

Table of Contents

Introduction 2

Bills 6

Calendar 8

Chamber Membership 9

Committees 11

Floor Speeches 13

Members of the House of Commons 16

Votes 18

Introduction

The project uses text data on day-to-day legislative activity to map together bills, motions, questions, speeches, and votes. The project also includes data on members, constituencies, and committees, among other things. The project covers XX parliaments: the XX Parliament (starting XX XX, 2XXX) through the most recent 57th Parliament (we end the day before the most recent 2019-20 session started on 14 October, 2019).

Legislative Process

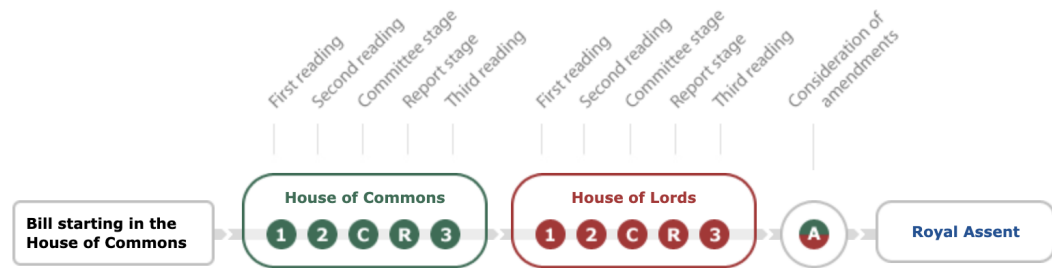
This project tracks the progression of bills through the legislative process and connects them with other legislative activity, including questions, speeches, and votes. The legislative process consists of the following steps, which are also outlined in Figure 1:

- **The introduction and first reading of a bill:** Though largely a formality, the short title of bills are read aloud and they are ordered to be “printed” in the official journal. This indicates that the bill can proceed to the next stage, which the first chance for MPs to debate the bill’s general principles and themes.
- **The second reading and referral to a committee:** The Government minister, spokesperson or MP sponsoring the bill begins the debate and the official Opposition spokesperson responds. The debate continues with other Opposition parties and backbenchers giving their opinions. At the end of the debate, the House of Commons as a whole decides by voting whether the bill should be given a second reading/proceed to the next stage. It is possible for a bill to have a second reading with no debate, as long as MPs agree. The second reading usually takes place no sooner than two weekends after first reading.
- **The committee stage:** The committee reviews the text of the bill and each clause or amendments may be debated. Every clause in the bill is agreed to, although this may happen without debate (particularly under a programme order). Amendments for discussion are selected by the committee chairman and only committee members can vote on amendments. Amendments proposed by MPs to the bill are published daily, so if the bill is amended it will be reprinted before its next stage. Once the committee stage is finished, the bill returns to the full chamber for further debate and amendment proposals.
- **The report stage and third reading:** During this stage, members can propose amendments or new clauses to the bill that was passed in committee. There is no set time period between the

end of committee stage and the start of the report stage. All MPs may speak and vote, and debates may last several days. This stage is typically followed directly by debate on the third reading, which is the final chance for MPs to debate the bills' contents. Amendments cannot be made during the third reading. At the end of the debate, MPs vote on whether to approve the bill and send it to the House of Lords for its first reading.

- **Consideration and passage by the House of Lords:** After the bill has passed through the third reading in both the House of Commons and House of Lords, it returns to the House of Commons first (where it started) for the second round of amendments. The Lords then debate the bill and can vote on amendments. A bill goes back and forth between The House of Commons and the House of Lords until both approve the same version of the bill.¹
- **Royal Assent and coming into force:** If the same version of a bill is passed by both the Commons and the Lords, it is granted Royal Assent by the Monarch's agreement to make the bill into an act.

Figure 1: Bill initiation and stages of completion in both upper and lower British chambers.



Source: www.parliament.uk.

¹ Rarely the bill fails when the two Houses do not reach agreement. If so, the Commons can use the "Parliament Acts" to pass the bill without the consent of the Lords in the following session.

