variable, we further restrict the sample to countries scored by that variable in 1996; 3) we analyze variables aggregated at the source level (for example, the *Global Competitiveness Report* corruption variables that contribute to WGI “Control of Corruption”) because WGI is inconsistent in whether variables are disaggregated further;⁵ and 4) we only consider the “Representative Sources.”

Table 3 reports source-level changes for 1996-2012 for this restricted but consistent set of countries and variables. To calculate these changes in the source data, we calculate the source-specific 1996 standard deviation and the 1996 and 2012 group means for each source and for each of “Control of Corruption,” “Rule of Law,” and “Government Effectiveness.” For example, the first number in the table indicates that the group mean of the ICRG corruption data fell by 0.605 of the 1996 standard deviation.

ICRG is by far the most pessimistic of the five data sources, and no other source approaches its pessimism. ICRG shows a substantial decline in “Control of Corruption” and “Rule of Law,” and a modest decline in “Government Effectiveness.” The table contains two other negative signs, compared with nine positive signs. The ICRG declines in “Control of Corruption” and “Rule of Law” are the largest changes in absolute magnitude in the table. ICRG appears to be the outlier within WGI.⁶

We can explore this further. Of the 102 historically developing countries, 89 have time series⁷ for QOG going back to at least 1985. Figure 1 reports the mean value of QOG from 1985 to 2023. The time series of the mean is flat from 1985-1991 and then again from 2002-2023. Between these periods, there is a large increase up through 1995 and then a collapse back down to the earlier level. That 1995-2001 decline in scores overall is the largest source of global variation in what APW measures as a decline occurring from 1996-2012. The standard deviation of the mean QOG score of the 89 countries in the “flat” years 1985-1991 and 2002-2023 is 0.0096. The difference between the mean value of the flat periods and the peak in 1995 is approximately nine standard deviations, using the standard deviation within the flat periods.

It is not clear whether this affected the QOG data, but Hoti and McAleer (2004: 18) report that PRS’s methodology was changed in 1997 to reflect changing reality after the end of the Cold War. The researchers producing the *Worldwide Governance Indicators* are hardly oblivious to their own variables and have examined whether there were breaks in the data around this time by comparing PRS with other sources (Kaufmann et al. 2006).

Nevertheless, for our purposes here, we still have a nine standard deviation discrepancy to explain, and however nostalgic we are for 1990s-era governance, the real world shift was very likely nowhere near nine standard deviations. We do not see any reason why APW chose 1996, other than the obvious: it is the first year for which *Worldwide Governance Indicators* data are

⁵Multiple indicators from the EIU inform “Control of Corruption,” but the replication materials only report a column labeled “EIU23CC,” which is the aggregate of the two sub-indicators from EIU for “Control of Corruption” in the year 2023. The exact number of variables underlying a source is not consistent across vintages. The example given here is consistent with the 2025 description found at <https://www.worldbank.org/content/dam/sites/govindicators/doc/CC-2025.pdf>

⁶The low model weight ICRG tends to receive in WGI supports the idea that it is not very effective at measuring the latent dimensions of governance.

⁷This counts “USSR” towards Russia.

available. However, if we were to pick a theoretically or narratively salient year to mark the beginning of the modern era, it would not be 1996, but 1991 or 1992, marking the end of the Cold War; the round number benchmark of 2000, or perhaps 2001, given that 9/11 may have marked a structural break in the trajectory of institutions worldwide.

We can momentarily abstract from WGI to calculate, using only ICRG data with *alternative starting years*, 1) how many countries would be improving at a rapid rate of at least 0.05 points per year on the 10.00-point scale and 2) how many countries are on pace to reach strong institutions by 2100. How many countries would be classified as “Business as Usual” in APW if we used only ICRG data and simply varied the initial year? Leaving out WGI allows us to show the sensitivity of the analysis to the choice of 1996 as the start year, in part because we can go back further in time. Figure 2 presents three time series in which the x-axis varies the starting year used to calculate the trajectory of institutional quality, with the possible initial years 1991-2008.⁸ Each of the time series reflects a different end year for the analysis: 2012 (as in APW), 2018 (as in Pritchett 2020), or 2023 (most recently available).

To recapitulate, Figure 2 tests the robustness of the initial year. The initial year, set to 1996 in APW, is a and the ending year is b . The value of a varies along the x-axis. The value of b is one of 2012, 2018, or 2023, and is represented as one of three time series. The year 1996 is not the absolute nadir for any of the three time series, but it is near the bottom in each case. Extending the x-axis before 1996 is only possible because we narrow the measure of institutions to ICRG, which covers countries over a longer period.

Figure 2 examines the robustness of the number of countries classified as “Rapid Positive” as in APW’s Table 1.1. That is not the only relevant question. Figure 3 then applies the same procedure, but reports the number of countries on pace to reach strong institutions by 2100. Again, 1996 is near, though not at, the absolute nadir.

The start date chosen for the ICRG analysis appears to have a substantial effect on the results, as the 1995 peak in Figure 1 corresponds to the nadirs found in Figure 2 and Figure 3. Some of these curves in Figure 2 and Figure 3 are flatter than others. The large counts at the right side of each graph likely reflect short-run rates of improvement, rather than medium-run rates of improvement. However, the 1996 starting point clearly contributes to APW’s pessimistic findings, and this same problem is relevant for Pritchett (2020) as well.

It remains possible, though unlikely, that ICRG was measuring a real late 1990s worldwide institutional collapse relative to any other period in the sample, and there are certainly narrow examples in which this may have been true (for example, in Russia). However, the collapse is not reflected in the other “Representative Sources” used in WGI, and casual empiricism does not suggest it was a uniquely bad period in recent global history.

⁸Individual regressions were run for each starting and ending year pair, implicitly at least through Microsoft Excel’s SLOPE() function.

4 *An Alternative Measure of State Quality and the Interpretation of Pritchett (2020)*

Pritchett (2020) extends APW by shoring up APW’s Table 1.1 against various empirical and conceptual objections. In this section, we describe the figures in Pritchett (2020) and treat them as the most recent statement of the APW/Pritchett position. Using this analysis as the benchmark, we then apply data from a different recent dataset on state capacity, showing considerable differences between our findings and those of Pritchett, suggesting, at minimum, that the result is less robust than Pritchett (2020) implies.

