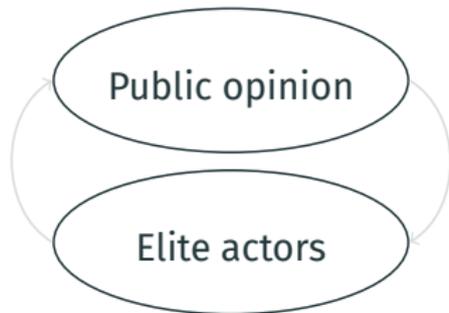# Preference Endogeneity and the Problem of Electoral Representation

Chitralekha Basu

July 5, 2022

Cologne Center for Comparative Politics
University of Cologne

Preference endogeneity
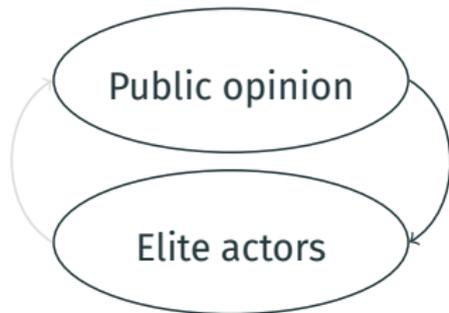

Public opinion

Elite actors

Preference endogeneity: the endogeneity of citizens' political preferences to elite communication and behavior.

May refer to:

- Position endogeneity (outcome: citizens' policy opinions).
  E.g. Lenz 2009, 2012; Broockman and Butler 2017; Barber and Pope 2019; Slothuus and Bisgaard 2021.

- Salience endogeneity (outcome: citizens' issue priorities).
  E.g. Druckman and Holmes 2004; Tesler 2015; Matthews 2019.

- Belief endogeneity (outcome: citizens' factual beliefs).
  E.g. Tesler 2018; Bisgaard and Slothuus 2018; Clayton et al. 2021.

Preference endogeneity

Preference endogeneity: the endogeneity of citizens' political preferences to elite communication and behavior.

May refer to:

- Position endogeneity (outcome: citizens' policy opinions).
  E.g. Lenz 2009, 2012; Broockman and Butler 2017; Barber and Pope 2019; Slothuus and Bisgaard 2021.

- Salience endogeneity (outcome: citizens' issue priorities).
  E.g. Druckman and Holmes 2004; Tesler 2015; Matthews 2019.

- Belief endogeneity (outcome: citizens' factual beliefs).
  E.g. Tesler 2018; Bisgaard and Slothuus 2018; Clayton et al. 2021.

Preference endogeneity
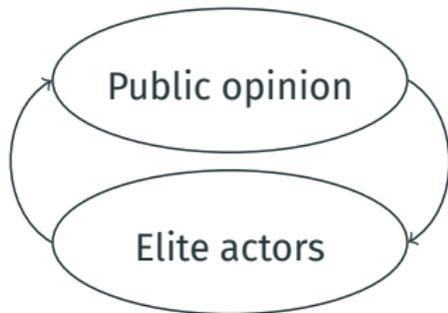

Public opinion

Elite actors

Preference endogeneity: the endogeneity of citizens' political preferences to elite communication and behavior.

May refer to:

- Position endogeneity (outcome: <u>citizens' policy opinions</u>).
  E.g. Lenz 2009, 2012; Broockman and Butler 2017; Barber and Pope 2019; Slothuus and Bisgaard 2021.

- Salience endogeneity (outcome: <u>citizens' issue priorities</u>).
  E.g. Druckman and Holmes 2004; Tesler 2015; Matthews 2019.

- Belief endogeneity (outcome: <u>citizens' factual beliefs</u>).
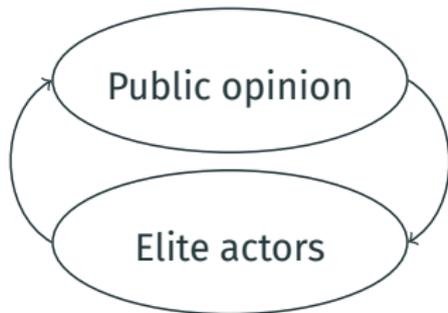 E.g. Tesler 2018; Bisgaard and Slothuus 2018; Clayton et al. 2021.

Preference endogeneity: the endogeneity of citizens' political preferences to elite communication and behavior.

May refer to:

- Position endogeneity (outcome: citizens' policy opinions).
    E.g. Lenz 2009, 2012; Broockman and Butler 2017; Barber and Pope 2019; Slothuus and Bisgaard 2021.

- Salience endogeneity (outcome: citizens' issue priorities).
    E.g. Druckman and Holmes 2004; Tesler 2015; Matthews 2019.

- Belief endogeneity (outcome: citizens' factual beliefs).
    E.g. Tesler 2018; Bisgaard and Slothuus 2018; Clayton et al. 2021.
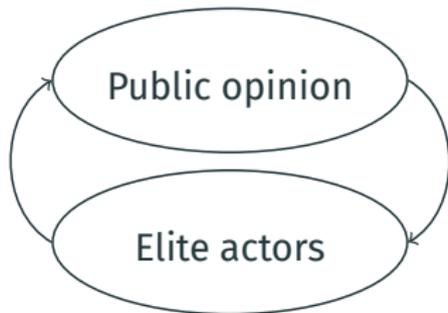
Preference endogeneity

Preference endogeneity: the endogeneity of citizens' political preferences to elite communication and behavior.

May refer to:

- Position endogeneity (outcome: citizens' policy opinions).
    E.g. Lenz 2009, 2012; Broockman and Butler 2017; Barber and Pope 2019; Slothuus and Bisgaard 2021.

- Salience endogeneity (outcome: citizens' issue priorities).
    E.g. Druckman and Holmes 2004; Tesler 2015; Matthews 2019.