Overview of Datasets

The project includes XX datasets organized into X sectors. Each sector focuses on a different aspect of legislative activity. Each dataset has four categories of variables: documentation variables, grouping variables, sorting variables, and substantive variables. Documentation variables include a variable that indicates the version of the dataset and a variable that indicates the location of the dataset within the project directory, which is organized by sector. Grouping variables can be used to collapse or merge datasets. Sorting variables can be used to sort observations. Substantive variables include all variables.

Sector 1: Calendars

Sector 1 contains data related to sittings of the House of Commons and committees in the House of Commons. This sector includes two datasets that record the calendar of sittings for the House of Commons and the calendar of sittings for standing committees in the House of Commons.

- `uk_calendar.csv` records the dates of all sittings of the House of Commons. There is one observation per sitting per parliament.
-

Sector 2: Official Publications

Sector 2 contains data on the day-to-day activity of the House of Commons based the **Order Paper**. The **Order Paper** are published each sitting day.

- `uk_order_papers.csv` tracks items of business on the **Order Paper**. There is one observation per item per edition of the **Order Paper**. See `uk_calendar.csv`.
-

Sector 3: Members

Sector 3 includes seven datasets that contain information about members of the House of Commons.

- `uk_members.csv` records all unique members of the House of Commons. There is one observation per unique member of the House of Commons across all parliaments covered in the project (the XXth Parliament through the XXth Parliament). This dataset includes the name, constituency, and party of each member. A unique path, `member_path`, is assigned to each

member. This path is used in a variety of other datasets to uniquely identify members. In other datasets, it is called `member_ID`.

- `uk_chamber_membership.csv` tracks the membership of the House of Commons across parliament sessions. There is one observation per member per parliament. This dataset includes the name, constituency, and party of each member.
 - `uk_constituencies.csv` records all of the unique constituencies that members are elected by. There is one observation per constituency. The dataset includes the name of the constituency. A unique path, `constituency_path`, is assigned to each constituency. This path is used in a variety of other datasets to uniquely identify constituencies. In other datasets, it is called `constituency_ID`.
 - `uk_elections.csv` records election results for all general elections and by-elections. There is one observation per candidate per race per election. There is one race per constituency in general elections. This dataset includes all general elections and all by-elections. In each election, there is one race for each constituency. The data includes the date and type of the election, name and party of each candidate, the constituency and province of the constituency, and whether the candidate was defeated, elected, or reelected.
 - `uk_ministries.csv` tracks the composition of the government. There is one observation per member per ministerial portfolio per government. Each observation indicates the member associated with the portfolio. Members can hold multiple posts at the same time, so there can be multiple observations per member per government. Posts can also change hands, so there can be multiple observations per post per government. The dataset includes the start and end dates of the period of time that each member held each post. The source of the data is `SOURCE`.
-

Bills

uk_house_bills.csv

0 MB

This dataset includes all bills introduced from the 93th Parliament through the 116th Parliament.

Grouping Variables

observation_path A path that uniquely identifies each observation in the dataset in the format /parliament-56/chamber-1/bill-1. Identical in this dataset to **bill_ID**.

parliament_path A path that uniquely identifies each parliament in the format /parliament-56.

chamber_path A path that uniquely identifies each chamber within a session in the format /parliament-56/chamber-1.

bill_path A path that uniquely identifies each bill in the format /parliament-56/chamber-1/bill-1.

Sorting Variables

observation_number The number of the observation. Indicates the default sorting order. Sorted by **congress**, then by **chamber_number**, then by **bill_number**.

congress_number The number of the parliament.

chamber_number A number assigned to each chamber. The House of Commons is coded **1** and the House of Lords is coded **2**.

bill_number The number of the bill. Note that this variable does not uniquely identify bills because bill numbers repeat across chambers and sessions.

Variables

`date` The date the bill was introduced in the format `YYYY-MM-DD`.

