

## What can we learn by collecting more complete UNGA voting records?

Joshua Fjelstul†

Simon Hug‡

Christopher Kilby\*

September 30, 2021

### Abstract:

Existing databases for UNGA voting (e.g., Bailey *et al.* 2017) provide excellent coverage of United Nations General Assembly voting on resolutions that were adopted by roll call vote. These databases, however, have some known limitations: They were built on historical dataset (e.g., the ICPSR dataset) that have varying coverage and definitions. They do not cover resolutions adopted by consensus. They do not consistently cover votes on draft resolutions that failed, i.e., votes on resolutions that were not adopted, or votes leading up to consideration by the UNGA. In this paper, we present an updated dataset that attempts to document and address these shortcomings, as well as explore the implications of more complete data for existing and future research.

† University of Geneva. [Joshua.Fjelstul@unige.ch](mailto:Joshua.Fjelstul@unige.ch)

‡ University of Geneva. [Simon.Hug@unige.ch](mailto:Simon.Hug@unige.ch)

\* Villanova University. [chkilby@yahoo.com](mailto:chkilby@yahoo.com)

## **I. Introduction**

A wide range of scholars have used United Nations General Assembly (UNGA) voting information for a variety of purposes. Some scholars study the UNGA itself, asking what major shifts in the international order imply for the functioning of the UN and what voting patterns in the UN reveal about such shifts (e.g., Bailey and Voeten 2018; Voeten 2004). Voting is also used to capture alignment between countries, with some asking whether countries trade votes for financial or other material advantages (e.g., Dreher and Jensen 2013, Dreher and Sturm 2012) and others exploring the role of alignment more generally (Alesina and Dollar 2000; Ball and Johnson 1996; Barro and Lee 2005; Dreher and Jensen 2007; Thacker 1999; Wang 1999). According to Voeten (2013), between 1998 and 2013, over 50 published studies use UN voting data; the number of such studies has surely more than doubled since that time.

Apart from studies that seek to understand shifts in behavior at the UN over time, a major reason to use UNGA voting behavior when studying international political economy is that the data are available over a long period of time and for essentially all countries. What exactly should be included in UNGA voting measures has received some attention. Wang (1999) and Thacker (1999) highlight the set of votes that the United States State Department designates as “key votes” to the U.S.; Andersen *et al.* (2006) suggest that differences between alignment on these key votes and on other votes reflect concessions to the U.S. position and hence may be useful for capturing vote buying. Häge and Hug (2016) include measures adopted without recorded roll call votes (e.g., consensus measures) and demonstrate that this can generate important differences in preference similarity measures. Beyond this, relatively little attention has been paid to which decisions go into the making of alignment or ideology indices constructed from UNGA data.

The data used in the majority of studies come from one of two main lineages. Some scholars assemble data from the UN website, the current version of which is the United Nations Digital Library. With the exception of a few small errors and omissions, this site provides voting information resolutions adopted by the UNGA. It does not, however, include draft resolutions not adopted (i.e., voted down), nor does it include information on the drafting process, such as committee decisions, amendments, and paragraph votes. The second lineage builds on a dataset developed by generations of scholars, most recently under the direction of Erik Voeten. These data include some draft resolutions not adopted, as well as some amendment and paragraph votes. While this dataset is richer than that directly available from the UN Digital Library, it reflects its history of successive stewardship with varying definitions of coverage.

Our paper presents an attempt to construct a broader, deeper, and more internally consistent set of data on decisions at the UNGA. We draw on UN records (the UN Digital Library, UNGA session indices, and UNGA minutes) as well as early ICPSR data to construct a new dataset with more clearly defined inclusion criteria and broader, more consistent coverage. The paper first describes the existing datasets (identifying both strengths and shortcomings). We then present our data collection methodology in detail. Next, we compare the resulting dataset with the two alternative datasets, highlight differences session-by-session and also at the individual vote level. We then present several replication studies to demonstrate the impact of using a more consistent and complete dataset for constructing measures of alignment and similarity. The paper concludes with a summary and observations for future research.