Pritchett (2020) does not combine WGI and QOG (or WGI, QOG, and FSI), as APW do. Instead, he reports the correlations between the various indices and the consistency of their results. Table 1 in Pritchett (2020) updates APW’s Table 1.1 from 2012 to 2018, using only WGI to construct both levels and trajectories, and still finds stagnation. (One noteworthy difference is terminology. “Middle” capability is now “Moderate,” “Very Weak” is now “Fragile,” and “Rapid Negative” is now “Collapsing.” Other labels remain the same.) Pritchett then shows a high correlation between 2018 levels of ICRG/QOG and WGI (Pritchett’s [2020] Figure 1), and high correlations between their trajectories from the earlier period⁹ to 2018 (Pritchett’s [2020] Figure 2). Pritchett’s (2020) Table 2 then reproduces APW’s Table 1.1 using ICRG/QOG alone and finds an even more pessimistic picture. Pritchett (2020) examines correlations with other state capability variables: namely “Public Services” from FSI and two variables, “Steering Capability” and “Efficient Use of Resources,” from the Bertelsmann Stiftung Transformation Index (BTI). The BTI measures are the only variables that suggest an upward trend rather than stagnation or decline.

Earlier, we compared WGI’s “Representative Sources” over 1996-2012 using a consistent sample of countries. Those results appear above as Table 3. Table 4 below updates Table 3 by extending the window from 2012 to 2018. All of the “Representative Sources” shown in Table 3 and now in Table 4, aside from ICRG, remain far more optimistic than ICRG, and in some cases, the differences between ICRG and the others are more pronounced for 2018 than they were for 2012. The full calculation of WGI is far more complicated than this exercise. Nevertheless, among the 102 historically developing countries with scores available in 1996, and among the “Representative Sources” that extend back to 1996, ICRG remains the most pessimistic source through 2018.

To sharpen this claim, we can apply the dataset developed in O’Reilly and Murphy (2022), which measures the quality of state institutions. This dataset has the advantage of extending back through 1996 and much farther into the past. Though it is less widely used than some other measures of institutions, recent applications of this dataset include Fergusson et al. (2023), Tiganasu et al. (2025), and Balcells and Kalyvas (2026).

A large part of the point of Pritchett (2020) is that ICRG/QOG is not mechanically forced to show no aggregate improvement. Pritchett therefore reproduces APW’s original Table 1.1 using only the ICRG/QOG data in his Table 2. Our Table 5, reproducing Pritchett (2020)’s Table 2, only reports results for 90 countries, rather than the 102 found in APW’s Table 1.1. Countries that appeared in APW but do not appear in Pritchett’s (2020) Table 2 are Albania, The Bahamas, Belarus, Brunei,

⁹For whatever reason, Figure 2 of Pritchett (2020) uses 1998 instead of 1996 as its initial year.

Croatia, Democratic Republic of the Congo (“ZAR” in APW’s Table 1.1), Guyana, Israel, Moldova, Russia, Suriname, and Ukraine. These countries appear in the current ICRG/QOG data, so it is not readily apparent why they did not appear in Pritchett (2020). Of the 90 countries in Table 5, six are highlighted in blue to indicate that they will not subsequently appear in Table 6 because of data limitations.

O’Reilly and Murphy (2022) use variables from the *Varieties of Democracy* (V-Dem) project (Coppedge et al. 2024) related to state capacity to create a composite index closely related to the measures used by both APW and Pritchett (2020). Using Hendrix (2010) and Savoia and Sen (2015) as their starting points, they focus on bureaucratic and administrative quality, legal capacity, infrastructural capacity, military capacity, and fiscal capacity as core dimensions of state capacity. The variables from V-Dem that capture these dimensions are “Rigorous and Impartial Public Administration;” “Educational Equality” (an evaluation of the universality of education, not its average quality); the V-Dem “Rule of Law” index; “Particularist or Public Goods;” “State Fiscal Source of Revenue”; and “State Authority Over Territory.”

Drumm (2015) serves as an empirical starting point for APW (2017: 16). O’Reilly and Murphy (2022) also cite Drumm (2015) as an empirical and conceptual source, along with Hendrix (2010) and Savoia and Sen (2015). Drumm (2015) argues that “effectiveness” is the primary dimension (more precisely, a statistical factor) along which institutions differ in quality worldwide, with the other factors being (to simplify somewhat) democracy, internal discord, and external threat. All of the variables used in O’Reilly and Murphy (2022) can be interpreted (especially in the context of North et al. [2009]) as expressions of “effectiveness,” although “State Authority Over Territory” also captures other dimensions of institutions identified by Drumm (2015), namely external threats and internal discord.

An initial comparison of trends in the O’Reilly and Murphy (2022) measure and WGI appears as Figure 4, using WGI as a benchmark (although ICRG will be the primary point of comparison). Ninety-three of the 102 historically developing countries are included.¹⁰ In their original releases, both data sources are standardized to have a global mean of 0 and a standard deviation of 1, so Figure 4 reports them in those standardized units. The 1996 mean across the 93 countries is nearly identical in the O’Reilly and Murphy (2022) measure and WGI. The WGI mean shows little change over time, whereas the mean of the O’Reilly and Murphy (2022) measure exceeds the 75th percentile of WGI by 2010. The rise in the mean of O’Reilly and Murphy (2022) is accompanied by increases in its 75th percentile values, and the 75th percentile still appears to be rising; the value in 2023 is just short of its pre-COVID-19 high in 2019. Both the 25th percentile values of the O’Reilly and Murphy (2022) measure and WGI are stagnant, suggesting that the difference arises because the O’Reilly and Murphy (2022) measure observes countries graduating into higher levels of institutional quality, while showing no progress among many countries. On that point, there is agreement across data sources.

We can reproduce Table 2 from Pritchett (2020) in full using the O’Reilly and Murphy data. This appears below as Table 6, following all the same classification rules, but with the O’Reilly and Murphy data substituted for ICRG/QOG. One change was necessary, however. Simply following

¹⁰Nine countries were removed from all time series in this analysis because they had incomplete data in V-Dem as of 1996 (and came online later) or they have never been scored by V-Dem. These countries are The Bahamas, Brunei, Moldova, Papua New Guinea, Peru, Senegal, Sudan, Tanzania, and Uganda. These will also be subsequently excluded.

the procedure used by APW and Pritchett (2020) of placing countries on the 0-10 scale and using 6.5 as the minimum threshold of what constitutes the minimal value for “Strong” state institutions sets the threshold far too low. Pritchett (2020) highlights Spain and Czechia as two countries reflecting the lower bound of “Strong” state institutions, so we use the average score of those two countries in the O’Reilly and Murphy data, which yields a threshold of 9.31 out of 10 for demarcating “Strong” state institutions, rather than 6.5. We then set the threshold for “Moderate” and “Weak” institutions at three-quarters (that is, 6.98) and one-half (4.66) of 9.31, with countries scoring below 4.66 classified as having “Fragile” institutions.

