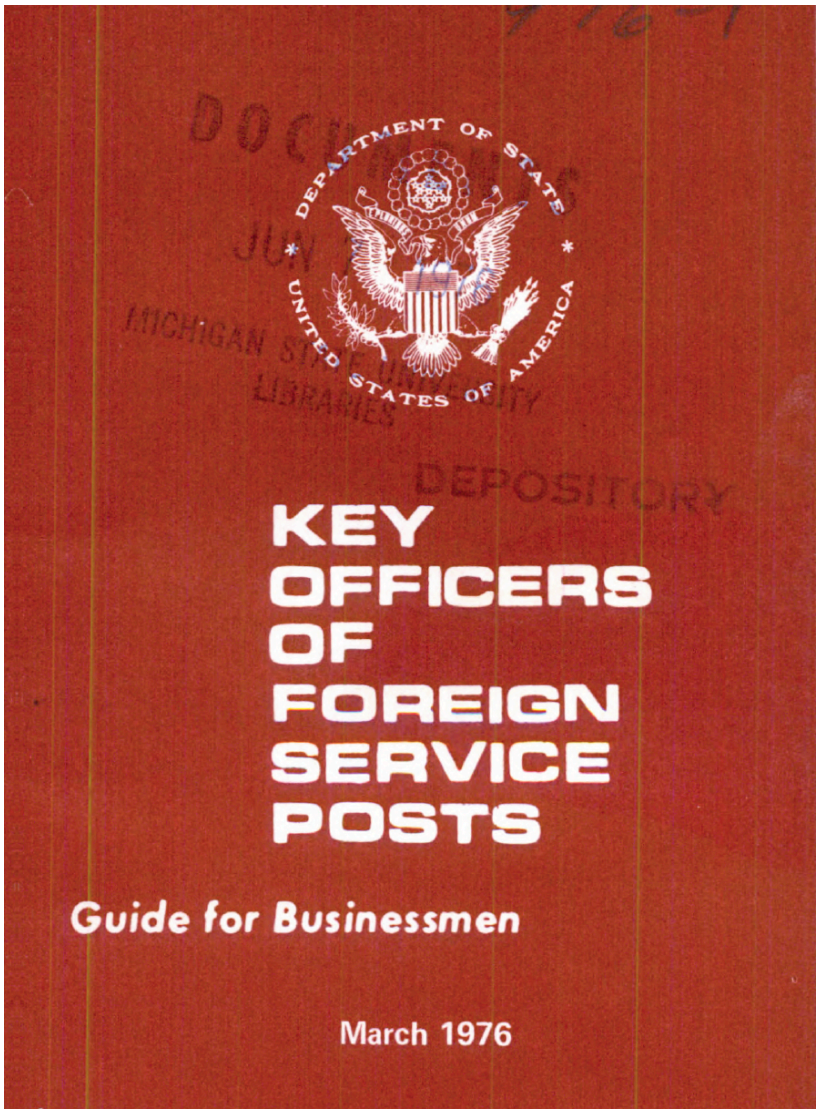
Key Officers of U.S. Foreign Service Posts (KOFSP): Codebook

David Lindsey* Matt Malis† Calvin Thrall‡

Version 1.1 (March 2026)



Key Officers Directory, March 1976.

*Associate Professor, Baruch College; https://blogs.baruch.cuny.edu/davidlindsey/
†Assistant Professor, Texas A&M University; https://mattmalis.github.io/
‡Assistant Professor, Columbia University; https://www.calvinthrall.com/

Suggested citation

We make this data freely available for noncommercial use. If you use the KOFSP data, we only ask that you please cite the academic paper in which it was introduced. Here we provide suggested citations for this paper in two different formats:

APSA

Lindsey, David, Matt Malis, and Calvin Thrall. 2026. "Key Officers of U.S. Foreign Service Posts: A Dataset of U.S. Diplomatic Representation, 1966-2017." Working Paper. URL: https://www.calvinthrall.com/assets/kofsp.pdf.