- Belief endogeneity (outcome: citizens' factual beliefs).
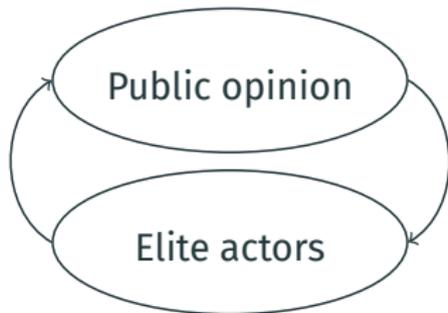   E.g. Tesler 2018; Bisgaard and Slothuus 2018; Clayton et al. 2021.

Preference endogeneity: the endogeneity of citizens' political preferences to elite communication and behavior.

May refer to:

- Position endogeneity (outcome: citizens' policy opinions).
  E.g. Lenz 2009, 2012; Broockman and Butler 2017; Barber and Pope 2019; Slothuus and Bisgaard 2021.

- Salience endogeneity (outcome: citizens' issue priorities).
  E.g. Druckman and Holmes 2004; Tesler 2015; Matthews 2019.

- Belief endogeneity (outcome: citizens' factual beliefs).
  E.g. Tesler 2018; Bisgaard and Slothuus 2018; Clayton et al. 2021.

Preference endogeneity

Preference endogeneity: the endogeneity of citizens' political preferences to elite communication and behavior.

May refer to:

- Position endogeneity (outcome: citizens' policy opinions).
    E.g. Lenz 2009, 2012; Broockman and Butler 2017; Barber and Pope 2019; Slothuus and Bisgaard 2021.

- Salience endogeneity (outcome: citizens' issue priorities).
    E.g. Druckman and Holmes 2004; Tesler 2015; Matthews 2019.

- Belief endogeneity (outcome: citizens' factual beliefs).
    E.g. Tesler 2018; Bisgaard and Slothuus 2018; Clayton et al. 2021.
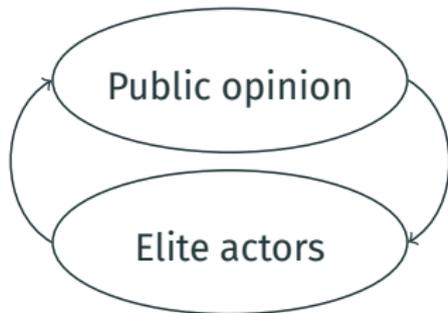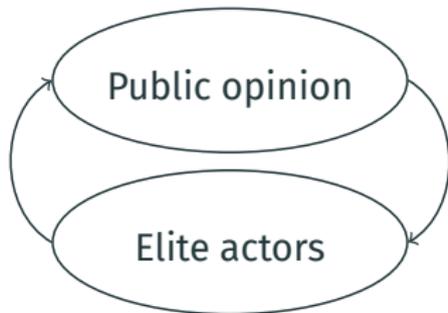
Preference endogeneity

Preference endogeneity: the endogeneity of citizens' political preferences to elite communication and behavior.

May refer to:

- Position endogeneity (outcome: citizens' policy opinions).
  E.g. Lenz 2009, 2012; Broockman and Butler 2017; Barber and Pope 2019; Slothuus and
  Bisgaard 2021.

- Salience endogeneity (outcome: citizens' issue priorities).
  E.g. Druckman and Holmes 2004; Tesler 2015; Matthews 2019.

- Belief endogeneity (outcome: citizens' factual beliefs).
  E.g. Tesler 2018; Bisgaard and Slothuus 2018; Clayton et al. 2021.

- Political equality: a widely held commitment among theorists of liberal democracy and of political representation.

  - Also motivates much empirical research on representation.

- Asks that we have an equal distribution of effective political power.

  - I.e. that all citizens have an equal opportunity to influence political decisions.
  - Though not necessarily actual influence over the decisions themselves.

How do we measure the distribution of effective political power in the presence of preference endogeneity?

- Political equality: a widely held commitment among theorists of liberal democracy and of political representation.
  - Also motivates much empirical research on representation.

- Asks that we have an equal distribution of effective political power.
  - I.e. that all citizens have an equal opportunity to influence political decisions.
  - Though not necessarily actual influence over the decisions themselves.

How do we measure the distribution of effective political power in the presence of preference endogeneity?

- Political equality: a widely held commitment among theorists of liberal democracy and of political representation.

  - Also motivates much empirical research on representation.

- Asks that we have an equal distribution of effective political power.

  - I.e. that all citizens have an equal opportunity to influence political decisions.

  - Though not necessarily actual influence over the decisions themselves.

How do we measure the distribution of effective political power in the presence of preference endogeneity?

Motivation

- Political equality: a widely held commitment among theorists of liberal democracy and of political representation.
    - Also motivates much empirical research on representation.

- Asks that we have an equal distribution of effective political power.
    - I.e. that all citizens have an equal opportunity to influence political decisions.
    - Though not necessarily actual influence over the decisions themselves.

How do we measure the distribution of effective political power in the presence of preference endogeneity?

Motivation

- Political equality: a widely held commitment among theorists of liberal democracy and of political representation.

  - Also motivates much empirical research on representation.

- Asks that we have an equal distribution of effective political power.

  - I.e. that all citizens have an equal opportunity to influence political decisions.
  - Though not necessarily actual influence over the decisions themselves.

How do we measure the distribution of effective political power in the presence of preference endogeneity?

Motivation

- Political equality: a widely held commitment among theorists of liberal democracy and of political representation.

  - Also motivates much empirical research on representation.

- Asks that we have an equal distribution of effective political power.

  - I.e. that <u>all</u> citizens have an equal opportunity to influence political decisions.
  - Though not necessarily actual influence over the decisions themselves.