## **II. Existing datasets**

In order to situate and characterize the various existing datasets on voting in the UNGA, it is useful to start with clear definitions of the various ways in which the UNGA makes decisions related to

resolutions.<sup>1</sup> Bailey (1960), drawing on the *Rules of Procedures of the UNGA*, identifies several decision-making methods, to which he adds to additional ones that are not explicitly envisioned but used extensively. First, voting can take place by show of hands or alternatively by standing up.<sup>2</sup> In either case, final tallies of the yes- and no-votes, as well as abstentions, are counted. Alternatively, votes may occur by roll-call in which each member-state delegate is called in alphabetical order to state her vote. In addition to these voting methods in the *Rules of Procedures*, Bailey (1960, 154) identifies two more used in practice: decision by acquiescence and voting by mechanical means.<sup>3</sup>

For votes using the first two methods (hand or stand), the minutes of the UNGA (and other sources provided by the UN, e.g., the digital library) information on the number of yes- and no-votes, as well as abstentions, is available. This information is also available for the votes based on roll-call votes; in addition, information on each member-states delegate's voting decision appears in the minutes and is reported on various UN websites. As one of the method not envisioned by the *Rules of Procedure* according to Bailey (1960, 154), namely mechanical voting, can be used both for roll-call votes and votes by “show of hands” (or standing), information on voting decisions by delegates may be available for some “mechanical votes” but not for others. The other method not explicitly mentioned in the *Rules of Procedure* (even today) is the decision by acquiescence.

---

<sup>1</sup> Thus, we do not discuss other decisions reached in the UNGA, like elections etc. For a fuller discussion of these methods of decision-making see Bailey (1960, 151ff). It is useful to note that Bailey (1960) separates this presentation of decision-making procedures from decisions on the agenda (see also below).

<sup>2</sup> On this latter option, Bailey (1960, 152) states: “I know of no case in which the Assembly or one of the committees has actually used this method.”

<sup>3</sup> While Bailey (1960) provides a detailed discussion of these voting methods in place at the time of writing, it is useful to note that the only substantive difference to the current *Rules of Procedure* (see <https://www.un.org/en/ga/about/ropga/plenary.shtml>, accessed September 30, 2021) concerns the explicit reference to votes by “mechanical means.”

The UN itself provides the following characterization

(<https://research.un.org/c.php?g=98268&p=636558> , accessed September 30, 2021) of voting methods used when resolution-related matter is decided upon in the UNGA:

The majority of General Assembly resolutions are adopted without a vote. If a vote is taken, it can be documented in two ways: either as a recorded vote or as a summary of the result. Only a recorded vote, which must be requested before the voting is conducted, will clearly identify the stand that a Member State took on the issue under discussion. If such a request is not put forth, only the voting summary (i.e., the number of countries which voted for or against a resolution as well as those who abstained) will be made available, without identification of how an individual Member State voted.

Based on this, for instance, the UN Digital library provides for every resolution adopted in the plenary of the UNGA the information whether it was adopted in a recorded (i.e., roll-call) or unrecorded (i.e., show of hands and by standing, respectively the corresponding mechanical vote) vote, or adopted without a vote (by acquiescence in Bailey's (1960, 152) terms). This is also the way in which Hovet (1960, 14f) and Peterson (2005, 54) categorize the voting methods, and we will rely on it as well to characterize the existing datasets on UNGA voting.

With these concepts established, we turn to existing datasets. The Inter-University Consortium for Political and Social Research (ICPSR) datasets originated with a 1967 study of UN roll call voting by Charles Wrigley. The ICPSR provides updated versions of this data, first drawing on U.S. State Department records and then validated with UN official records. As a result, there are multiple archived datasets that are based on different inclusion criteria. The most comprehensive

one contains all recorded votes from plenary meetings for the first 26 sessions of the UNGA.<sup>4</sup> Three additional datasets provide the same information based on the identical selection criteria for the 27<sup>th</sup>, 28<sup>th</sup>, and 29<sup>th</sup> plenary sessions. Eleven additional datasets cover plenary sessions 30 to 40, but only report recorded final passage votes on draft resolutions (excluding other recorded votes). An additional dataset covers the first five special and emergency special sessions; this dataset contains again information on all recorded votes, whether on draft resolutions or on resolution-related matters or on other business. This same file also contains the same information on recorded votes from five sessions that were joint meetings of more than one of the main committees. These include one joint session of committees 1 and 6 and four joint sessions of committees 2 and 3. Finally, for the first 20 UNGA sessions there are also six datasets covering the recorded votes having occurred in the six main committees (meeting individually rather than jointly). It is important to note that these main UNGA committees are all committees of the whole (see Peterson 2005, 60).

These datasets were subsequently updated and corrected by Kim and Russett (1996) and Gartzke and Jo (2006) for their respective studies drawing on recorded votes at the UNGA.