.bib

@techreport{lindsey_etal_2026,
address = {New York},
author = {Lindsey, David and Malis, Matt and Thrall, Calvin},
type = {Working Paper},
title = {Key Officers of U.S. Foreign Service Posts: A Dataset of U.S. Diplomatic Representation, 1966-2017},
year = {2026},
url = {https://www.calvinthrall.com/assets/kofsp.pdf}
}

Contents

1 About the data **5**

1.1 Data formats and reference files . . . 5
1.2 Sources . . . 6
1.3 Additional information . . . 6

2 Variable descriptions: Individual-level data **7**

2.1 Geographic and Post-related . . . 7
2.1.1 `country` . . . 7
2.1.2 `ccode` . . . 7
2.1.3 `in_COW_system` . . . 7
2.1.4 `city` . . . 8
2.1.5 `post` . . . 8
2.1.6 `post_type` . . . 8
2.1.7 `post_type_raw` . . . 10
2.2 Time . . . 10
2.2.1 `year` . . . 10
2.2.2 `quarter` . . . 10
2.2.3 `qy` . . . 11
2.3 Individual and position-related . . . 11
2.3.1 `officer_ID` . . . 11
2.3.2 `officer_name` . . . 11
2.3.3 `officer_name_raw` . . . 12
2.3.4 `suffix` . . . 12
2.3.5 `military` . . . 12
2.3.6 `residence_location` . . . 12
2.3.7 `officer_is_at_post` . . . 12
2.3.8 `residence_certainty` . . . 13
2.3.9 `position_category` . . . 14
2.3.10 `position_agg` . . . 15
2.3.11 `officer_position` . . . 16
2.3.12 `senior_officer` . . . 16
2.3.13 `political_appointee` . . . 17
2.3.14 `spell_ID` . . . 17

3 Variable descriptions: Country-level data **19**

3.1 Diplomatic Representation Aggregates . . . 19
3.1.1 `n_posts` . . . 19
3.1.2 `post_names` . . . 19
3.1.3 `has_emb` . . . 19
3.1.4 `has_amb` . . . 20
3.1.5 `n_officers` . . . 20
3.1.6 `n_generalists` . . . 20
3.1.7 `n_officers_emb` . . . 20
3.1.8 `n_generalists_emb` . . . 21
3.1.9 `pol_amb` . . . 21
3.1.10 `attache_` variables . . . 21

3.1.11 hardship 21
3.1.12 danger 22
3.1.13 normal_relations 22

4 Variable descriptions: Post-level data **22**
4.0.1 senior_officer_position 23
4.0.2 senior_officer_name 23

1 About the data

The Key Officers of U.S. Foreign Service Posts (KOFSP) dataset contains information on the following, observed quarterly between 1966 and 2017:

1. the names, locations, and classifications of each U.S. overseas diplomatic post (embassies, consulates, branch offices, missions to international organizations, commercial offices, and other);[1]
2. the diplomatic position categories represented at each post;
3. the identities of the top-ranking ("key") officers at each post in each position category, as well as unique identifiers for each officer.

It is therefore the most comprehensive micro-level dataset on diplomatic representation to ever be compiled. There are many ways in which this data might be used to study the determinants of diplomatic representation, the effects of resident diplomacy, the factors that shape individual diplomatic career trajectories, and a variety of other research questions.

1.1 Data formats and reference files

We share four versions of the dataset:

- `KOFSP_individual_v1.1.csv`: This is the disaggregated version of the data, structured at the officer-position-quarter-year level, which most closely reflects the original Key Officers publications.
- `KOFSP_post_quarter_year_v1.1.csv`: This version aggregates the individual-level data up to the post-year level.
- `KOFSP_country_quarter_v1.1.csv`: This version aggregates the individual-level data up to the country-quarter-year level.
- `KOFSP_country_year_v1.1.csv`: This version aggregates the country-quarter-year-level data up to the country-year level.

[1] We use the term "post" throughout this codebook generically to refer to any establishment listed within the Key Officers directories, whether or not these are "posts" as that term is used by the State Department.

Additionally, the KOFSP Dataverse site contains two additional (sets of) reference files:

- `KOFSP_position_mappings_v1.1.csv`: Most values of the `officer_position` variable are reported using State's system of abbreviations and acronyms; this file contains a correspondence between each common value of `officer_position`, the position title to which it refers, as well as the value of `position_agg` assigned to each position. While the correspondence between `officer_position` and `position_agg` can easily be constructed from the main data files, users less familiar with the foreign service may find the full titles helpful.
- `key_officers_directories_1966_2017.zip`: This file contains all 156 Key Officers directories from which we collected the KOFSP data.

1.2 Sources

The KOFSP data were collected from directories, regularly issued by the U.S. State Department, that contain contact information for all U.S. diplomatic posts; the names of all position categories represented at a given post; and the names of the highest-ranking officer in each position category at each post. These directories, titled "Key Officers of Foreign Service Posts," were primarily created for distribution to representatives of U.S. businesses operating abroad; until the 2000s, the directories were subtitled "Guide for Business Representatives."[2] Prior to 1998, the directories were only issued in print format; for these years, we rely on digitized versions accessed via HathiTrust. In the 21st century, we use both HTML and native PDF versions of the directories, accessed via the Internet Archive's Wayback Machine (`archive.org`). We collect one directory per quarter-year whenever available, resulting in 156 total directories between 1966 and 2017.