How do we measure the distribution of effective political power in the presence of preference endogeneity?

The problem

Preference endogeneity introduces new asymmetries of power into the principal-agent relationship that links representatives and their constituents.

- Elected representatives can use their position to reshape citizens' preferences in a favourable direction.

- Often desirable for citizens who face constraints of time, information and expertise.

- Might even be retrospectively approved by the represented as part of a 'mutually educative' process (Mansbridge 2003).

- But...

    - Same power asymmetries mean citizens do not receive time or information to evaluate if their new preferences are more aligned with their interests.

    - Not fully addressed by the treatment and selection mechanisms of democracy (Landa and Pevnick 2020).

This is distinct from the challenges associated with estimating the causal effect of public opinion on policy if preferences are endogenous.

The problem

Preference endogeneity introduces new asymmetries of power into the principal-agent relationship that links representatives and their constituents.

- Elected representatives can use their position to reshape citizens' preferences in a favourable direction.

- Often desirable for citizens who face constraints of time, information and expertise.

- Might even be retrospectively approved by the represented as part of a 'mutually educative' process (Mansbridge 2003).

- But...

    - Same power asymmetries mean citizens do not receive time or information to evaluate if their new preferences are more aligned with their interests.

    - Not fully addressed by the treatment and selection mechanisms of democracy (Landa and Pevnick 2020).

This is distinct from the challenges associated with estimating the causal effect of public opinion on policy if preferences are endogenous.

The problem

Preference endogeneity introduces new asymmetries of power into the principal-agent relationship that links representatives and their constituents.

- Elected representatives can use their position to reshape citizens' preferences in a favourable direction.

- Often desirable for citizens who face constraints of time, information and expertise.

- Might even be retrospectively approved by the represented as part of a 'mutually educative' process (Mansbridge 2003).

- But...

  - Same power asymmetries mean citizens do not receive time or information to evaluate if their new preferences are more aligned with their interests.

  - Not fully addressed by the treatment and selection mechanisms of democracy (Landa and Pevnick 2020).

This is distinct from the challenges associated with estimating the causal effect of public opinion on policy if preferences are endogenous.

The problem

Preference endogeneity introduces new asymmetries of power into the principal-agent relationship that links representatives and their constituents.

- Elected representatives can use their position to reshape citizens' preferences in a favourable direction.

- Often desirable for citizens who face constraints of time, information and expertise.

- Might even be retrospectively approved by the represented as part of a 'mutually educative' process (Mansbridge 2003).

- But...

  - Same power asymmetries mean citizens do not receive time or information to evaluate if their new preferences are more aligned with their interests.

  - Not fully addressed by the treatment and selection mechanisms of democracy (Landa and Pevnick 2020).

This is distinct from the challenges associated with estimating the causal effect of public opinion on policy if preferences are endogenous.

The problem

Preference endogeneity introduces new asymmetries of power into the principal-agent relationship that links representatives and their constituents.

- Elected representatives can use their position to reshape citizens' preferences in a favourable direction.

- Often desirable for citizens who face constraints of time, information and expertise.

- Might even be retrospectively approved by the represented as part of a 'mutually educative' process (Mansbridge 2003).

- But...

  - Same power asymmetries mean citizens do not receive time or information to evaluate if their new preferences are more aligned with their interests.

  - Not fully addressed by the treatment and selection mechanisms of democracy (Landa and Pevnick 2020).

This is distinct from the challenges associated with estimating the causal effect of public opinion on policy if preferences are endogenous.

The problem

Preference endogeneity introduces new asymmetries of power into the principal-agent relationship that links representatives and their constituents.

- Elected representatives can use their position to reshape citizens' preferences in a favourable direction.

- Often desirable for citizens who face constraints of time, information and expertise.

- Might even be retrospectively approved by the represented as part of a 'mutually educative' process (Mansbridge 2003).

- But...

  - Same power asymmetries mean citizens do not receive time or information to evaluate if their new preferences are more aligned with their interests.

  - Not fully addressed by the treatment and selection mechanisms of democracy (Landa and Pevnick 2020).

This is distinct from the challenges associated with estimating the causal effect of public opinion on policy if preferences are endogenous.

The problem

Preference endogeneity introduces new asymmetries of power into the principal-agent relationship that links representatives and their constituents.

- Elected representatives can use their position to reshape citizens' preferences in a favourable direction.

- Often desirable for citizens who face constraints of time, information and expertise.

- Might even be retrospectively approved by the represented as part of a 'mutually educative' process (Mansbridge 2003).

- But...

    - Same power asymmetries mean citizens do not receive time or information to evaluate if their new preferences are more aligned with their interests.

    - Not fully addressed by the treatment and selection mechanisms of democracy (Landa and Pevnick 2020).

This is distinct from the challenges associated with estimating the causal effect of public opinion on policy if preferences are endogenous.

The problem

Preference endogeneity introduces new asymmetries of power into the principal-agent relationship that links representatives and their constituents.

- Elected representatives can use their position to reshape citizens' preferences in a favourable direction.

- Often desirable for citizens who face constraints of time, information and expertise.

- Might even be retrospectively approved by the represented as part of a 'mutually educative' process (Mansbridge 2003).

- But...

  - Same power asymmetries mean citizens do not receive time or information to evaluate if their new preferences are more aligned with their interests.

  - Not fully addressed by the treatment and selection mechanisms of democracy (Landa and Pevnick 2020).

This is distinct from the challenges associated with estimating the causal effect of public opinion on policy if preferences are endogenous.

Given the possibility of preference endogeneity, **any measure of effective political power must also seek to measure** individuals' preference and influence within public political debate.

- Materially affects citizens' opportunities to influence political outcomes, via their influence on the opinions of others.