Erik Voeten and collaborators built on ICPSR data through 1984 (with corrections derived from Gartzke and Jo (2002) and Kim and Russett (1996)) to extend the dataset by drawing on the United Nations Bibliographic Information System (UNBIS) and various official documents. Nonetheless, several holes in the dataset persist. Voeten notes that “ICPSR (and Gartzke/Jo) data miss votes on amendments since 1975, votes on paragraphs since 1978, and votes on unsuccessful resolutions since 1985.” Looking at Voeten data (published May 6, 2021), these data remain incomplete. No

---

<sup>4</sup> We note already here that a large majority of these votes are on draft resolutions and related matter, but also includes votes that are not related to any draft resolution debated by the UNGA.

amendments are reported for 27 sessions since 1974: 1975 (Session 30), 1976 (Session 31), 1980 (Session 35), 1981 (Session 36), 1984-1996 (Sessions 39-51), 2003 (Session 58), 2005 (Session 60), 2008 (Session 63), 2009 (Session 64), 2013-2015 (Sessions 68-70), and 2018-2020 (Sessions 73-75). In addition, for the years after 1975 that do report amendment votes, amendments account for 1.5% of all votes; for comparison, prior to 1975, they account for 17% of all votes. This suggests that even for sessions covered after 1974, not all amendments have been identified. Repeating this analysis with paragraph votes, the percentage of paragraph votes drops notably after 1975 (Session 30), with no paragraph votes recorded for 18 sessions since 1978 (1985-1996 (Sessions 40-51), 2013-2015 (Sessions 68-70), and 2018-2020 (Sessions 73-75)). There is good coverage for 1997-2012 (Sessions 52-67) and 2016-2017 (Sessions 71-72). Using a simple majority rule for passage, the Voeten data show no failed resolutions, amendments or paragraph votes between 1975 (Session 30) and 1996 (Session 51), as well as scattered years since then, including a number of years that also list no amendment votes (2003, 2005, 2013-2015, and 2018-2020). While failed resolutions have become less common, this pattern still suggests that a second look at the data is warranted.

Discuss Hug data, too.

### **III. Data collection strategy**

dataset We start our data collection process by developing a clear set of inclusion criteria. First, we only include resolution-related decisions in the General Assembly, and therefore exclude decisions on procedural issues that cannot be tied to a particular resolution. For instance, we would not include a decision on a motion with respect to the agenda, which occur at the start of each session,

while we would include a motion to adjourn during the discussion of a resolution.<sup>5</sup> Resolution-related decisions fall into four categories: decisions on the final passage of a resolution; decisions on amendments to draft resolutions; decisions on paragraphs or other sub-components of draft resolutions, which are also called separate votes; and decisions on procedural motions directly related to a resolution.<sup>6</sup>

We do not currently include decisions in meetings of the six main committees. Decisions on resolutions in these main committees are confirmed (or not) in the plenary, so excluding these decisions for resolutions adopted in the main committees eliminates largely redundant information, as the same decisions are made in the plenary. This is, however, not the case with respect to draft resolutions and other items that failed in committee, as these are normally not decided on by the plenary. Thus, in the next iteration of our data collection, we will add information on decisions on approximately 60 negative decisions on draft resolutions.

Second, we include decisions taken without a vote, by a non-recorded vote (i.e., a vote where the total number of yeses, noes, and abstentions is reported, but not each country's vote), or by a recorded vote (i.e., a vote in which each country's vote is reported). We include successful votes and unsuccessful votes.

Third, we cover the period from the start of the first session in 1946 through the end of the 75th session in 2021, including all decisions in the 75 ordinary sessions and all decisions in all special

---

<sup>5</sup> Some of our sources also include the former for some periods, so we provide an easy way to merge them with our data (for researchers also interested in these non-resolution-related votes).

<sup>6</sup> Decisions on motions may concern innocuous things, like adjournments, or weightier topics, like declaring a draft resolution as important and thus requiring a two-thirds majority in the final passage vote.



sessions and emergency special sessions. We plan to update the data regularly using an automated process. Updates will also be hand-validated before publication to ensure accuracy.

Given our inclusion criteria, there is no single source that we can rely on for all time periods. While the minutes of the UNGA (officially called meeting records) in principle contain all the information we need and are (with a few exceptions) all available electronically either from the UN Official Document System (ODS) or the UN Digital Library, it would be prohibitively expensive, in terms of time and effort, to extract the data we are interested in only from the minutes.<sup>7</sup>

Therefore, we adopt a multi-pronged data collection strategy. For every observation in our data, we report the source and method (e.g., automated parsing, hand-coding) that we use to code the information. For every decision, we report the ID number of the meeting record so users can track down the vote in the minutes if need be.<sup>8</sup> For decisions relating to resolutions that are adopted, we report the resolution ID number and the ID number for the corresponding draft resolution.<sup>9</sup> For decisions relating to draft resolutions, we report the ID number of the draft.