1.3 Additional information

Those seeking additional information on the construction and contents of the KOFSP data, as well as steps the authors have taken to validate its contents, are directed to read the working paper in which the data are introduced.

[2] Prior to 1977: "Guide for Businessmen."

2 Variable descriptions: Individual-level data

2.1 Geographic and Post-related

2.1.1 `country`

The country in which the post is located, standardized using R package countrycode (Arel-Bundock, Enevoldsen, and Yetman, 2018). For continuity and ease of merging in other variables, we use modern rather than historical country names when dealing with successor states. The example of East vs. West Germany is instructive. Under international law, the modern state of Germany is considered to be the legal continuation of the state of West Germany (Federal Republic of Germany) that existed prior to reunification in 1990. Thus, we code the `country` of the U.S. post in West Berlin prior to 1990 as Germany, even though the historical country name was the Federal Republic of Germany. However, East Germany (German Democratic Republic) has no legal continuator state; as a result, we code the `country` of the U.S. Embassy in East Berlin as German Democratic Republic.

2.1.2 `ccode`

The Correlates of War (COW) country code corresponding to the country in which the post is located. There are several entities which are listed in the Key Officers publications as countries, but are not included in the COW state system, Bermuda, French Caribbean Department, Hong Kong, Jerusalem, Netherlands Antilles, and Ryukyu Islands. We keep these country names in the data, and assign an NA value of `ccode`.

2.1.3 `in_COW_system`

An indicator variable denoting whether the country-year in question is part of the COW state system. This takes on a value of 0 in the cases of `ccode` = NA described above, and in cases where a country appears in the Key Officers data prior to its appearance in the COW state system. For example, Key Officers lists a CG in Luanda, Angola from 1966 until Angola's independence in 1975; these pre-independence observations take a value of `in_COW_system` = 0, with `ccode` = 540 (Angola's country code once it appears in the COW state system).

2.1.4 `city`

The city in which the post is located. Similarly to the way we code the `country` variable, when a city has changed names without changing its territorial boundaries we use the most recent name available for all years. For example, the U.S. established a consulate general in Leningrad, USSR in 1973 and maintained it after the USSR collapsed and the city was renamed St. Petersburg. Because Leningrad and St. Petersburg refer to the same geographic area, we set the `city` variable for this consulate general to St. Petersburg (the city's most recent name) all the way back to 1973. Conversely, while East and West Berlin are both contained within post-reunification Berlin, none of the three municipalities have the same territorial boundaries; thus, all three receive a separate value of the `city` variable.

2.1.5 `post`

The name of the post itself. For all embassies, consulates general, and consulates, this will be identical to the `city` variable. The `post` variable may differ from the `city` variable in the following cases: (i) missions to international organizations; (ii) commercial/trade promotion offices; (iii) other posts with a specific purpose, such as USIS offices; (iv) "American Interests Sections" within non-U.S. missions to countries with which the U.S. does not have formal diplomatic relations (e.g. "AMERICAN INTERESTS SECTION, SWISS EMBASSY – HAVANA").

2.1.6 `post_type`

The classification of the diplomatic post. The potential values are as follows:

1. E – **Embassy**: the most important bilateral diplomatic post in a given country. Embassies are headed by ambassadors, with Deputy Chiefs of Mission (DCMs) as second in command; they tend to have full slates of foreign service generalists and specialists, as well as a variety of attachés from other federal agencies. With rare exceptions, embassies hold supervisory authority over consular posts within the same country.[3]

[3]Collectively, the embassy to a country and subordinate posts constitute the "mission" to that country, with the ambassador serving as chief of mission with overall supervisory authority. In certain unusual cases, the principal officer of a Consulate General may hold chief of mission authority, reporting directly to the State Department rather than the ambassador (see 2 FAH-2 H-112). At present, this applies only to Curacao and Hong Kong.

2. CG – **Consulate General**: consulates general, headed by consuls general,[4] resemble "mini-embassies" with diverse functions. In addition to providing consular services, they often have several foreign service specialists and generalists outside of the consular cone, and sometimes they host attachés as well. Large CGs can be larger than small embassies.

3. C – **Consulate**: consulates are small posts that primarily (and sometimes exclusively) provide consular services. While consulates sometimes host generalist foreign service officers and occasionally officers from the foreign commercial service, they rarely host any other type of attaché.

4. CA – **Consular Agency**: Consular agencies provide limited consular services under the supervision of another post, and are staffed by consular agents who are not foreign service officers and often work part-time.[5] In most analyses, it will make sense to drop consular agencies from the sample.