- May be indirect, via the reflection of individuals' preferences by elected representatives and other elite actors.

Empirical research on representation does not typically seek to measure this quantity.

Given the possibility of preference endogeneity, **any measure of effective political power must also seek to measure** individuals' preference and influence within public political debate.

- Materially affects citizens' opportunities to influence political outcomes, via their influence on the opinions of others.

- May be indirect, via the reflection of individuals' preferences by elected representatives and other elite actors.

Empirical research on representation does not typically seek to measure this quantity.

Given the possibility of preference endogeneity, **any measure of effective political power must also seek to measure** individuals' preference and influence within public political debate.

- Materially affects citizens' opportunities to influence political outcomes, via their influence on the opinions of others.

- May be indirect, via the reflection of individuals' preferences by elected representatives and other elite actors.

Empirical research on representation does not typically seek to measure this quantity.

Given the possibility of preference endogeneity, **any measure of effective political power must also seek to measure** individuals' preference and influence within public political debate.

- Materially affects citizens' opportunities to influence political outcomes, via their influence on the opinions of others.

- May be indirect, via the reflection of individuals' preferences by elected representatives and other elite actors.

Empirical research on representation does not typically seek to measure this quantity.

The solution

A measurable criterion for the distribution of effective political power in such contexts is the degree of **communicative representation** in a polity: how much speech by elite actors proportionately represents *and* responds to the entire distribution of preferences on a large number of issues.

- If public opinion is partly endogenous to elite communication, a high level of communicative representation is a necessary but not sufficient condition for political equality.

  - Low communicative representation $\implies$ elite actors with certain preferences and priorities are overrepresented in public debate $\implies$ some voters have disproportionate political power.

  - High communicative representation $\implies$ voter preferences and priorities are proportionately represented in public debate $\implies$ greater political equality.

A measurable criterion for the distribution of effective political power in such contexts is the degree of **communicative representation** in a polity: how much speech by elite actors proportionately represents *and* responds to the entire distribution of preferences on a large number of issues.

- If public opinion is partly endogenous to elite communication, a high level of communicative representation is <u>a necessary but not sufficient condition</u> for political equality.

    - Low communicative representation $\implies$ elite actors with certain preferences and priorities are overrepresented in public debate $\implies$ some voters have disproportionate political power.

    - High communicative representation $\implies$ voter preferences and priorities are proportionately represented in public debate $\implies$ greater political equality.

A measurable criterion for the distribution of effective political power in such contexts is the degree of **communicative representation** in a polity: how much speech by elite actors proportionately represents *and* responds to the entire distribution of preferences on a large number of issues.

- If public opinion is partly endogenous to elite communication, a high level of communicative representation is a necessary but not sufficient condition for political equality.

    - Low communicative representation $\implies$ elite actors with certain preferences and priorities are overrepresented in public debate $\implies$ some voters have disproportionate political power.

    - High communicative representation $\implies$ voter preferences and priorities are proportionately represented in public debate $\implies$ greater political equality.

The solution

A measurable criterion for the distribution of effective political power in such contexts is the degree of **communicative representation** in a polity: how much speech by elite actors proportionately represents *and* responds to the entire distribution of preferences on a large number of issues.

- If public opinion is partly endogenous to elite communication, a high level of communicative representation is <u>a necessary but not sufficient condition</u> for political equality.

    - Low communicative representation $\implies$ elite actors with certain preferences and priorities are overrepresented in public debate $\implies$ some voters have disproportionate political power.

    - High communicative representation $\implies$ voter preferences and priorities are proportionately represented in public debate $\implies$ greater political equality.

The solution (cont'd)

I argue that the degree of communicative representation provides a useful barometer for how far the political environment in a given place and time deviates from the ideal of political equality.

- Informative even if preferences *or* priorities are endogenous.

- *Not* intended as a measure of the causal effect of voter preferences on policymakers!

- Rather:
    - If there is high equality of influence over public opinion, we should observe high communicative representation.
    - Observed high communicative representation $\implies$ (likely) high equality of influence over public opinion.

      $\because$ High communicative representation $\implies$ all (current) opinions have approx. equal weight in key public fora, even if those opinions are endogenous.

I argue that the degree of communicative representation provides a useful barometer for how far the political environment in a given place and time deviates from the ideal of political equality.

- Informative <u>even if preferences *or* priorities are endogenous.</u>

- *Not* intended as a measure of the causal effect of voter preferences on policymakers!

- Rather:
    - If there is high equality of influence over public opinion, we should observe high communicative representation.
    - Observed high communicative representation $\implies$ (likely) high equality of influence over public opinion.

        $\therefore$ High communicative representation $\implies$ all (current) opinions have approx. equal weight in key public fora, even if those opinions are endogenous.

I argue that the degree of communicative representation provides a useful barometer for how far the political environment in a given place and time deviates from the ideal of political equality.

- Informative <u>even if preferences *or* priorities are endogenous.</u>

- *Not* intended as a measure of the causal effect of voter preferences on policymakers!

- Rather:
    - If there is high equality of influence over public opinion, we should observe high communicative representation.
    - Observed high communicative representation $\implies$ (likely) high equality of influence over public opinion.

      $\therefore$ High communicative representation $\implies$ all (current) opinions have approx. equal weight in key public fora, even if those opinions are endogenous.

The solution (cont'd)

I argue that the degree of communicative representation provides a useful barometer for how far the political environment in a given place and time deviates from the ideal of political equality.

- Informative <u>even if preferences *or* priorities are endogenous.</u>

- *Not* intended as a measure of the causal effect of voter preferences on policymakers!

- Rather:
  - If there is high equality of influence over public opinion, we should observe high communicative representation.
  - Observed high communicative representation $\implies$ (likely) high equality of influence over public opinion.