To be clear about the differences in the sample, there were 102 countries in APW’s original Table 1.1 and 90 countries in Pritchett’s (2020) Table 2. There are 93 countries in Table 6. Of the nine countries that appear in APW but not in Table 6, three (The Bahamas, Brunei, and Moldova) do not appear in either Pritchett’s (2020) Table 2 or the O’Reilly and Murphy (2022) dataset. The remaining six are highlighted in blue in our Table 5. As noted earlier, nine countries appear in APW but not in Pritchett’s (2020) Table 2, and those are highlighted in blue in our Table 6. The difference between APW’s Table 1.1, our Table 6, and Pritchett’s (2020) Table 2 can be tracked by starting with APW’s 102 countries, subtracting the nine that do not appear in the O’Reilly and Murphy (2022) data set (leaving 93, as in Table 6). Six of those countries appear in Pritchett’s (2020) Table 2, bringing the count to 99. But there are nine additional countries that do not appear in Pritchett (2020), bringing the count back down to 90.

In Table 6, few countries match the exact square found in Table 5, only 10 of the 90. These appear in green. Of the 90, 31 countries are in a square adjacent to where they appear in Table 5. These appear in yellow. Nine appear in APW’s Table 1.1 but not in Pritchett’s (2020) Table 2. These appear in blue. The remaining 43 (46%) countries are at least two steps away from where they are located in Table 5. These appear in red. For comparison, when matching our Table 2 to APW’s Table 1.1, only one country, Algeria, was equally distant from its APW classification.

We bold countries that were on track as of 2018 to reach a “Strong” level of institutional quality by 2100, an exercise that is carried out by APW but not Pritchett (2020). A total of 23 countries are in bold in Table 6. These are the “Business as Usual” countries from the vantage point of O’Reilly and Murphy (2022): Albania, Angola, Colombia, Costa Rica, Côte d’Ivoire, Croatia, Ecuador, The Gambia, Guyana, Indonesia, Kenya, South Korea, Liberia, Malaysia, Myanmar, Qatar, Sierra Leone, Sri Lanka, Togo, Trinidad and Tobago, Tunisia, Uruguay, and Vietnam. In addition, Somalia technically reaches the threshold by 2100 as well, but is excluded because of its distance from the threshold.

For a handful of countries, the O’Reilly and Murphy data are more pessimistic, not less. Chile just misses achieving strong institutional status by 2100 with a predicted score of 9.25, just below the 9.31 threshold. *Varieties of Democracy* does not look as kindly as other measures do on Singapore, so Singapore is not over the threshold in 2018, with the country essentially projected to remain institutionally stagnant after 2018.¹¹ Botswana and Armenia are calculated to have a downward

¹¹ Singapore receives strong scores in Rigorous and Impartial Public Administration and Educational Equality. It receives good, but not great, scores in Particularist or Public Goods (perhaps due to lingering concerns over how the People’s Action Party continues to hold power), state fiscal source of revenue (presumably because of its direct ownership stakes in much of the economy, which provide it significant revenue), and the rule of law (where the definition for V-Dem differs from the rule of law needed to support a strong business environment). The sixth variable, state control over territory, is

trajectory rather than the upwards one they have in APW. South Korea and Costa Rica are also calculated to be on a downward trajectory, but on a slope so gentle that they remain classified as “Strong” in 2100.

While it does not affect any of the preceding findings, there is a “visual” bias in Table 6. Nine countries were omitted from Table 6 because of data availability but appeared in Table 1, which reproduces APW’s Table 1.1. The omission visually distorts the table because none of the nine appear in the upper-right quadrant of Table 1, so the omissions were systematic. These countries that were omitted are The Bahamas, Brunei, Moldova, Papua New Guinea, Peru, Senegal, Sudan, Tanzania, and Uganda. Our conclusions would not change if all of these countries were placed in Table 6 the same as they were placed in Table 1, but Table 6 would look visually different.

In Table 7, each of the “Business as Usual” countries shown in bold in Table 6 is listed alphabetically. To the right of each country, characterizations of each country’s level of institutional quality and trajectory are given, first as scored by O’Reilly and Murphy (2022), then by WGI, and finally by ICRG/QOG. The three measures agree for some countries, namely Colombia, Croatia, Indonesia, Liberia, Myanmar, Qatar, and Uruguay. In terms of trajectories, the O’Reilly and Murphy measure is most optimistic, followed by WGI, then ICRG/QOG, though there are exceptions. Focusing on instances where O’Reilly & Murphy (2022) and ICRG/QOG diverge most sharply concerning a trajectory of a country (e.g., one evaluates a country as “Positive,” while the other evaluates it as “Collapsing”), we identify twelve countries: Albania, Angola, Côte d’Ivoire, Ecuador, The Gambia, Guyana, Kenya, Sierra Leone, Sri Lanka, Togo, Trinidad and Tobago, and Tunisia. For the purposes of this exercise, we consider only trajectory and only ICRG/QOG in comparison to O’Reilly and Murphy.

These differences exist, but what underlies the O’Reilly and Murphy data? Table 8 details the underlying *Varieties of Democracy* data. In all cases, *Varieties of Democracy*, as operationalized by O’Reilly and Murphy (2022), offers the more optimistic view of trends. For three of the 12 (Angola, Sierra Leone, and Sri Lanka), the change in the O’Reilly and Murphy (2022) score can be substantively explained by the state gaining greater control of its territory. If we hew closely to Drumm (2015) and exclude these three as cases, we are left with nine countries where there is strong disagreement between ICRG/QOG and O’Reilly and Murphy (2022). All 12 countries saw increases in their “Particularist or Public Goods,” with nontrivial increases in most cases. Clear increases ($> +0.50$) in “Rigorous and Impartial Public Administration” occurred in Albania, Angola, Côte d’Ivoire, Ecuador, The Gambia, Kenya, Sierra Leone, Togo, and Tunisia; clear increases in “State Fiscal Source of Revenue” occurred in Albania, Kenya, Sierra Leone, and Togo; clear increases in “Educational Equality” occurred in Ecuador, The Gambia, Kenya, Sierra Leone, Togo, and Trinidad and Tobago; a clear increase in the V-Dem “Rule of Law” index occurred only in Tunisia, but all 12 countries had some positive gain on the variable. The experts surveyed by *Varieties of Democracy* may be mistaken about these countries, but they demonstrably believe governance is improving in these countries, whereas ICRG sees a negative trend or collapse.