5. USINT – **U.S./American Interests Section**: when the U.S. has no formal diplomatic relations with a partner state, they may negotiate to form an American Interests Section in a non-U.S. diplomatic mission to that state. These sections, which the U.S. has formed in Algeria, Sudan, Egypt, Iraq, and Cuba, typically consist of foreign service generalists and assorted attachés.

6. Commercial Office – **Commercial Office**: offices dedicated primarily to trade and investment promotion. Commercial offices are staffed by officers in both the State Department's foreign service and the Commerce Department's foreign service. These posts are treated inconsistently in the Key Officers directories over time: they are sometimes listed as separate posts, and they are sometimes combined with other posts in the same city. We therefore advise caution when conducting post-level analyses using commercial offices.

7. IO – **Mission to International Organization**: U.S. missions to international organizations

[4]The title of "consul general" is (confusingly) used both by the principal officer of a consulate general and by certain other officers performing purely consular functions (such as the head of the consular section at a large embassy). These are distinct roles and should not be conflated.

[5]Under State Department regulations: "A consular agent may be appointed by the Secretary [of State] to provide limited consular services at a location within a consular district. Consular agents are limited, non-career appointees of the Foreign Service already residing abroad, who act under the supervision and direction of the supervisory consular officer. They are generally employed on a part-time intermittent basis."

or other multilateral summits or negotiations. In addition to the standard foreign service generalists, specialists, and external attachés, these missions occasionally have specialized positions not found elsewhere. Missions to international organizations are independent of the diplomatic mission to the country where they are located, but they may share some administrative and other resources with the mission.

8. Other – **Other Post**: this catchall category contains a range of other posts. Examples include branch offices of embassies, offices of the now-defunct U.S. Information Service (USIS), and offices of the U.S. Immigration and Naturalization Service (INS).

2.1.7 `post_type_raw`

The classification of the diplomatic post as it appears in the Key Officers directory. In most cases, this is equivalent to `post_type`. However, the two will differ in cases where we either (i) fixed errors in the directories or (ii) aggregated smaller post types into broader categories. We include this variable in the data because users may wish to use the more disaggregated post types. However, we note that this variable is *raw*—it includes errors and inconsistencies, and we urge interested parties to tread carefully when using it.

2.2 Time

2.2.1 `year`

The calendar year in which the directory was issued.

2.2.2 `quarter`

The quarter of the year in which the directory was issued.

1 = Jan – Mar;

2 = Apr – June;

3 = July – Sep;

4 = Oct - Dec.

2.2.3 qy

Short for quarter-year, this is a numeric variable that combines calendar year with quarter. For example, Q1 2005 would have a qy value of 2005; Q2 2005 would have a qy value of 2005.25; Q3 2005 would have a qy value of 2005.5; Q4 2005 would have a qy value of 2005.75.

2.3 Individual and position-related

2.3.1 officer_ID

A unique numeric identifier for each individual diplomat in the data. This is the only variable in the KOFSP data that uniquely identifies diplomats; while there are many instances in which different diplomats share the exact same name,[6] for example, identifiers are specific to the individual.

To create this variable, we used both automated and manual record-linking techniques to identify unique officers based primarily on (i) their names; (ii) their positions; (iii) the dates at which they appear in the Key Officers directories. We went to great lengths to link officers even in the presence of nicknames/shortenings, last name changes due to marriage, multiple officers with the same exact name, and a variety of other edge cases (as discussed in depth in the data introduction paper). We stress, however, that this variable is our own creation rather than an official administrative identifier.

2.3.2 officer_name

The cleaned, modal name associated with each individual officer. We begin by recording the officer name as it appears in the Key Officers directory (this is what is listed in the officer_name_raw variable); we then remove prefixes, suffixes, titles, middle initials, and nicknames (e.g. Dwight "Ike" Eisenhower), and perform some additional cleaning. For consistency, the officer_name variable is set to the modal name for each unique officer. For example, suppose that the same individual is listed as "Tom Warner" 6 times and "Thomas Warner" 18 times throughout his career; in this case, the officer_name variable would list him as "Thomas Warner," as this is the most common variant of his name.

[6] There are 941 names (first and last, e.g. "David Brown") shared by at least 2 diplomats; in total, 2,172 officers have the same name as at least one other officer in the data.

2.3.3 officer_name_raw

The officer's name as it appears in the respective Key Officers directory. Note that, due to inconsistencies in how individual officers were listed over time, this variable often takes multiple different values within a single officer_ID.

2.3.4 suffix

The suffix listed with an officer's name, if any. This variable is coded directly from officer_name_raw.