    $\therefore$ High communicative representation $\implies$ all (current) opinions have approx. equal weight in key public fora, even if those opinions are endogenous.

I argue that the degree of communicative representation provides a useful barometer for how far the political environment in a given place and time deviates from the ideal of political equality.

- Informative <u>even if preferences *or* priorities are endogenous.</u>

- ***Not*** intended as a measure of the causal effect of voter preferences on policymakers!

- Rather:
    - If there is high equality of influence over public opinion, we should observe high communicative representation.
    - Observed high communicative representation $\implies$ (likely) high equality of influence over public opinion.
    - $\because$ High communicative representation $\implies$ all (current) opinions have approx. equal weight in key public fora, even if those opinions are endogenous.

The solution (cont'd)

I argue that the degree of communicative representation provides a useful barometer for how far the political environment in a given place and time deviates from the ideal of political equality.

- Informative <u>even if preferences *or* priorities are endogenous.</u>

- *Not* intended as a measure of the causal effect of voter preferences on policymakers!

- Rather:
    - If there is high equality of influence over public opinion, we should observe high communicative representation.
    - Observed high communicative representation $\implies$ (likely) high equality of influence over public opinion.

        $\because$ High communicative representation $\implies$ all (current) opinions have approx. equal weight in key public fora, even if those opinions are endogenous.

The solution (cont'd)

I argue that the degree of communicative representation provides a useful barometer for how far the political environment in a given place and time deviates from the ideal of political equality.

- Informative <u>even if preferences *or* priorities are endogenous.</u>

- *Not* intended as a measure of the causal effect of voter preferences on policymakers!

- Rather:
    - If there is high equality of influence over public opinion, we should observe high communicative representation.
    - Observed high communicative representation $\implies$ (likely) high equality of influence over public opinion.

    $\therefore$ High communicative representation $\implies$ all (current) opinions have approx. equal weight in key public fora, even if those opinions are endogenous.

Measuring communicative representation

The concept of communicative representation has two components:

- **Communicative congruence:** proportionate representation of the entire distribution of voter preferences, across issues, in elite speech.
- **Communicative responsiveness:** changes in the distribution of preferences produce corresponding changes in preferences expressed in elite speech.

As proof of concept, I analyse communicative representation on immigration in the United Kingdom between 2015 and 2019, using legislative speech as a proxy for elite speech.

- Split into: May 2015-June 2016, June 2016-June 2017, June 2017-July 2019.

The concept of communicative representation has two components:

- **Communicative congruence:** proportionate representation of the entire distribution of voter preferences, across issues, in elite speech.
- Communicative responsiveness: changes in the distribution of preferences produce corresponding changes in preferences expressed in elite speech.

As proof of concept, I analyse communicative representation on immigration in the United Kingdom between 2015 and 2019, using legislative speech as a proxy for elite speech.

- Split into: May 2015-June 2016, June 2016-June 2017, June 2017-July 2019.

The concept of communicative representation has two components:

- **Communicative congruence:** proportionate representation of the entire distribution of voter preferences, across issues, in elite speech.
- **Communicative responsiveness:** changes in the distribution of preferences produce corresponding changes in preferences expressed in elite speech.

As proof of concept, I analyse communicative representation on immigration in the United Kingdom between 2015 and 2019, using legislative speech as a proxy for elite speech.

- Split into: May 2015-June 2016, June 2016-June 2017, June 2017-July 2019.

The concept of communicative representation has two components:

- **Communicative congruence:** proportionate representation of the entire distribution of voter preferences, across issues, in elite speech.
- **Communicative responsiveness:** changes in the distribution of preferences produce corresponding changes in preferences expressed in elite speech.

As proof of concept, I analyse communicative representation on immigration in the United Kingdom between 2015 and 2019, using legislative speech as a proxy for elite speech.

- Split into: May 2015-June 2016, June 2016-June 2017, June 2017-July 2019.

Measuring communicative representation

The concept of communicative representation has two components:

- **Communicative congruence:** proportionate representation of the entire distribution of voter preferences, across issues, in elite speech.
- **Communicative responsiveness:** changes in the distribution of preferences produce corresponding changes in preferences expressed in elite speech.

As proof of concept, I analyse communicative representation on immigration in the United Kingdom between 2015 and 2019, using legislative speech as a proxy for elite speech.

- Split into: May 2015-June 2016, June 2016-June 2017, June 2017-July 2019.

Empirical strategy

Measuring communicative representation on an issue requires:

1. Identifying the distribution of voter preferences on that issue.

2. Identifying elite speech on that issue in a particular domain (e.g. legislative speech).

3. Estimating elite preferences on that issue using text scaling methods.

4. Rescaling elite and voter preferences so that they are on the same, comparable, scale.

5. Comparing the distribution of voter preferences and of elite *speech* on the issue.

6. Measuring the correlation between changes in the distribution of voter preferences and of elite speech.

Empirical strategy

Measuring communicative representation on an issue requires:

1. Identifying the distribution of voter preferences on that issue.
2. Identifying elite speech on that issue in a particular domain (e.g. legislative speech).
3. Estimating elite preferences on that issue using text scaling methods.
4. Rescaling elite and voter preferences so that they are on the same, comparable, scale.
5. Comparing the distribution of voter preferences and of elite *speech* on the issue.
6. Measuring the correlation between changes in the distribution of voter preferences and of elite speech.

Empirical strategy

Measuring communicative representation on an issue requires:

1. Identifying the distribution of voter preferences on that issue.
2. Identifying elite speech on that issue in a particular domain (e.g. legislative speech).
3. Estimating elite preferences on that issue using text scaling methods.
4. Rescaling elite and voter preferences so that they are on the same, comparable, scale.
5. Comparing the distribution of voter preferences and of elite *speech* on the issue.
6. Measuring the correlation between changes in the distribution of voter preferences and of elite speech.