`bill_type` The type of the bill. Possible values include:

- Bills (H.R. or S.)
- Amendments (H.Amdt. or S.Amdt.)
- Resolutions (H.Res. or S.Res.)
- Joint Resolutions (H.J.Res. or S.J.Res.)
- Concurrent Resolutions (H.Con.Res. or S.Con.Res.)

`bill_title` The official title of the bill.

`member_name` The full name of the member who sponsored the bill. This variable does not uniquely identify members because members can have the same name.

`member_ID` A path segment that uniquely identifies each member of the House of Representatives in the format `/chamber-1/member-1`. See `us_house_member.csv`.

`state` The constituency of who sponsored the bill.

`district` The constituency of who sponsored the bill.

`committee_name` The name of the committee. The committee the bill is under.

Calendar

us_calendar.csv

94KB

This dataset indicates the dates where the parliament is active from 1995 to 2020.

Variables

date

The date of the activity.

Chamber Membership

us_house_membership.csv

2.7 MB

This dataset indicates the members of the House of Commons in each parliament from the 93th parliament through the 116th parliament. There is one observation per member per parliament.

Grouping Variables

observation_path A path that uniquely identifies each observation in the dataset in the format `/parliament-56/chamber-1/member-1`. Identical in this dataset to **member_ID**.

parliament_path A path that uniquely identifies each parliament in the format `/parliament-56`.

chamber_path A path that uniquely identifies each chamber per parliament in the format `/parliament-56/chamber-1`.

member_path A path that uniquely identifies each member per chamber per parliament in the format `/parliament-56/chamber-1/member-1`.

Sorting Variables

observation_number The number of the observation. Indicates the default sorting order. Sorted by **parliament_number**, then by **chamber_number**, then by **member_number**.

parliament_number The number of the parliament.

chamber_number A number assigned to each chamber. The House of Commons is coded **1** and the House of Lords is coded **2**.

`member_number` A number assigned to each member. Assigned with members sorted by last name, then by first name, then by constituency. Not comparable across groups.

Variables

`chamber_name` The name of the chamber.

`full_name` The first and last name of the member.

`first_name` The first name of the member.

`middle_name` The middle name of the member.

`last_name` The last name of the member.

`member_ID` A path segment that uniquely identifies each member of the House of Representatives in the format `/chamber-1/member-1`. See `us_house_member.csv`.

`state` The state portion of member constituency.

`district` The district portion of member constituency

`party` The party of the member. Members sometimes switch parties while in office. This variable indicates their party at the time they were originally elected.

`year_start` The start date for each member in each parliament in the format `YYYY-MM-DD`. The start date of the parliament unless the member was elected in a by-election.

`year_end` The end date for each member in each parliament in the format `YYYY-MM-DD`. The end date of the parliament unless the member died in office or resigned.

Committees

us_committee.csv

1.2 MB

This dataset includes all committees in the House of Representatives and the Senate. No committees changed during this period. There is one observation per committee and subcommittee.

Grouping Variables

observation_path A path that uniquely identifies each observation in the dataset in the format `/chamber-1/committee-1`. Identical in this dataset to **committee_path**.

chamber_path A path that uniquely identifies each chamber in the format `/chamber-1`.

committee_path A path that uniquely identifies each committee in the format `/chamber-1/committee-1`.

Sorting Variables

observation_number The number of the observation. Indicates the default sorting order. Sorted by **chamber_number**, then by **committee_number**.

chamber_number A number assigned to each chamber. The House of Commons is coded `1` and the House of Lords is coded `2`.

committee_number A number assigned to each committee within a session. Assigned with committees sorted by name.

Variables

committee_name The name of the committee.

parent_name The parent committee of the sub committee.

past_names The past names of the committee.

type House or Senate

committee_id Committee ID as listed on the parliament website.

Floor Speeches

us_house_speech_xx.csv

2.8 GB

This dataset includes all speeches delivered by members of the House of Representatives from the 97th Parliament through the 114th Parliament (includes speeches by the presiding officer but not speeches by non-members). There is one observation per paragraph per speech. The speeches are put into separate files, one for each parliament, to avoid too huge of a dataset.