ICRG/QOG has a long history of academic use and pre-dates the *Worldwide Governance Indicators* (e.g., Knack and Keefer 1995). V-Dem is relatively recent as a published data source, but it is one of the few data sources that provides broad pre-2000 country coverage. It goes well beyond the scope of this comment to assess conditions on the ground of these 12 countries, but the

only 85% as of 2018 and had been there for decades. It increases in 2019 and in 2022, and it is now over 99%. It is not clear what geopolitical events led to this change.

ICRG/QOG time series pattern shown above in Figure 1 is strange, to say the least. WGI currently uses V-Dem only on a limited basis, but it may warrant expanded use. At present, there is no overlap between the variables from V-Dem that WGI uses and the V-Dem variables that are applied by O'Reilly and Murphy (2022).

In responding to “The Big Stuck” hypothesis advanced by APW, we see evidence of confusion and ambiguity within the data, rather than evidence of stagnation or regression, whether the endpoint is 2012, 2018, or 2023. V-Dem and the Bertelsmann Transformation Index indicate that “Business as Usual” has continued 2000-present, whereas the trends in ICRG/QOG and the Fragile States Index are negative. The “Representative Sources” used by WGI other than ICRG/QOG appear to be, if anything, positive when we focus on a consistent sample of countries from 1996 to either 2012 or 2018. These comparisons are summarized in full in Table 9.

5 Conclusion

Pritchett (2020) shows that the pessimistic results found in APW were not contingent on the normalization of WGI data that precludes aggregate global improvement in institutional quality, as the ICRG/QOG data alone show a negative trend. However, we have seen that the ICRG/QOG data yielding this result come from a time series that appears flat from 1985 to 1991, increases dramatically from 1991 to 1995, decreases from 1995 to 2001, and is once again stagnant from 2002 forward. The regressions that draw a downward trend from 1996 to 2012 or 2018 are, as a group, driven by the unexplained downward trajectory from 1995 to 2001. Using ICRG start years outside the mid-1990s yields different results, as shown in Figure 2 and Figure 3 above.

Even if we take the global rise and decline in institutional quality across the 1990s, followed by stagnation, to be “real” rather than a statistical artifact or the result of a methodology change, such a pattern is not reproduced using the measure of state capacity developed by O'Reilly and Murphy (2022) from *Varieties of Democracy* data. The 12 countries whose trajectories are evaluated most differently by the O'Reilly and Murphy (2022) and ICRG/QOG data are Albania, Angola, Côte d'Ivoire, Ecuador, The Gambia, Guyana, Kenya, Sierra Leone, Sri Lanka, Togo, Trinidad and Tobago, and Tunisia. We find in the underlying data that the variables accounting for these increases chiefly concern governance quality, as shown in Table 8. There are 23 countries that are classified as “Business as Usual,” i.e., on pace for high quality governance by 2100, when assessed using the O'Reilly and Murphy measure.

We do not believe the evidence in favor of Business as Usual to be completely clear-cut and unambiguous. The 12 cases above are not all decisive evidence in favor of the O'Reilly and Murphy (2022) measure and V-Dem over QOG and ICRG. There is insufficient evidence to reject “The Big Stuck.” But overall, taking ICRG, the Bertelsmann Transformation Index, the Fragile States Index, and the O'Reilly and Murphy (2022) measure into account, we conclude the directional evidence is *ambiguous*, rather than strongly suggestive of stagnation.

References

Andrews, M., Pritchett, L., and Woolcock, M. (2017). *Building State Capability: Evidence, Analysis, Action*. Oxford: Oxford University Press. DOI: [10.1093/acprof:oso/9780198747482.001.0001](https://doi.org/10.1093/acprof:oso/9780198747482.001.0001).

Balcells, L. and Kalyvas, S. N. (2026). “Ideology and Revolution in Civil Wars: The Marxist Paradox.” *American Political Science Review*, vol. 120, no. 1, pp. 141–159. DOI: [10.1017/S0003055425000188](https://doi.org/10.1017/S0003055425000188).

Coppedge, M., Gerring, J., Knutsen, C. H., Lindberg, S. I., Teorell, J., Altman, D., Angiolillo, E., Bernhard, M., Borella, C., Cornell, A., Fish, M. S., Fox, L., Gastaldi, L., Gjerløw, H., Glynn, A., Good God, A., Grahn, S., Hicken, A., Kinzelbach, K., Marquardt, K. L., McMann, K., Mechkova, V., Neundorff, A., Paxton, P., Pemstein, D., Ryd'en, O., von R'omer, J., Seim, B., Sigman, R., Skaaning, S.-E., Staton, J., Sundstr'om, A., Tzelgov, E., Uberti, L., Wang, Y.-T., Wig, T., and Ziblatt, D. (2024). “V-Dem Codebook v14.” *Varieties of Democracy (V-Dem) Project, Codebook*. PDF: https://v-dem.net/documents/38/V-Dem_Codebook_v14.pdf.

Drumm, B. R. (2015). “Distinguishing Earth, Water, Fire, and Air: Factor Analysis to Determine the Four Fundamental Elements of State Capability.” B.A. thesis, Harvard College. Available at: <https://dash.harvard.edu/entities/publication/73120378-eb33-6bd4-e053-0100007fdf3b>.

Fergusson, L., Mejia, J., Robinson, J. A., and Torres, S. (2023). “Constitutions and Order: A Theory and Evidence from Colombia and the United States.” *NBER Working Paper*, no. 31501. DOI: [10.3386/w31501](https://doi.org/10.3386/w31501).

Hendrix, C. S. (2010). “Measuring State Capacity: Theoretical and Empirical Implications for the Study of Civil Conflict.” *Journal of Peace Research*, vol. 47, no. 3, pp. 273–285. DOI: [10.1177/0022343310361838](https://doi.org/10.1177/0022343310361838).

Hoti, S. and McAleer, M. (2004). “An Empirical Assessment of Country Risk Ratings and Associated Models.” *Journal of Economic Surveys*, vol. 18, no. 4, pp. 539–588. DOI: [10.1111/j.0950-0804.2004.00230.x](https://doi.org/10.1111/j.0950-0804.2004.00230.x).

Kaufmann, D. and Kraay, A. (2024). “The Worldwide Governance Indicators: Methodology and 2024 Update.” *World Bank Policy Research Working Paper*, no. 10952. DOI: [10.1596/1813-9450-10952](https://doi.org/10.1596/1813-9450-10952).

Kaufmann, D., Kraay, A., and Mastruzzi, M. (2006). “Governance Matters V: Aggregate and Individual Governance Indicators for 1996–2005.” *World Bank Policy Research Working Paper*, no. 4012. DOI: [10.1596/1813-9450-4012](https://doi.org/10.1596/1813-9450-4012).