Empirical strategy

Measuring communicative representation on an issue requires:

1. Identifying the distribution of voter preferences on that issue.
2. Identifying elite speech on that issue in a particular domain (e.g. legislative speech).
3. Estimating elite preferences on that issue using text scaling methods.
4. Rescaling elite and voter preferences so that they are on the same, comparable, scale.
5. Comparing the distribution of voter preferences and of elite *speech* on the issue.
6. Measuring the correlation between changes in the distribution of voter preferences and of elite speech.

Empirical strategy

Measuring communicative representation on an issue requires:

1. Identifying the distribution of voter preferences on that issue.
2. Identifying elite speech on that issue in a particular domain (e.g. legislative speech).
3. Estimating elite preferences on that issue using text scaling methods.
4. Rescaling elite and voter preferences so that they are on the same, comparable, scale.
5. Comparing the distribution of voter preferences and of elite *speech* on the issue.
6. Measuring the correlation between changes in the distribution of voter preferences and of elite speech.

Empirical strategy

Measuring communicative representation on an issue requires:

1. Identifying the distribution of voter preferences on that issue.
2. Identifying elite speech on that issue in a particular domain (e.g. legislative speech).
3. Estimating elite preferences on that issue using text scaling methods.
4. Rescaling elite and voter preferences so that they are on the same, comparable, scale.
5. Comparing the distribution of voter preferences and of elite *speech* on the issue.
6. Measuring the correlation between changes in the distribution of voter preferences and of elite speech.

Empirical strategy

Measuring communicative representation on an issue requires:

1. Identifying the distribution of voter preferences on that issue.
2. Identifying elite speech on that issue in a particular domain (e.g. legislative speech).
3. Estimating elite preferences on that issue using text scaling methods.
4. Rescaling elite and voter preferences so that they are on the same, comparable, scale.
5. Comparing the distribution of voter preferences and of elite *speech* on the issue.
6. Measuring the correlation between changes in the distribution of voter preferences and of elite speech.

Measuring voter and elite preferences on immigration

- I measure the preferences of British voters on immigration in 2016, 2017 and 2018 using data from the British Election Study (waves 7, 11 and 14).

  - I use the sampling weights provided to ensure the resulting distribution is representative of UK electorate.

- To analyze elite speech, I source the relevant parliamentary speeches from the ParlSpeech database (Rauh and Schwalbach 2020).

  - Identify speeches dealing with *immigration* using the cross-domain supervised topic classifier proposed by Osnabrüegge, Ash and Morelli (2021).

  - Easily scaled to other issues and countries/periods for which labelled manifesto data and digitized speech data is available.

- To further reduce measurement error, restrict attention to references to immigration within these speeches.

- Use Wordscores to scale legislator preferences on immigration based on speech extracts.

Measuring voter and elite preferences on immigration

- I measure the preferences of British voters on immigration in 2016, 2017 and 2018 using data from the British Election Study (waves 7, 11 and 14).

    - I use the sampling weights provided to ensure the resulting distribution is representative of UK electorate.

- To analyze elite speech, I source the relevant parliamentary speeches from the ParlSpeech database (Rauh and Schwalbach 2020).

    - Identify speeches dealing with *immigration* using the cross-domain supervised topic classifier proposed by Osnabrüegge, Ash and Morelli (2021).

    - Easily scaled to other issues and countries/periods for which labelled manifesto data and digitized speech data is available.

- To further reduce measurement error, restrict attention to references to immigration within these speeches.

- Use Wordscores to scale legislator preferences on immigration based on speech extracts.

Measuring voter and elite preferences on immigration

- I measure the preferences of British voters on immigration in 2016, 2017 and 2018 using data from the British Election Study (waves 7, 11 and 14).

  - I use the sampling weights provided to ensure the resulting distribution is representative of UK electorate.

- To analyze elite speech, I source the relevant parliamentary speeches from the ParlSpeech database (Rauh and Schwalbach 2020).

  - Identify speeches dealing with *immigration* using the cross-domain supervised topic classifier proposed by Osnabrüegge, Ash and Morelli (2021).

  - Easily scaled to other issues and countries/periods for which labelled manifesto data and digitized speech data is available.

- To further reduce measurement error, restrict attention to references to immigration within these speeches.

- Use Wordscores to scale legislator preferences on immigration based on speech extracts.

Measuring voter and elite preferences on immigration

- I measure the preferences of British voters on immigration in 2016, 2017 and 2018 using data from the British Election Study (waves 7, 11 and 14).

  - I use the sampling weights provided to ensure the resulting distribution is representative of UK electorate.

- To analyze elite speech, I source the relevant parliamentary speeches from the ParlSpeech database (Rauh and Schwalbach 2020).

  - Identify speeches dealing with *immigration* using the cross-domain supervised topic classifier proposed by Osnabrüegge, Ash and Morelli (2021).

  - Easily scaled to other issues and countries/periods for which labelled manifesto data and digitized speech data is available.

- To further reduce measurement error, restrict attention to references to immigration within these speeches.

- Use Wordscores to scale legislator preferences on immigration based on speech extracts.

Measuring voter and elite preferences on immigration

- I measure the preferences of British voters on immigration in 2016, 2017 and 2018 using data from the British Election Study (waves 7, 11 and 14).

  - I use the sampling weights provided to ensure the resulting distribution is representative of UK electorate.

- To analyze elite speech, I source the relevant parliamentary speeches from the ParlSpeech database (Rauh and Schwalbach 2020).

  - Identify speeches dealing with *immigration* using the cross-domain supervised topic classifier proposed by Osnabrüegge, Ash and Morelli (2021).