Based on high quality pdfs for sessions 49-75, we scrape and parse the UNGA minutes for these sessions to extract information about all resolution-related decisions, including both decisions taken without a vote and decisions with a recorded vote (note that there are no resolution-related

---

<sup>7</sup> There are several reasons for this. First, digital PDFs, which are more accurately parsed than scans of paper copies, are only available from around 1994. The quality of the scans for earlier documents varies widely, even within a session. The quality of the OCR done by the UN also varies widely. For some documents, the OCRed text is not usable. Even with the best-available tools, OCR is more accurate for some documents than others. Second, the format of the minutes changes periodically and some documents are more structured, in terms of their formatting, than others. Consequently, it is easier to accurately extract information on resolution-related decisions from the minutes for some periods than for others.

<sup>8</sup> Meeting record ID numbers are in the format A/PV.# or A/#/PV.#, depending on the session.

<sup>9</sup> Resolution ID numbers are in the format A/RES/# or A/RES/##/#, depending on the session.

decisions take by a non-recorded vote during this period). The data on recorded votes include how each country voted.<sup>10</sup>

Future updates to the data will rely on the minutes (as all the information we need is available there). The minutes-based data on final passage recorded votes will also be cross-checked against UN Digital Library data on final passage votes (see below) to ensure accuracy. By cross-validating final passage votes that we extract from minutes, we can be more confident that we also accurately extract amendment, motion, and paragraph votes from minutes. None of the latter votes are available from other sources but fortunately their format in the minutes matches that for final passage votes.<sup>11</sup> See appendix for details on the process we use for automated scraping of data from minutes.

Second, we use vote data from the UN Digital Library, which has (almost) complete voting records for all resolutions that are adopted by the General Assembly (sessions 1-75) that indicate whether the resolution was adopted without a vote, adopted with a non-recorded vote, or adopted with a recorded vote. If the resolution is adopted with a recorded vote, the UN Digital Library record for that resolution also includes information about how each country voted. We scrape the country-level vote data for recorded votes and use this source for all final passage votes. For sessions 49-

---

<sup>10</sup> We can cleanly and accurately extract the text of the minutes that are available as digital PDFs (i.e., PDFs that do not require OCR). Digital PDFs are consistently available starting with the 49<sup>th</sup> session. There are a small number of meeting records in sessions 49-75 that are not available as digital PDFs. When these minutes contain votes, we code them by hand.

<sup>11</sup> For this purpose, we developed a parsing algorithm in R that extracts the relevant information for all minute formats used in sessions 49-75. The entire process, from downloading the minutes to the creation of the final data, is fully automated and fully replicable. We apply a number of validation checks (see below) at each step of the process to identify parsing errors caused by formatting inconsistencies or errors in the minutes and correct any errors in a replicable way.

75, we use this data to cross-check the data on final passage votes that we extract from the minutes (described above).

Third, we rely on ICPSR datasets for sessions 1-29.<sup>12</sup> ICPSR data for this period contain (almost) all recorded votes, including final passage votes and non-final passage votes (votes on amendments, votes on motions, and paragraph votes).<sup>13</sup>

Fourth, since the ICPSR data only covers recorded votes, we hand code all resolution-related decisions in sessions 1-48 that were adopted by an unrecorded vote (e.g., by a voice vote or a show of hands) or that were adopted without a vote.<sup>14</sup>

Finally, using the indices to proceedings of the UNGA, we identify all resolutions that failed in ordinary, special, or emergency special sessions. For all such votes that are recorded votes, we collect the country-level vote data by hand from the minutes, provided they are not already covered in the ICPSR data.

We validate the data by cross-checking information available from multiple sources. We also check for internal consistency where possible. All corrections that we make to the data we extract from the minutes or the data we scrap from the UN Digital Library (which are both automated processes) are done via an R script so the entire process remains fully automated and fully replicable. For details on the validation process, see the appendix.

---

<sup>12</sup> The ICPSR archive contains voting data for sessions 1-40, but the inclusion criteria changed with session 30. Starting with that session, ICPSR data only include information on recorded final passage votes that led to the adoption of a resolution.

<sup>13</sup> We identified five final passage votes on resolutions not covered in the ICPSR data and add these. In addition, the ICPSR data also contains recorded votes on things that are not related to resolutions. In line with our inclusion criteria, we remove these votes from the ICPSR data. For researchers interested in these votes, we provide a way to add them to our dataset.