Kaufmann, D., Kraay, A., and Mastruzzi, M. (2011). “The Worldwide Governance Indicators: Methodology and Analytical Issues.” *Hague Journal on the Rule of Law*, vol. 3, no. 2, pp. 220–246. URL: <https://link.springer.com/article/10.1017/S1876404511200046>.

Knack, S. and Keefer, P. (1995). “Institutions and Economic Performance: Cross-Country Tests Using Alternative Institutional Measures.” *Economics & Politics*, vol. 7, no. 3, pp. 207–227. DOI: [10.1111/j.1468-0343.1995.tb00111.x](https://doi.org/10.1111/j.1468-0343.1995.tb00111.x).

O'Reilly, C. and Murphy, R. H. (2022). “An Index Measuring State Capacity, 1789–2018.” *Economica*, vol. 89, no. 355, pp. 713–745. DOI: [10.1111/ecca.12411](https://doi.org/10.1111/ecca.12411).

Pritchett, L. (2020). “An Update in 2020 of the Big Stuck in State Capability.” Blog post. Available at: <https://lantpritchett.org/an-update-in-2020-of-the-big-stuck-in-state-capability/>.

Savoia, A. and Sen, K. (2015). “Measurement, Evolution, Determinants, and Consequences of State Capacity: A Review of Recent Research.” *Journal of Economic Surveys*, vol. 29, no. 3, pp. 441–458. DOI: [10.1111/joes.12065](https://doi.org/10.1111/joes.12065).

Țigănașu, R., Bănică, A., and Wong, P.-H. (2025). “Catalyzing Digital and Environmental Transformations by Institutions in a Diverse Socio-Economic World.” *World Development*, vol. 193, article 107052. DOI: [10.1016/j.worlddev.2025.107052](https://doi.org/10.1016/j.worlddev.2025.107052).

Table 1: Reproduction of Table 1.1 from Andrews et al. (2017)

	Trajectory of institutions, ICRG/QOG 1996–2012/2013			
	Rapid Negative < -0.05	Negative -0.05 < x < 0	Positive 0 < x < 0.05	Rapid Positive > 0.05
Strong (Capability > 6.5) (ICRG, QOG, FSI, 2012/2013)	–	BHR, BHS , BRN	CHL , SGP, KOR, QAT	ARE
Middle (4 < Capability < 6.5)	MDA , GUY, IRN, PHL, LKA, MNG, ZAF, MAR, THA, NAM, TTO, ARG, CRI	PER , EGY, CHN, MEX, LBN, VNM, BRA, IND, JAM, SUR, PAN, CUB, TUN, JOR, OMN, MYS, KWT, ISR	KAZ, GHA, UKR, ARM, RUS, BWA, IDN , COL , TUR , DZA , ALB , SAU , URY , HRV	–
Weak (2.5 < Capability < 4)	GIN, VEN, MDG, LBY, PNG , KEN, NIC, GTM, SYR, DOM, PRY, SEN , GMB, BLR	MLI, CMR, MOZ, BFA, HND, ECU, BOL, PAK, MWI, GAB, AZE, SLV	UGA , AGO, BGD, TZA , ETH, ZMB	–
Very Weak (Capability < 2.5)	YEM, ZWE, CIV	SOM, HTI, PRK, NGA, COG, TGO, MMR	SDN , SLE, ZAR, IRQ	NER, GNB, LBR

Source: Reproduced from Andrews et al. (2017, p. 20), Table 1.1. Cells contain ISO-3 codes for 102 historically developing countries. Rows reflect levels of institutional quality circa 2012–2013, using *Worldwide Governance Indicators*, the *Quality of Government* index, and the *Fragile States Index*, rescaled to 0–10. Columns reflect the trajectory of institutions as an average annual rate of improvement over 1996–2012 or 2013. Bold denotes “Business as Usual countries on pace to reach Strong institutions by the year 2100. Countries highlighted in blue will not appear in Section 4 of the paper for issues of data availability.

Table 2: Attempted Replication of Table 1.1 from Andrews et al. (2017)

	Trajectory of institutions, ICRG/QOG 1996–2012/2013			
	Rapid Negative < -0.05	Negative -0.05 < x < 0	Positive 0 < x < 0.05	Rapid Positive > 0.05
Strong (Capability > 6.5) (ICRG, QOG, FSI, 2012/2013)	-	BHS, BRN, ISR	CHL, KOR, QAT, SGP	ARE
Middle (4 < Capability < 6.5)	ARG, CRI, GUY, IRN, MDA, MNG, MAR, NAM, PHL, ZAF, LKA, THA, TTO	BHR, BRA, CHN, CUB, EGY, IND, JAM, JOR, KWT, LBN, MYS, MEX, OMN, PAN, PER, SUR, TUN, UKR, VNM	ALB, ARM, BWA, COL, HRV, GHA, IDN, KAZ, RUS, SAU, TUR, URY	-
Weak (2.5 < Capability < 4)	BLR, DOM, GMB, GTM, KEN, LBY, MDG, NIC, PNG, PRY, SEN, SYR, VEN	DZA, AZE, BOL, BFA, CMR, ECU, SLV, GAB, HND, MWI, MLI, MOZ, PAK, UGA	BGD, ETH, TZA, ZMB	-
Very Weak (Capability < 2.5)	CIV, GIN, SOM, YEM, ZWE	COG, HTI, PRK, MMR, NGA, TGO	AGO, IRQ, SLE, SDN, ZAR	GNB, LBR, NER

Results of attempted full replication of Andrews et al. (2017, p. 20), Table 1.1. Cells contain ISO-3 codes for 102 historically developing countries. Rows reflect levels of institutional quality circa 2012–2013, using *Worldwide Governance Indicators*, the *Quality of Government* index, and the *Fragile States Index*, rescaled to 0–10. Columns reflect the trajectory of institutions as an average annual rate of improvement over 1996–2012 or 2013. Bold denotes “Business as Usual countries on pace to reach Strong institutions by the year 2100. Countries highlighted in green are fully replicated. Countries highlighted in yellow are off by one degree in one dimension. Algeria (DZA), which is not highlighted, is off by one degree in both dimensions.

Table 3: *Worldwide Governance Indicators* Representative Sources from 1996 and Standardized Changes, 1996–2012

	Control of Corruption	Rule of Law	Government Effectiveness
PRS/ICRG	-0.791	-0.610	-0.132
<i>Economist</i> Intelligence Unit (EIU)	+0.044	+0.201	-0.159
Global Competitiveness Report (GCS)	-0.124	+0.239	+0.270
Varieties of Democracy (VDM)	-0.063	+0.233	N/A
S&P Global Country Risk Service (WMO)	+0.078	+0.094	+0.220

Calculated trajectories of consistent data underlying three of the *Worldwide Governance Indicators*. Average change is calculated in terms of each variable's standard deviation as of 1996. Only countries scored in both 1996 and 2012 by each individual variable are included in the respective samples.