  - Easily scaled to other issues and countries/periods for which labelled manifesto data and digitized speech data is available.

- To further reduce measurement error, restrict attention to references to immigration within these speeches.

- Use Wordscores to scale legislator preferences on immigration based on speech extracts.

Measuring voter and elite preferences on immigration

- I measure the preferences of British voters on immigration in 2016, 2017 and 2018 using data from the British Election Study (waves 7, 11 and 14).

  - I use the sampling weights provided to ensure the resulting distribution is representative of UK electorate.

- To analyze elite speech, I source the relevant parliamentary speeches from the ParlSpeech database (Rauh and Schwalbach 2020).

  - Identify speeches dealing with *immigration* using the cross-domain supervised topic classifier proposed by Osnabrüegge, Ash and Morelli (2021).

  - Easily scaled to other issues and countries/periods for which labelled manifesto data and digitized speech data is available.

- To further reduce measurement error, restrict attention to references to immigration within these speeches.

- Use Wordscores to scale legislator preferences on immigration based on speech extracts.

Measuring voter and elite preferences on immigration

- I measure the preferences of British voters on immigration in 2016, 2017 and 2018 using data from the British Election Study (waves 7, 11 and 14).
  - I use the sampling weights provided to ensure the resulting distribution is representative of UK electorate.

- To analyze elite speech, I source the relevant parliamentary speeches from the ParlSpeech database (Rauh and Schwalbach 2020).
  - Identify speeches dealing with *immigration* using the cross-domain supervised topic classifier proposed by Osnabrüegge, Ash and Morelli (2021).
  - Easily scaled to other issues and countries/periods for which labelled manifesto data and digitized speech data is available.

- To further reduce measurement error, restrict attention to references to immigration within these speeches.

- Use Wordscores to scale legislator preferences on immigration based on speech extracts.

Comparing voter preferences and elite speech on immigration

- In order to compare the distributions of legislator and voter preferences on immigration, we need to place them on a common scale.

- Proposed solution: use some legislators as anchors, placing them on the BES scale by using (attentive) respondents' placements of their parties.

    - Anchor legislators: for Labour and the Conservatives, serving immigration and home affairs spokespeople; for the Liberal Democrats, the party leader.

- As in Basu et al. (2022), I use a generalized additive model to generate a mapping between the two scales.

    - Extremes are bridged using left and right-most legislators based on Wordscores.

- Finally, to assess how voters are represented in *speech*, I compare the distribution of voter **preferences** on immigration to the distribution of legislative **discourse**.

- In order to compare the distributions of legislator and voter preferences on immigration, we need to place them on a common scale.

- Proposed solution: use some legislators as anchors, placing them on the BES scale by using (attentive) respondents' placements of their parties.

  - Anchor legislators: for Labour and the Conservatives, serving immigration and home affairs spokespeople; for the Liberal Democrats, the party leader.

- As in Basu et al. (2022), I use a generalized additive model to generate a mapping between the two scales.

  - Extremes are bridged using left and right-most legislators based on Wordscores.

- Finally, to assess how voters are represented in *speech*, I compare the distribution of voter **preferences** on immigration to the distribution of legislative **discourse**.

- In order to compare the distributions of legislator and voter preferences on immigration, we need to place them on a common scale.

- Proposed solution: use some legislators as anchors, placing them on the BES scale by using (attentive) respondents' placements of their parties.

  - Anchor legislators: for Labour and the Conservatives, serving immigration and home affairs spokespeople; for the Liberal Democrats, the party leader.

- As in Basu et al. (2022), I use a generalized additive model to generate a mapping between the two scales.

  - Extremes are bridged using left and right-most legislators based on Wordscores.

- Finally, to assess how voters are represented in *speech*, I compare the distribution of voter **preferences** on immigration to the distribution of legislative **discourse**.

- In order to compare the distributions of legislator and voter preferences on immigration, we need to place them on a common scale.

- Proposed solution: use some legislators as anchors, placing them on the BES scale by using (attentive) respondents' placements of their parties.

  - Anchor legislators: for Labour and the Conservatives, serving immigration and home affairs spokespeople; for the Liberal Democrats, the party leader.

- As in Basu et al. (2022), I use a generalized additive model to generate a mapping between the two scales.

  - Extremes are bridged using left and right-most legislators based on Wordscores.

- Finally, to assess how voters are represented in *speech*, I compare the distribution of voter **preferences** on immigration to the distribution of legislative **discourse**.

- In order to compare the distributions of legislator and voter preferences on immigration, we need to place them on a common scale.

- Proposed solution: use some legislators as anchors, placing them on the BES scale by using (attentive) respondents' placements of their parties.

  - Anchor legislators: for Labour and the Conservatives, serving immigration and home affairs spokespeople; for the Liberal Democrats, the party leader.

- As in Basu et al. (2022), I use a generalized additive model to generate a mapping between the two scales.

  - Extremes are bridged using left and right-most legislators based on Wordscores.

- Finally, to assess how voters are represented in *speech*, I compare the distribution of voter **preferences** on immigration to the distribution of legislative **discourse**.