<sup>14</sup> For sessions 49-75, we automatically extract this information from the minutes, as described above, and validate it.

#### IV. Comparative descriptors

In this section, we compare our data to existing UNGA voting data collections, chiefly that provided by Bailey *et al.* (2017). For this dataset, we use the file “UNVotes.csv” file on Erik Voeten’s Harvard Dataverse page, dated May 6, 2021, which contains raw vote data for UNGA sessions 1 to 75. This data file contains votes on resolutions, amendments, and paragraphs; these include both cases where the measure passed (the “yes” vote tally exceeds the “no” vote tally) and where the measure failed (the “yes” vote tally does not exceed the “no” vote tally).<sup>15</sup>

Figure 1 provides a comparison of the number of recorded plenary votes per UNGA session in the Voeten dataset (top panel) and our dataset (FHK; bottom panel) broken out by amendment, paragraph, and resolution categories.

---

<sup>15</sup> For measures designated as important, a two-thirds majority is needed for passage. Since the Voeten data do not indicate which items were important, we use a simple majority rule for both datasets for this exercise. UN procedures specify that a tied vote (in some cases after a second try) means the measure does not pass. Also note that Voeten includes motions in the resolutions category so we do the same for this exercise.

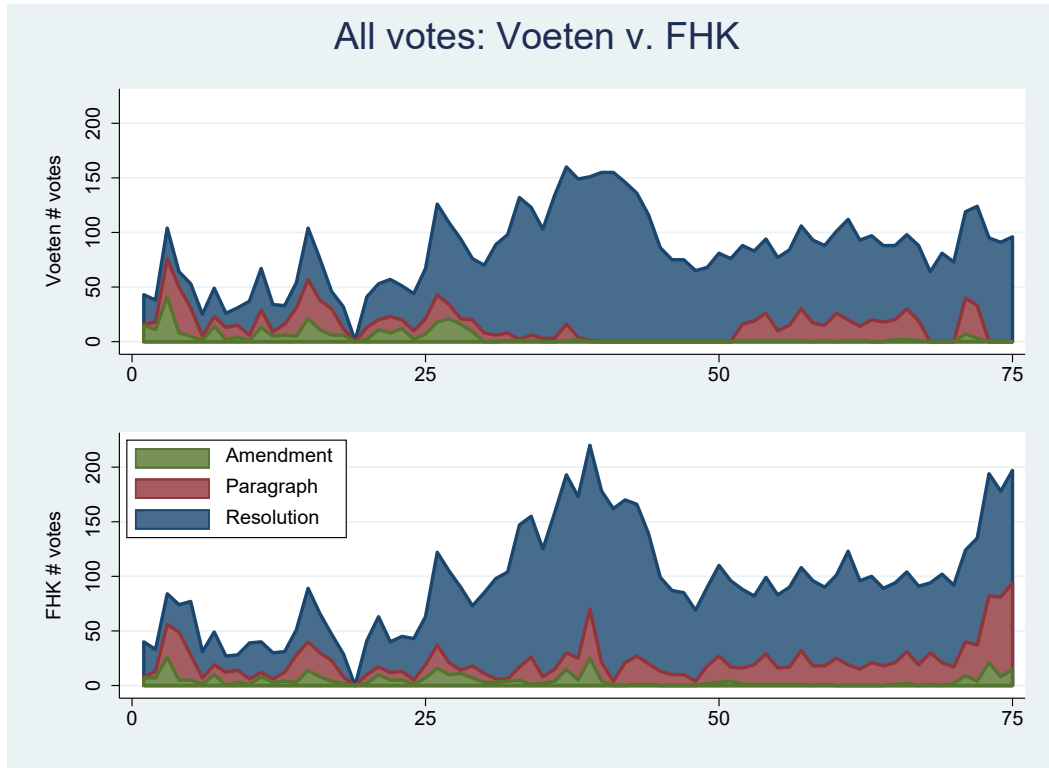


Figure 1: Number of votes by data source

Before examining this figure in detail, it is important to note that differences between the two panels likely *understate* differences between the two datasets. This level of aggregation (simple vote counts by session) may hide cases where there are both too many and too few votes. For example, a dataset might include duplicate entries for one vote and no entry for another vote. See Appendix for a non-exhaustive list of such examples. Nonetheless, this comparison is useful because it provides an overview free from too many details.