Grouping Variables

observation_path

A path that uniquely identifies each observation in the dataset in the format /parliament-56chamber-1/speech-1. Identical in this dataset to `speech_path`.

parliament_path

A path that uniquely identifies each parliament in the format /parliament-97.

chamber_path

A path that uniquely identifies each chamber within a session in the format /parliament-97/chamber-1.

speech_path

A path that uniquely identifies each speech in the format /parliament-97/chamber-1/speech-1/.

Sorting Variables

`parliament_number` The number of the parliament.

`chamber_number` A number assigned to each chamber. The House of Commons is coded `1` and the House of Lords is coded `2`.

`speech_number` The number of the speech within a sitting. Note that this variable does not uniquely identify speeches because speech numbers repeat across sittings.

Variables

`speech_date` The date of the speech in the format `YYYY-MM-DD`.

`chamber_name` The name of the chamber.

`full_name` The first and last name of the member who gave the speech.

`first_name` The first name of the member who gave the speech.

`middle_name` The middle name of the member who gave the speech.

`last_name` The last name of the member who gave the speech.

`gender` The gender of the member who gave the speech.

`state` The state portion of the member who gave the speech.

`district` The district portion of the member who gave the speech.

`party` The name of the party of the member who gave the speech.

`nonvoting` Whether the speaker is voting

`paragraph_text` The text of the paragraph.

`word_count` The word count of the paragraph.

`char_count` The character count of the paragraph.

Members of the House of Commons

uk_lower_members.csv

0 MB

This dataset includes all members of the House of Commons from the XXth parliament through the XXth parliament. There is one observation per member.

Grouping Variables

observation_path A path that uniquely identifies each observation in the dataset in the format `/chamber-1/member-1`. Identical in this dataset to **member_path**.

chamber_path A path that uniquely identifies each chamber in the format `/chamber-1`.

member_path A path that uniquely identifies each member in the format `/chamber-1/member-1`.

Sorting Variables

observation_number The number of the observation. Indicates the default sorting order. Sorted by **chamber_number**, then by **member_number**.

chamber_number A number assigned to each chamber. The House of Commons is coded `1` and the House of Lords is coded `2`.

member_number A number assigned to each member. Assigned with members sorted by last name, then by first name, then by constituency. Not comparable across groups.

Variables

chamber_name The name of the chamber.

full_name The first and last name of the member.

`first_name` The first name of the member.

`last_name` The last name of the member.

`constituency_name` The name of the constituency of the member.

`constituency_ID` A path segment that uniquely identifies each constituency in the format `/chamber-1/constituency-1`. See `uk_lower_constituencies.csv`.

`party_name` The party of the member. Members sometimes switch parties while in office. This variable indicates their party at the time they were originally elected. Possible values include:

`start_date` The date that the member started serving in parliament in the format `YYYY-MM-DD`.

`end_date` The date that the member stopped serving in parliament in the format `YYYY-MM-DD`.

Votes

us_house_votes_xxxx.csv

8.5 MB

This dataset includes all the votes that occurred from 1990 to 2020. Votes for each year are put into separate files to avoid huge file size.

Sorting

Variables

observation_number The number of the observation. Indicates the default sorting order. Sorted by **roll_number**.

roll_number An official number that uniquely identified each vote in each congress.

Variables

congress_number The number of the parliament.

session_number The number of the session within a parliament. This variable does not uniquely identify sessions because session numbers repeat across parliaments.

chamber_number A number assigned to each chamber. The House of Commons is coded **1** and the House of Lords is coded **2**.

year The year the vote happened in.

bill_number The number of the bill. Note that this variable does not uniquely identify bills because bill numbers repeat across chambers and sessions.

vote_question The question that the vote is addressing.

vote_type The type of the vote.

date The date of the vote.

last_name	Last name of the member who voted.
party	The party of the member
state	The constituency of the member
role	The role the member serves in the voting process.
vote	The vote of the member, can be "YEA", "NAY", or "NOT VOTING".
majority	The majority party of the members who voted.
vote_result	Whether the vote is passed.