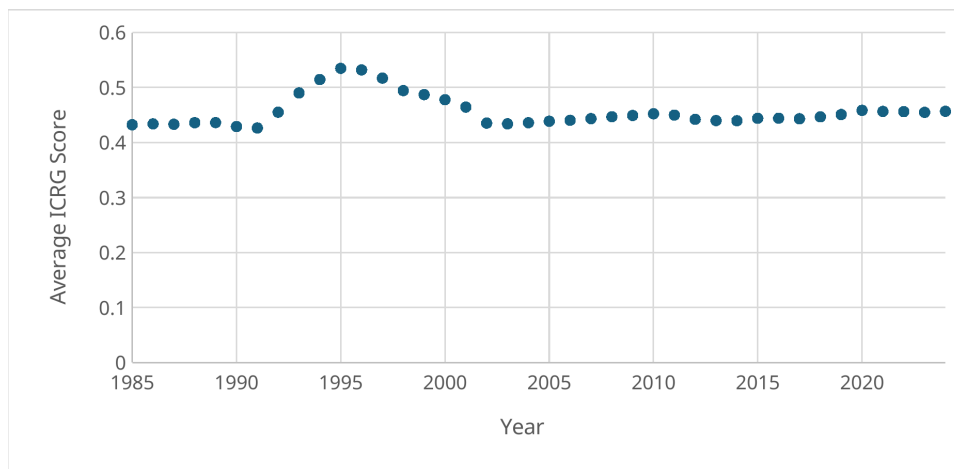


Figure 1: ICRG QOG, Average of 89 Historically Developing Countries, 1985–2023

Countries included are the 89 of the 102 countries that have *Quality of Government* scores derived from ICRG data going back to 1985. The group average of the 89 countries is taken for each year.

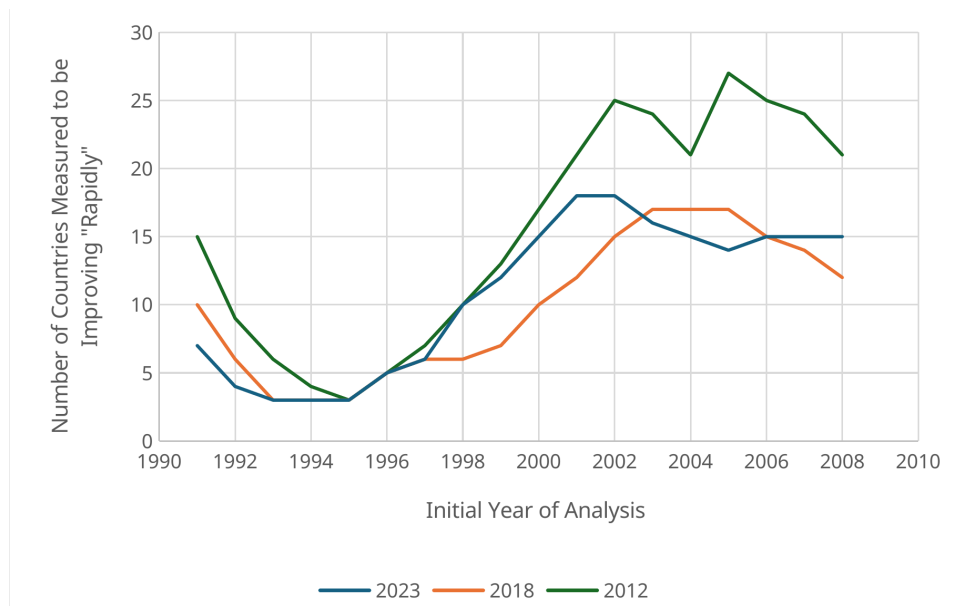


Figure 2: Sensitivity of Number of Countries Improving “Rapidly” to Initial Year Chosen, ICRG

The x-axis represents changing the starting year for the window considered in calculating the number of countries whose institutions are developing rapidly. Each time series corresponds to an ending year. The year used in Andrews et al. (2017) is 1996.

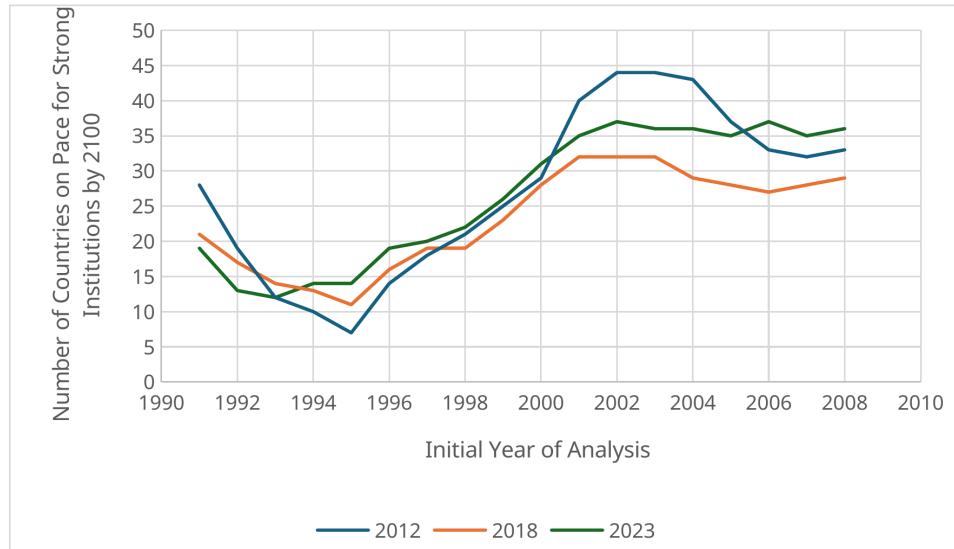


Figure 3: Sensitivity of Number of Countries on Pace for Strong Institutions by 2100 to Initial Year Chosen, ICRG

The x-axis represents changing the starting year for the window considered in calculating the number of historically developing countries whose institutions are projected to be “Strong” by the year 2100. Each time series corresponds to an ending year. The year used in Andrews et al. (2017) is 1996.