We start by noting similarities between the two datasets. Both show a bimodal distribution of votes before the 19<sup>th</sup> session, and both show the effects of a Soviet strategy to block UN action in 1964 (the 19<sup>th</sup> session). Both show a subsequent rise in reported voting activity, peaking around the 39<sup>th</sup> session. Finally, both show a drop in recorded amendment votes in the second half of the period covered.

Turning to differences, recorded amendment votes are completely absent in the Voeten data between sessions 39 and 51 whereas we find 40 recorded amendment votes in this period. More generally, for the first 29 sessions Voeten generally records more amendment votes than we do (an average of 9.7 per session as compared with 6.4 per session, with the greatest difference in session 3 with 41 amendments as compared to 26). This pattern reverses with Voeten recording a total of 20 recorded amendment votes from session 30 to session 70 as compared to 95 over the same period in our data.

For paragraph votes, figures are generally similar up through session 32 then diverge until session 52. During the intervening two decades, the Voeten data report 30 recorded paragraph votes while our data identify ten times that number (311). We again see substantially greater numbers in our data for the last five sessions.

Finally turning to recorded resolution votes, the two series track closely especially when considered in percentage terms. Differences between the two vote counts (totaling 330 votes) account for about 6% of the total number of votes and are not concentrated in a particular time period. Our data report 5430 recorded resolution votes while Voeten finds 5154.<sup>16</sup>

---

<sup>16</sup> Again, for this exercise, we mirror Voeten by combining motion and resolution votes so the number of “resolutions” indicated here is higher than elsewhere.

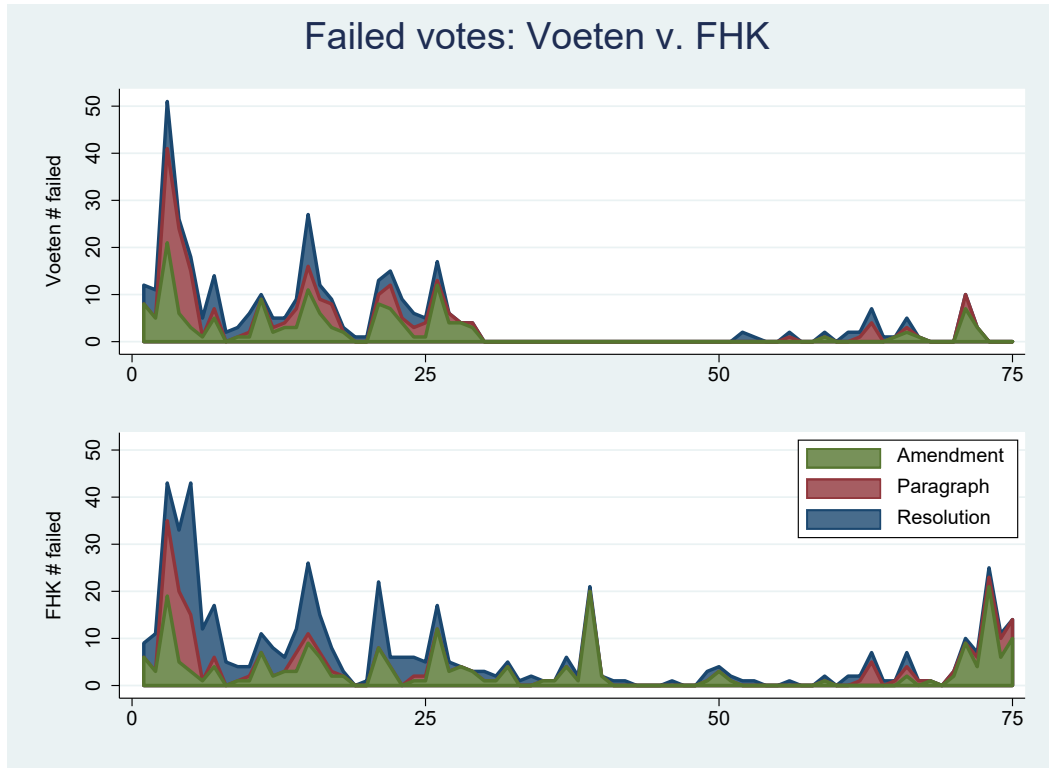


Figure 2: Number of failed votes by data source

The numbers above reflect all recorded votes from UNGA plenary sessions. Next, we focus just on votes where the measure failed to be approved as these are more likely to be missed given the UN reporting system. For consistency across both datasets, we define a failed vote as cases where the number of “no” votes exceeds the number of “yes” votes. Looking first at amendments that failed to pass, the differences are striking. From the 30<sup>th</sup> to the 58<sup>th</sup> session, Voeten reports no failed amendments whereas we find 40 cases. Again at the end of the sample, our data include a substantial number of failed amendments not captured in Voeten’s dataset. Overall, our data collection method identifies 60 more failed amendments (149 versus 209), 40% more.