2.3.5 military

A binary variable equal to 1 if an officer appears to be military based on the title included with their name in the corresponding Key Officers directory (e.g. "Lt", "Col", "Capt", etc), and zero otherwise. This variable is coded directly from officer_name_raw. As it is possible that some military officers are listed in the Key Officers directories without their military titles, and we do not use external sources to validate military status, this variable should be used with caution.

2.3.6 residence_location

If an officer is *not* in residence at the post listed in a given officer-position-quarter-year observation, the resident_location variable reports where the officer is located instead; if it is not clear where the officer is actually in residence, which occurs when an officer is listed simultaneously at two different posts with no indication of residence location, the variable will take a value of "unclear". This variable is coded NA for observations in which the officer is resident at the location given by the city variable.

Note that, at present, this variable is not cleaned extensively or standardized to allow for easy linking with the post variable. This is often a non-issue in practice, as officers are typically listed at both their post of residence *and* any non-resident posts in the same quarter-year.

2.3.7 officer_is_at_post

A binary variable equal to 1 if the officer appears to be in residence at this diplomatic post in this quarter-year, and zero otherwise. There are two main reasons why the value of this variable

might be zero for a given officer-position-quarter-year. First, diplomats occasionally hold a position at a given post while living elsewhere; they may live in a different region of the host state for official reasons, for example, or they may be moonlighting at a small post that is nearby their primary post.[7] This type of "multiple accreditation" even occurs with ambassadors; for example, the U.S. Ambassador to Cameroon is often accredited to Equatorial Guinea as well.

Second, it occasionally happens that (i) a diplomat's prior post continues listing them for more than one quarter after their departure, and/or (ii) a diplomats' future post lists them for more than one quarter prior to their arrival. When this happens, there is a period of overlap wherein the same diplomat is reported as being simultaneously at two different posts. Because we do not have access to a separate source of ground truth information about where diplomats are physically located at any given time, when this occurs we cannot definitively say that an office is at either post; we thus err on the side of caution and set the `officer_is_at_post` variable to 0 for overlapping officer-quarter years.[8]

2.3.8 `residence_certainty`

The level of specificity with which we can report the residence location for a given officer. Note that this varies at the officer-quarter-year level. This variable takes four potential values:

- **post_known**: We know both the country and the post where this officer is located. For example: Aaron Wilkins (Officer ID: 23920) is listed as FAA attaché in both Abu Dhabi and Manama in 2017q1; his listing in Manama says "Abu Dhabi Based". For both observations, we set `residence_certainty` to "post_known"; and we set `officer_is_at_post` to 1 for the Abu Dhabi listing, and 0 for the Manama listing.

- **country_known**: We don't know the specific post at which this officer is in residence, but we do know the country. This occurs when an officer is simultaneously listed at two posts in the same country (for more than 1 quarter), with no indication of their residence location.

[7] In some rare cases, the same external attaché will be listed at a large number of posts in a given quarter-year. For instance, Kathy Beck (officer ID: 18658) is simultaneously listed as the IRS attaché at 48 separate posts in 2007 and 2008.

[8] The exception to this is ambassadors, for whom we do have ground truth records of arrival and departure dates through the State Department Office of the Historian.

- **in_transition**: We list an officer as "in transition" if they are simultaneously listed at their previous and current posts for only 1 quarter. In these cases, it is possible that an officer transferred between postings mid-quarter and thus actually was in both places; thus, for officers in transition, we set `officer_is_at_post` to 1 for both posts.

- **unclear**: We do not have enough information to determine either the post or the country in which this officer was in residence. This occurs when an officer is simultaneously listed at two posts in *different* countries (for more than 1 quarter) with no indication of their residence location. In all cases with `residence_certainty` = "unclear", we also code `residence_location` = "unclear" and `officer_is_at_post` = 0.

2.3.9 `position_category`

This variable describes the type of officer that is most likely to fill a given position. It is coded based on the position, not the officer filling the position (i.e., a given officer may take on different values of this variable over time). It takes six potential values:

- **Generalist**: Generalists are foreign service officers within the State Department that handle most of the substantive work of diplomacy; they are hired based on potential rather than on their pre-existing skill set, and are thus qualified to fill any role within their career track (or "cone"). Generalists may also sometimes take on assignments outside their cone. The five generalist cones are: Political, Economic, Public, Management, and Consular.[9]

- **Specialist**: Foreign service specialists are employed by the State Department to provide a variety of administrative and support services for overseas posts; unlike generalists, they are hired to do a specific job. Examples of specialist positions include diplomatic security, information technology, and operations and maintenance.