Table 4: *Worldwide Governance Indicators* Representative Sources from 1996 and Standardized Changes, 1996–2018

	Control of Corruption	Rule of Law	Government Effectiveness
PRS/ICRG	-0.695	-0.638	-0.086
<i>Economist</i> Intelligence Unit (EIU)	+0.278	+0.344	+0.058
Global Competitiveness Report (GCS)	-0.068	+0.154	+0.303
Varieties of Democracy (VDM)	-0.150	+0.166	N/A
S&P Global Country Risk Service (WMO)	-0.244	+0.335	+0.601

Calculated trajectories of consistent data underlying three of the *Worldwide Governance Indicators*. Average change is calculated in terms of each variable's standard deviation as of 1996. Only countries scored in both 1996 and 2018 by each individual variable are included in the respective samples.

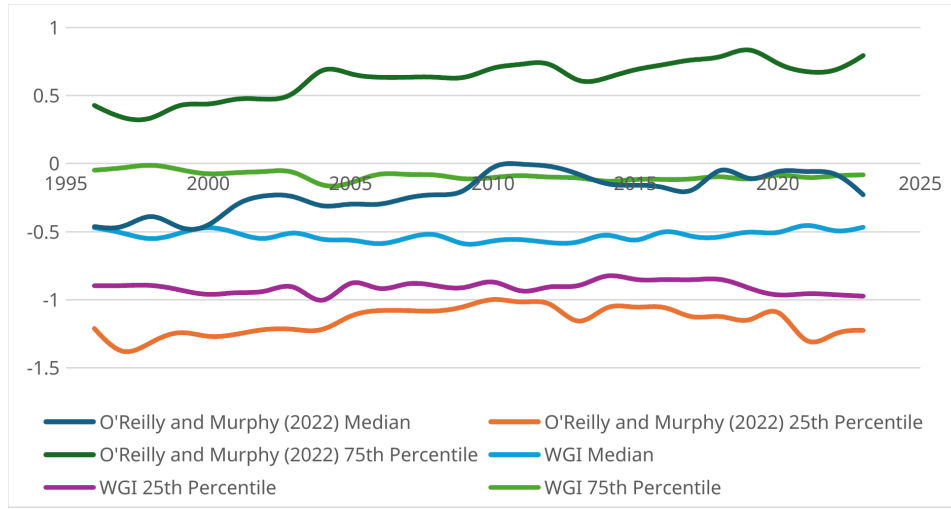


Figure 4: Comparison of Distribution of *Worldwide Governance Indicators* (WGI) and O'Reilly and Murphy (2022), 1996–2023

A comparison of the 25th percentile, median, and 75th percentile values of O'Reilly and Murphy (2022) data on state capacity and *Worldwide Governance Indicators* (WGI). Here, "WGI" is the average of WGI Control of Corruption, Rule of Law, and Government Effectiveness indicators.

Table 5: Reproduction of Table 2 from Pritchett (2020), ICRG Data, Level in 2018, Trajectory 1996–2018

Level of ICRG, 2018	Rate of Change in Standardized ICRG Data, 1996–2018			
	Collapsing ($g < -0.05$)	Slow Negative ($-0.05 < g < 0$)	Slow Positive ($0 < g < 0.05$)	Rapid Positive ($g > 0.05$)
Strong (Capability > 6.5)	–	KOR, SGP	CHL	IDN, ARE
Moderate (4 < Capability < 6.5)	GMB, CRI, LKA, ARG, IRN, TTO, MAR, NAM	EGY, GTM, PAK, PAN, PNG, UGA, BOL, MWI, TUR, KAZ, CHN, PER, TZA, CUB, JAM, MNG, VNM, JOR, KWT, PHL, BHR, TUN, MYS, OMN	DZA, LBN, ETH, GHA, BWA, URY, IND, QAT, SAU	–
Weak (2.5 < Capability < 4)	DOM, NIC, MDG, ZWE, PRY, SLV, MEX, GIN, LBY, ECU, THA, ZAF, KEN, SYR	CMR, SLE, AGO, AZE, SEN, BRA, GAB, HND, BFA, MOZ, ZMB	ARM, MMR, BGD, COL	GNB, NER
Fragile (Capability < 2.5)	SOM, HTI, VEN, CIV, COG, YEM, PRK	SDN, MLI, TGO, NGA	LBR	IRQ

Reproduced from Pritchett (2020), Table 2. Cells contain ISO-3 codes for 90 historically developing countries. Rows reflect levels of institutional quality circa 2018, using the *Quality of Government* index, rescaled to 0–10. Columns reflect the trajectory of institutions as an average annual rate of improvement over 1996–2018. Countries highlighted in blue do not appear in Table 6.

Table 6: Table 2 of Pritchett (2020), ICRG Data Replaced with O'Reilly and Murphy (2022) Data

Level of O'Reilly and Murphy, 2018	Rate of Change in Standardized O'Reilly and Murphy Data, 1996–2018			
	Rapid Negative ($g < -0.05$)	Negative ($-0.05 < g < 0$)	Positive ($0 < g < 0.05$)	Rapid Positive ($g > 0.05$)
Strong (Capacity > 9.31)	–	CRI, KOR	URY	–
Middle (6.98 < Capacity < 9.31)	–	ARM, BWA, ISR, MNG, NAM, ZAF	ARG, BLR, BRA, CHL, GMB, GUY, JAM, KWT, MYS, OMN, QAT, SGP, LKA, TTO, ARE	ALB, CIV, HRV, IDN, KEN, SLE, TUN
Weak (4.66 < Capacity < 6.98)	TUR, ZWE	BGD, BOL, CHN, CUB, DOM, GAB, HND, IRN, KAZ, MEX, MOZ, SUR, THA, ZMB	DZA, BFA, COG, EGY, SLV, ETH, GHA, GTM, IND, JOR, LBN, MWI, MLI, MAR, NER, NGA, PAN, PRY, PHL, RUS, SAU, UKR, VNM	AGO, COL, ECU, LBR, MMR, TGO
Very Weak (Capacity < 4.66)	LBY, SYR, VEN, YEM	AZE, BHR, HTI, IRQ, MDG, NIC	CMR, ZAR, GIN, GNB, PRK, PAK	SOM

Attempt at reproducing Pritchett (2020) using O'Reilly and Murphy (2022) as an alternative data source. Cells contain ISO-3 codes for 93 historically developing countries. Rows reflect levels of institutional quality circa 2018, using the O'Reilly and Murphy (2022) dataset, rescaled to 0–10. Columns reflect the trajectory of institutions as an average annual rate of improvement over 1996–2018. Countries highlighted in green match Table 5. Countries highlighted in yellow are one unit distance away. Countries highlighted in red are two categorical distances away. Countries highlighted in blue do not appear in Table 5.