The pattern is somewhat reversed failed paragraph votes; Voeten identifies 99 while we find 79. In general, the Voeten count is higher up through session 30 while our count is generally higher

after session 62, particular in the last four years when we record 10 failed paragraph votes while Voeten records none.

Finally, we look at resolutions that fail to pass. In total Voeten reports 100 final votes on resolutions that failed to pass; our data include 212. Our data report substantially more failed resolutions before the 25<sup>th</sup> session. We also find a small but steady number between the 30<sup>th</sup> and 50<sup>th</sup> sessions (as well as in recent sessions) where Voeten reports none (all failed motions). Perhaps the most striking aspect of this is that gaps evident in the Voeten data are not present in our data.

To do:

- Comparison of vote passage margins by session (Voeten v. FHK)
- Smaller examples of differences
- Comparison with UN Digital Library
  - o Repeat figure 1 but without breakdown by vote type
  - o Repeat comparison of vote passage margins by session

## **V. Replication studies**

TBD

## **VI. Conclusions and future research**

Existing data sets appear to miss important aspects of UN voting. There has been substantial change over time in UN voting patterns. Existing datasets both understate this (e.g., by omitted



failed votes) or overstate this (because their coverage of different vote types is inconsistent over time). These problems run two risks: introducing noise that undermines otherwise statistically significant relationships and introducing apparent changes where none exist.

## References

- Alesina, Alberto, and David Dollar. "Who gives aid to whom and why?" *Journal of Economic Growth* 5.1 (2000): 33-63.
- Andersen, Thomas Barnebeck, Thomas Harr, and Finn Tarp. "On US politics and IMF lending." *European Economic Review* 50.7 (2006): 1843-1862.
- Bailey, Sydney Dawson. 1960. *The General Assembly of the United Nations: A study of procedure and practice*. New York, NY: Praeger.
- Bailey, Michael A., and Erik Voeten. "A two-dimensional analysis of seventy years of United Nations voting." *Public Choice* 176.1-2 (2018): 33-55.
- Bailey, Michael A., Anton Strezhnev, and Erik Voeten. "Estimating dynamic state preferences from United Nations voting data." *Journal of Conflict Resolution* 61.2 (2017): 430-456.
- Ball, Richard, and Christopher Johnson. "Political, economic, and humanitarian motivations for PL 480 food aid: Evidence from Africa." *Economic Development and Cultural Change* 44.3 (1996): 515-537.
- Barro, Robert J., and Jong-Wha Lee. "IMF programs: Who is chosen and what are the effects?" *Journal of Monetary Economics* 52.7 (2005): 1245-1269.
- Dreher, Axel, and Nathan M. Jensen. "Independent actor or agent? An empirical analysis of the impact of US interests on International Monetary Fund conditions." *The Journal of Law and Economics* 50.1 (2007): 105-124.
- Dreher, Axel, and Nathan M. Jensen. "Country or leader? Political change and UN General Assembly voting." *European Journal of Political Economy* 29 (2013): 183-196.
- Dreher, Axel, and Jan-Egbert Sturm. "Do the IMF and the World Bank influence voting in the UN General Assembly?" *Public Choice* 151.1 (2012): 363-397.
- Gartzke, Erik, and Dong-Joon Jo. *The affinity of nations index, 1946-2002*. New York, NY: Columbia University (2006).
- Häge, Frank, and Simon Hug. "Consensus decisions and similarity measures in international organizations." *International Interactions* 42.3 (2016): 503-529.
- Hovet, Jr Thomas. *Bloc politics in the United Nations*. (1960) Cambridge, MA: Harvard University Press.
- Kim, Soo Yeon, and Bruce Russett. "The new politics of voting alignments in the United Nations General Assembly." *International Organization* 50.4 (1996): 629-652.
- Peterson, M.J. *The United Nations General Assembly*. (2005) Abingdon, Oxon: Routledge.
- Thacker, Strom. "The High Politics of IMF Lending." *World Politics* 52.1 (1999): 38-75.

Voeten, Erik. "Resisting the lonely superpower: Responses of states in the United Nations to U.S. dominance." *Journal of Politics* 66.3 (2004): 729-754.