- **Nonstate FSO**: This category covers foreign service officers who are employed by agencies

[9]Note that this variable is coded to reflect organizational changes in the State Department workforce. For instance, the Foreign Commercial Service moved from the State Department to the Department of Commerce in 1980, so their `position_category` variable changes from "Generalist" to "Nonstate FSO". Likewise, Public Diplomacy officers are coded as "Nonstate FSO" prior to the incorporation of the U.S. Information Agency (USIA) into the State Department in 1999, and "Generalist" thereafter.

other than the State Department.[10] This group covers officers from USAID, the US Commercial Service, the Foreign Agricultural Service, the Animal and Plant Health Inspection Service (APHIS), and the US Agency for Global Media.

- **External**: This category covers officers who are not employed by the State Department *and* are not foreign service officers. This is a broad category that includes, but is not limited to: attachés from military agencies, Treasury and IRS, FAA, Peace Corps, law enforcement agencies like the FBI and DHS, EPA, and others.

- **Undetermined FSO**: This category covers positions that can be held by foreign service officers across agencies. It includes CLOs (community liason officers), AFSA representatives (representatives of the American Foreign Service Association, the labor union for foreign service officers), and other similar positions. They are typically auxiliary positions for officers serving in a different primary capacity at a post, which may or may not itself be listed among the key officers.

- **Unclear**: These are positions for which we are unsure about the correct categorization. This typically occurs for one of two reasons. First, transcription errors, omissions, and poor scan quality in the Key Officers directories occasionally render a position abbreviation either incorrect or illegible. Second, for a number of historical and/or unique position titles, we have not been able to find information on the type of officer that is qualified to fill them.

2.3.10 `position_agg`

This variable represents our attempt to aggregate the 1,500+ unique position titles found in the Key Officers directories into 29 coherent thematic categories. A correspondence between the most common values of `officer_position` and their corresponding `position_agg` can be downloaded from the MAD website.

[10] The Foreign Service Act of 1980 determines which non-State agencies can use the foreign service personnel system; other agency personnel working abroad are not considered foreign service officers, and are not eligible for appointment to career Ambassadorships.

2.3.11 officer_position

This variable reports the officer's position as listed in the Key Officers directories. The primary position titles are relatively stable over time; for example, there are 64 unique position titles with >1,000 officer-position-quarter-year observations, and which collectively account for 93% of the KOFSP data. However, there are also over 1,000 position titles that appear fewer than 10 times in the data. Most values of `officer_position` are abbreviations or acronyms; for brief descriptions of over 150 of the most common positions, see the `officer_position:position_agg` correspondence table.

2.3.12 senior_officer

A binary variable equal to 1 if this officer holds a position that makes them the senior officer of their post, and zero otherwise. We code this variable differently for embassies and consulates general:

- **Embassies**: The senior officer at an embassy is the ambassador (AMB). If the ambassadorship is vacant, the deputy chief of mission (DCM) typically becomes the acting chief of mission ("chargé d'affaires"). In rare cases, a different officer is assigned the role of chargé d'affaires (CHG), in which case we list this officer as the senior officer. NB: while consuls general may be listed at embassies, this is *not* the senior officer position at an embassy.
- **Consulates General**: The senior officer at a consulate general is the consul general (CG). If a consulate general does not have a CG, another officer may be listed as the principal officer (PO) in which case we list this individual as the senior officer. In rare cases, a consulate general will lack both a CG and a PO but have a deputy principal officer (DPO), and in such cases we list the DPO as the senior officer.

This variable is coded `NA` for all officers at all post types other than embassies and consulates general.

2.3.13 `political_appointee`

A binary variable equal to 1 if the officer is a political appointee, and zero otherwise. While the vast majority of officers in the KOFSP data are career diplomats, ∼29% of ambassadors and a limited number of other positions are filled by political appointees. We identify political appointees by merging the KOFSP data with the State Department Office of the Historian's "Principal Officers and Chiefs of Mission" database.[11] The Principal Officers database includes the most senior officers in the department (generally, ambassadors and officers at the rank of Assistant Secretary and above), who are appointed by the President with the advice and consent of the Senate. This database identifies each officer as either a "Career Foreign Service Officer" or a "Non-career appointee".[12] We linked officers from Principal Officers to their corresponding entries in Key Officers, and coded `political_appointee` = 0 for those listed in Principal Officers as "Career Foreign Service Officer", 1 for those listed as "Non-career appointee", and NA for everyone else. Political appointees are almost exclusively ambassadors, high-level officers at IO missions, or principal officers at other high-profile posts (such as the U.S. "Liaison Office" in Beijing in the mid-1970s); in only a handful of cases do political appointees appear in other embassy positions such as DCM or as USAID mission director.