Table 7: Select Differences Between WGI and ICRG in Pritchett (2020) and O'Reilly and Murphy (2022) Concerning 1996–2018

Country	O'Reilly & Murphy	WGI	ICRG
Albania*	Rapid Positive / Moderate Capability	Positive Trend / Moderate Capability	Negative Trend / Moderate Capability
Angola	Rapid Positive / Weak Capability	Positive Trend / Weak Capability	Negative Trend / Weak Capability
Colombia	Rapid Positive / Weak Capability	Positive Trend / Moderate Capability	Positive Trend / Weak Capability
Costa Rica	Negative Trend / Strong Capability	Negative Trend / Strong Capability	Collapsing Trend / Moderate Capability
Côte d'Ivoire	Rapid Positive / Moderate Capability	Positive Trend / Weak Capability	Collapsing Trend / Fragile
Croatia*	Rapid Positive / Moderate Capability	Rapid Positive / Moderate Capability	Negative Trend / Moderate Capability
Ecuador	Rapid Positive / Weak Capability	Positive Trend / Weak Capability	Collapsing Trend / Weak Capability
The Gambia	Positive Trend / Moderate Capability	Negative Trend / Weak Capability	Negative Trend / Weak Capability
Guyana*	Positive Trend / Moderate Capability	Negative Trend / Weak Capability	Collapsing Trend / Weak Capability
Indonesia	Rapid Positive / Moderate Capability	Rapid Positive / Moderate Capability	Positive Trend / Moderate Capability
Kenya	Rapid Positive / Moderate Capability	Positive Trend / Weak Capability	Collapsing Trend / Weak Capability
Liberia	Rapid Positive / Weak Capability	Rapid Positive / Weak Capability	Positive Trend / Fragile
Malaysia	Positive Trend / Moderate Capability	Positive Trend / Moderate Capability	Negative Trend / Moderate Capability
Myanmar	Rapid Positive / Weak Capability	Positive Trend / Weak Capability	Positive Trend / Weak Capability
Qatar	Positive Trend / Moderate Capability	Rapid Positive / Moderate Capability	Positive Trend / Moderate Capability
Sierra Leone	Rapid Positive / Moderate Capability	Positive Trend / Weak Capability	Negative Trend / Weak Capability
Somalia	Rapid Positive / Fragile	Negative Trend / Fragile	Collapsing / Fragile
South Korea	Negative Trend / Strong Capability	Positive Trend / Strong Capability	Negative Trend / Strong Capability
Sri Lanka	Positive Trend / Moderate Capability	Negative Trend / Moderate Capability	Collapsing Trend / Moderate Capability
Togo	Rapid Positive / Weak Capability	Negative Trend / Weak Capability	Negative Trend / Fragile
Trinidad & Tobago	Positive Trend / Moderate Capability	Collapsing Trend / Moderate Capability	Collapsing Trend / Moderate Capability
Tunisia	Rapid Positive / Moderate Capability	Negative Trend / Moderate Capability	Negative Trend / Moderate Capability
Uruguay	Positive Trend / Strong Capability	Positive Trend / Strong Capability	Positive Trend / Moderate Capability
Vietnam	Positive Trend / Weak Capability	Positive Trend / Moderate Capability	Negative Trend / Moderate Capability

*Does not appear in Pritchett (2020), but reconstructed retroactively.

Table 8: Underlying V-Dem Data for Countries with Large Differences of Trajectory between V-Dem and PRS/ICRG

Country	Year	Albania	Angola	Côte d'Ivoire	Ecuador	The Gambia	Guyana
Particularist or Public Goods	1996	-0.44	-1.96	-0.04	-1.10	+0.33	-0.82
	2018	+0.41	-0.39	+2.07	+0.53	+1.65	+0.92
Rigorous & Impartial Public Administration	1996	+0.10	-1.08	+0.41	-0.60	-0.91	+0.87
	2018	+0.70	-0.07	+1.36	+0.31	+0.93	+0.69
Control Over State Territory	1996	89.0%	63.3%	94.6%	99.8%	97.2%	99.5%
	2018	95.9%	90.8%	89.4%	95.4%	97.2%	99.4%
State Fiscal Source of Revenue	1996	-0.71	-0.07	+0.70	+0.49	-0.40	+1.23
	2018	+0.83	-0.07	+0.88	+0.49	-0.10	+1.23
Educational Equality	1996	+0.96	-2.00	-0.97	-0.48	-0.49	+0.68
	2018	+1.42	-1.31	-0.78	+0.86	+0.14	+0.68
Rule of Law	1996	+0.50	+0.17	+0.45	+0.39	+0.29	+0.62
	2018	+0.55	+0.34	+0.56	+0.52	+0.77	+0.66

Country	Year	Kenya	Sierra Leone	Sri Lanka	Togo	Trinidad & Tobago	Tunisia
Particularist or Public Goods	1996	-0.21	+1.66	+0.56	+1.14	+0.42	+0.45
	2018	+0.60	+2.11	+1.06	+2.14	+1.35	+2.45
Rigorous & Impartial Public Administration	1996	-1.02	-0.45	+0.21	-0.50	+0.87	-0.78
	2018	+0.94	+1.74	+0.18	+0.14	+1.30	+0.84
Control Over State Territory	1996	90.8%	50.9%	71.7%	87.2%	98.5%	91.5%
	2018	92.5%	93.5%	100.0%	83.8%	95.7%	95.2%
State Fiscal Source of Revenue	1996	+1.23	-0.53	+1.68	-0.21	+1.14	+0.95
	2018	+1.79	+0.11	+1.41	+1.98	+1.14	-0.32
Educational Equality	1996	+0.01	-1.52	+0.68	+1.09	+1.15	+0.54
	2018	+0.57	-0.38	+0.94	+1.61	+1.89	+0.84
Rule of Law	1996	+0.23	+0.24	+0.60	+0.15	+0.86	+0.16
	2018	+0.61	+0.50	+0.68	+0.24	+0.90	+0.81

Countries selected by largest differences in assessed trajectory according to the *Quality of Government* index versus O'Reilly and Murphy (2022).

Table 9: Summary of Assessed Trajectories of Institutions by Data Source

Data source	Trajectory
<u>Variables Used by Pritchett (2020)</u>	
<i>Worldwide Governance Indicators</i>	Flat (by construction)
<i>Fragile States Index</i>	Negative
<i>Quality of Government/International Country Risk Guide</i>	Negative (with caveats)
<i>Bertelsmann Stiftung Transformation Index</i>	Positive
<u>Other Underlying Worldwide Governance Indicators Variables</u>	
<i>The Economist Intelligence Unit (EIU)</i>	Positive
<i>Global Competitiveness Report</i>	Mixed
<i>Varieties of Democracy</i> (select variables)	Mixed
S&P Global Country Risk Service	Mixed
<u>Novel Variables</u>	
O'Reilly and Murphy (2022)	Positive