Voeten, Erik. "Data and analyses of voting in the United Nations General Assembly." *Routledge Handbook of International Organization*. Routledge, 2013. 80-92.

Voeten, Erik. *Ideology and International Institutions*. (2021) Princeton, NJ: Princeton University Press.

Wang, T. Y. "U.S. foreign aid and UN voting: An analysis of important issues." *International Studies Quarterly* 43.1 (1999): 199-210.

## Appendix

### Detailed description of methods for scraping data from minutes:

We start by extracting the text of the minutes from the PDF line by line (eliminating the standard two-column format used during this period). We then use the structure and formatting of the minutes (e.g., indentation, italics, bolding) to group these lines of text into three types of items: headings, speech paragraphs, and procedural notes. We code the exact location of each item in the minutes (page, column, and line). We use parsing rules that are broad enough that we do not miss-code items with small formatting irregularities (false negatives) but that are also narrow enough to avoid incorrectly including irrelevant text (false positives). We validate this algorithm by hand to ensure it is performing accurately.

The procedural items include, among other things, decisions taken without a vote and decisions taken with a recorded vote. For resolution-related recorded votes, the text of the minutes includes the subject of the vote (the draft resolution number and, for separate votes, the paragraph number), the outcome of the vote (the total number of yeses, noes, and abstentions), and the lists that indicate how each country voted. We also associate each decision with the heading (and agenda item) that it falls under. We extract and clean all of this metadata and make it available for users to include with our country-level vote data for decisions in sessions 49-75 if they are interested in it.

For decisions taken without a vote, we include the text of the question that is put by the presiding officer so users can identify what the decision is about. For recorded votes, we include the procedural text that indicates the outcome of the vote, including the vote totals and whether the vote was adopted (which might not be clear from the totals when a two-thirds decision-rule is being used). We use the lists of countries in the minutes to create country-level vote data. For each country, we use the official UN country name on the day of the vote, which is usually (but not always) the name used in the minutes (as the minutes are inconsistent for some countries). We map these country names to Correlates of War (COW) country codes. Again, for final passage votes, we check our country-level vote data against the data reported by the UN Digital Library.

### Detailed description of data validation steps:

First, for recorded votes and non-recorded votes, we ensure that the number of yeses, noes, and abstentions is equal to the number of voting members less the number of countries recorded as not voting. When there is a discrepancy, we check each source by hand and determine the correct information. As part of this process, we code the official start date (to the exact day) of each member and generate a list of voting members on the day of each vote. For final passage votes, which are available from the UN Digital Library, we ensure that the number of voting members reported in the UN Digital Library metadata matches the number that we calculate (as there are errors in the UN Digital Library metadata).

Second, for recorded votes, we ensure that the vote totals that we have collected from the UN Digital Library match the country-level data (e.g., the number of countries listed as voting yes add up to the total number of yes votes, etc., which is not always the case because of errors in the

UN data). For sessions 49-75, where we also extract recorded votes from the minutes, we make this same comparison with the totals reported in the minutes. When there is a discrepancy, we check the minutes by hand to determine the right information and then correct the data. In a very small number of cases, there is an obvious error in the minutes, in which case we use all available information to determine what the correct coding should be. We produce detailed documentation that tracks all discrepancies we have identified and resolved.

### List of known errors

In the process of deploying our data collection strategy we were already able to identify in detail some specific issues with existing data. Our ultimate goal is to provide a full list of these issues based on a systematic comparison of datasets, while here we only provide those of which we currently are aware. The ICPSR dataset, and thus by implication also the Voeten *et al.* dataset, misses five recorded final passage votes on draft resolutions:

- A/RES/62(I)[I] Constitution of the International Refugee Organization, and Agreement on Interim Measures to be taken in respect of refugees and displaced persons: [...] (adopted in a recorded vote on 1946-12-15)
- A/RES/62(I)[II] Arrangements and measures to be taken by Members of the United Nations in connection with displaced persons, refugees, prisoners of war and persons of [...] (adopted in a recorded vote on 1946-12-15)
- A/RES/2443(XXIII) Respect for and implementation of human rights in occupied territories: resolution (adopted in a recorded vote on 1968-12-19)
- A/RES/2996(XXVII) International responsibility of States in regard to the environment: resolution (adopted in a recorded vote on 1972-12-15)
- A/RES/3151(XXVIII)[G] Situation in South Africa resulting from the policies of apartheid: resolution (adopted in a recorded vote on 1973-12-14)

In recent sessions, there is apparent coding error where final vote is missed but paragraph votes are captured.