2.3.14 `spell_ID`

This variable uniquely identifies each stretch of time that one diplomat spent at one post, which we refer to as a "spell." For example, Amb. Herbert Horowitz first appears in the KOFSP data as Economic Officer in Beijing from 1973-1975; during this period, Horowitz receives a `spell_ID` value of 3894_BEIJING_1 ({officer_ID}_{post name}_{nth tour at this post}). Horowitz would return to Beijing in 1985 as DCM; during this spell, `spell_ID` would be equal to 3894_BEIJING_2.

We caution that this variable is *not* equivalent to a unique identifier for each position that a given diplomat has held during her career. The primary reason for this is that a number of officers simultaneously hold multiple positions at the same post; in these cases, spell IDs will not uniquely identify diplomat-positions. For example, Edward Heartney is simultaneously listed as Economic

[11] https://history.state.gov/departmenthistory/people/principals-chiefs.

[12] We have fixed a number of errors in the State Department historian's coding of this variable, so users may notice that it is not a perfect match to what is listed on the State Department website.

section chief, Commercial section chief, and State ICASS chair during his tour in Colombo from 2009–2012; all such listings are grouped together as `spell_ID` = 21125_COLOMBO_1. There are also instances in which an officer holds two different subsequent positions at the same post during the same spell, such as a DCM who becomes chargé during an ambassadorial vacancy. Those interested in creating a unique ID for the diplomat-position could do so by concatenating `spell_ID` and `officer_position`.

3 Variable descriptions: Country-level data

The country-quarter-year dataset (`KOFSP_country_quarter_v1.1.csv`) takes the individual-level dataset and aggregates up to the country-quarter-year level. It includes the same `country`, `ccode`, `year`, `quarter`, and `qy` variables as the individual-level dataset, along with the additional variables listed below. All aggregate measures are based on the set of officers within a country-quarter-year who are (i) resident in the location they are listed (i.e. coded as `officer_is_at_post` =1), and (ii) listed at any post type other than an IO mission. Note that the countries included in this dataset in a given quarter-year are only the countries with at least one U.S. key officer listed as resident in the country; for example, the Central African Republic does not appear in our country-level data in 2004, despite the U.S. maintaining diplomatic recognition of the country, because the U.S. temporarily suspended embassy operations and withdrew diplomatic personnel following a coup in late 2002 (and resumed operations in 2005).[13]

The country-year dataset includes all of the same variables from the country-quarter-year dataset (other than `quarter` and `qy`), with values averaged across the quarters of a given calendar year.

3.1 Diplomatic Representation Aggregates

3.1.1 `n_posts`

The number of distinct diplomatic posts in the country (excluding IO missions).

3.1.2 `post_names`

A list of the posts in the country, separated by a semicolon. For example: "ABU DHABI (E); DUBAI (CG)".

3.1.3 `has_emb`

A binary variable denoting whether the Key Officers publication lists an embassy with officers (possible resident elsewhere) in this quarter-year (1 for yes, 0 for no). This variable takes a value

[13] In this case, the Key Officers publication that we obtained from May 2003, via the Internet Archive, still lists the pre-coup embassy personnel as being present at the embassy. We believe these sorts of delays in updating the directories are fairly common throughout the data.

of 0 when the U.S. has severed diplomatic relations with the country, or when a U.S. diplomatic mission has opened in the country prior to formal diplomatic recognition. For example: the U.S. recognized the independence of Belize in September 1981; the U.S. diplomatic post in Belize City is listed as a Consulate General before that, and upgraded to Embassy afterward. All observations of Belize are coded as `has_emb` = 0 through the end of 1981, and 1 starting in 1982.

3.1.4 `has_amb`

A binary variable denoting whether there is an officer listed at an embassy, with the position "AMB", in this quarter year. This takes on a value of zero both during routine ambassadorial vacancies, and in situations in which the ambassador accredited to a given country is resident elsewhere. For example: the U.S. has a staffed embassy in Apia, Samoa starting in 1989, but the ambassador accredited to Samoa is always resident in Wellington, New Zealand; so all observations for Samoa are are coded as `has_emb` =1, and `has_amb` = 0.

3.1.5 `n_officers`

The total number of distinct key officers listed across all posts in the country (excluding IO missions). Note that this is the variable that we have validated as being highly correlated with the total number of U.S. government personnel assigned to a country. That is, `n_officers` is a strong proxy for the scale of diplomatic representation in a given country. By extension, it may proxy for the importance of the relevant bilateral relationship, although some caution is necessary.

3.1.6 `n_generalists`

The number of foreign service generalists (per the `position_category` variable) in the country.

3.1.7 `n_officers_emb`

The total number of key officers listed at the embassy in the country. (This variable is coded zero whenever `has_emb` = 0.)

3.1.8 n_generalists_emb

The number of foreign service generalists (per the position_category variable) at the embassy in the country.

3.1.9 pol_amb

Binary. Takes on a value of 1 if the ambassador in the country is a political appointee, and 0 if the ambassador is a career FSO (and NA if has_amb = 0).

3.1.10 attache_ variables

Binary. Denotes the presence or absence of an attaché from the specified category in the country. Categories include: attache_agricultural, attache_broadcasting, attache_commercial, attache_EPA, attache_FAA, attache_health, attache_law_enforcement, attache_military, attache_peace_corps, attache_scientific, attache_treasury_IRS, attache_USAID. Note that these variables, like all other aggregated measures in this dataset, only account for officers that are resident in the country.

Importantly: these variables are set to NA, rather than zero, for quarter-years in which *no* posts report hosting an attaché of a given type. This is to account for the fact that some attaché types were not included at all in the *Key Officers* directories in certain quarter-years, for reasons that are not known to us, and thus we cannot confidently determine which posts hosted them and which did not.

3.1.11 hardship

The average rate of hardship pay for posts in the country.

Hardship pay is additional compensation (paid as a percentage of an employee's base salary) "to provide additional compensation to employees for service in foreign areas where environmental conditions differ substantially from environmental conditions in the continental United States and warrant additional compensation as a recruitment and retention incentive" (3 FAM 3261.1). Hardship differentials are set at the post level (and can vary within countries). We have aggregated these to the country level as the simple average of post hardship differentials.

Values for the `hardship` and `danger` variables were recorded from the State Department website (from 1997–present: https://allowances.state.gov/web920/hardship.asp) and from "U.S. Department of State Indexes of Living Costs Abroad, Quarters Allowances, and Hardship Differentials" publications (from 1982–1996: https://www.google.com/books/edition/_/72yuwY1_ET4C?gbpv=0&kptab=editions).

3.1.12 `danger`

Binary. Takes on a value of 1 if employees are eligible for danger pay in the country.

Danger pay is additional compensation "for service at places in foreign areas where there exist conditions of civil insurrection, civil war, terrorism or war, when these conditions threaten physical harm or imminent danger to the health or well-being of an employee" (3 FAM 3272.1). Danger pay is set at the post level (and can vary within countries). Our binary coding takes on a value of 1 if danger pay is in place for any post in the given country.[14]

3.1.13 `normal_relations`

Binary. Coding comes from Malis (2021), and takes on a value of 0 if any of the following hold (and 1 otherwise): (i) the country did not have full relations with the U.S. for the entire calendar year; or (ii) there was a diplomatic dispute during an ambassadorial vacancy that extended into this year, or within a year prior to the vacancy; or (iii) the country and the U.S. have established diplomatic relations but have not yet exchanged ambassadors; or (iv) the U.S. ambassador accredited to this country is resident elsewhere.

4 Variable descriptions: Post-level data

The post-quarter-year dataset (`KOFSP_post_quarter_v1.1.csv`) takes the individual-level dataset and aggregates up to the post-quarter-year level; this means that, for example, there would be separate observations for U.S. Embassy Mexico City (Q1 2008) and U.S. Consulate General Guadalajara (Q1 2008). It includes the same post, post type, country, country code, year, quarter, and qy vari-

[14] Regulations governing danger pay have changed over time. At present, danger pay varies from 15% to 35% of base salary for applicable posts. Previously, the State Department treated danger pay as binary, awarding only a single level (25%) to applicable posts. We treat this as a binary variable to provide greater consistency over time.

ables as the individual-level dataset; the same officer count and attaché presence variables as the country-level dataset; and the additional variables listed below. As is true of the country-level data, the post-level data only includes post-quarter-years with key officers resident at the post.

4.0.1 `senior_officer_position`

The position title of the senior officer at the given post in the given quarter-year. For a more thorough discussion of the values that this variable may take, see discussion of the `senior_officer` variable in Section 2.3.12.

4.0.2 `senior_officer_name`

The full name of the senior officer at the given post in the given quarter-year. This is taken directly from the `officer_name` variable in the individual-level dataset; see Section 2.3.2 for further details.

References

Arel-Bundock, Vincent, Nils Enevoldsen, and CJ Yetman. 2018. "countrycode: An R package to convert country names and country codes." *Journal of Open Source Software* 3 (28): 848.

Malis, Matt. 2021. "Conflict, cooperation, and delegated diplomacy." *International Organization* 75 (4): 1018–1057.