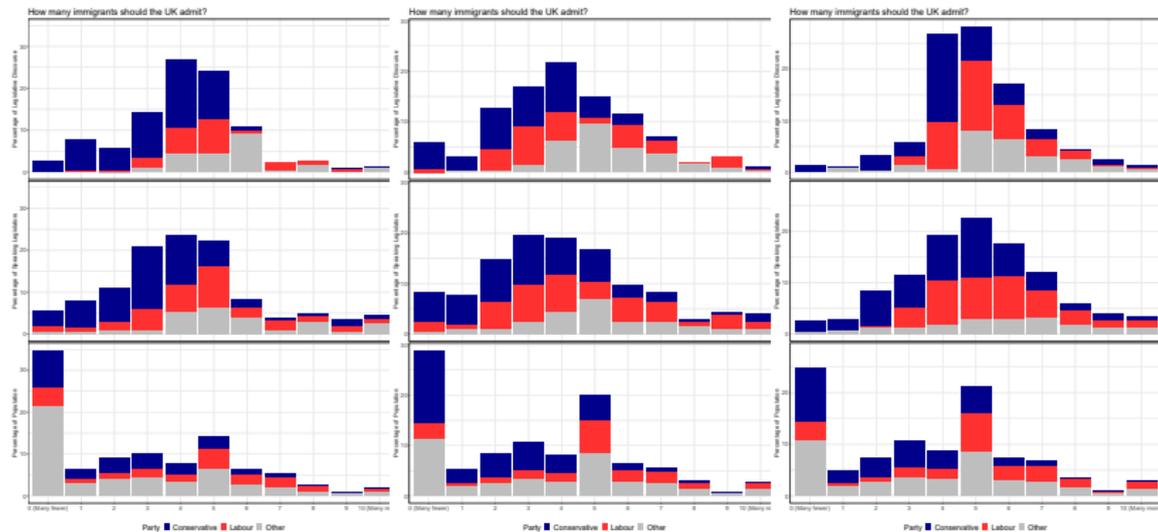# Comparing voter preferences and elite speech on immigration

- In order to compare the distributions of legislator and voter preferences on immigration, we need to place them on a common scale.

- Proposed solution: use some legislators as anchors, placing them on the BES scale by using (attentive) respondents' placements of their parties.

    - Anchor legislators: for Labour and the Conservatives, serving immigration and home affairs spokespeople; for the Liberal Democrats, the party leader.

- As in Basu et al. (2022), I use a generalized additive model to generate a mapping between the two scales.

    - Extremes are bridged using left and right-most legislators based on Wordscores.

- Finally, to assess how voters are represented in *speech*, I compare the distribution of voter **preferences** on immigration to the distribution of legislative **discourse**.

Additional results

1. **Communicative malapportionment**, based on measure of electoral malapportionment proposed by Samuels and Snyder (2001).

   - Intuition: prop. of speech delivered by legislators who would not have received floor time under perfect comm. congruence.
   - Estimated to be about 30-40% in this period.

2. **Relative communicative power**, or how much different opinions and social groups are over-represented in elite (legislative) speech relative to their numbers in the electorate.

   - Liberal and centrist opinions on immigration are consistently more over-represented than conservative ones.
     - Ages 18-29 most over-represented.
     - Leavers, non-graduates, the white working class under-represented.

3. **Communicative responsiveness**, or correlation between lagged Δ prop. of voters (t-2,t-1) and Δ prop. legislative speech (t-1,t) in each response category.

   - Perfect responsiveness $\implies \rho = 1$.
   - Estimated as 0.61 given data coded so far (p<0.05).

Additional results

1. **Communicative malapportionment**, based on measure of electoral malapportionment proposed by Samuels and Snyder (2001).
   - Intuition: prop. of speech delivered by legislators who would not have received floor time under perfect comm. congruence.
   - Estimated to be about 30-40% in this period.

2. **Relative communicative power**, or how much different opinions and social groups are over-represented in elite (legislative) speech relative to their numbers in the electorate.
   - Liberal and centrist opinions on immigration are consistently more over-represented than conservative ones.
     - Ages 18-29 most over-represented.
     - Leavers, non-graduates, the white working class under-represented.

3. **Communicative responsiveness**, or correlation between lagged Δ prop. of voters (t-2,t-1) and Δ prop. legislative speech (t-1,t) in each response category.
   - Perfect responsiveness $\implies \rho = 1$.
   - Estimated as 0.61 given data coded so far (p<0.05).

1. **Communicative malapportionment**, based on measure of electoral malapportionment proposed by Samuels and Snyder (2001).
   - Intuition: prop. of speech delivered by legislators who would not have received floor time under perfect comm. congruence.
   - Estimated to be about 30-40% in this period.

2. **Relative communicative power**, or how much different opinions and social groups are over-represented in elite (legislative) speech relative to their numbers in the electorate.
   - Liberal and centrist opinions on immigration are consistently more over-represented than conservative ones.
      - Ages 18-29 most over-represented.
      - Leavers, non-graduates, the white working class under-represented.

3. **Communicative responsiveness**, or correlation between lagged Δ prop. of voters (t-2,t-1) and Δ prop. legislative speech (t-1,t) in each response category.
   - Perfect responsiveness $\implies \rho = 1$.
   - Estimated as 0.61 given data coded so far (p<0.05).

1. **Communicative malapportionment**, based on measure of electoral malapportionment proposed by Samuels and Snyder (2001).
   - Intuition: prop. of speech delivered by legislators who would not have received floor time under perfect comm. congruence.
   - Estimated to be about 30-40% in this period.

2. Relative communicative power, or how much different opinions and social groups are over-represented in elite (legislative) speech relative to their numbers in the electorate.
   - Liberal and centrist opinions on immigration are consistently more over-represented than conservative ones.
     - Ages 18-29 most over-represented.
     - Leavers, non-graduates, the white working class under-represented.

3. Communicative responsiveness, or correlation between lagged Δ prop. of voters (t-2,t-1) and Δ prop. legislative speech (t-1,t) in each response category.
   - Perfect responsiveness $\implies \rho = 1$.
   - Estimated as 0.61 given data coded so far (p<0.05).

Additional results

1. **Communicative malapportionment**, based on measure of electoral malapportionment proposed by Samuels and Snyder (2001).
   - Intuition: prop. of speech delivered by legislators who would not have received floor time under perfect comm. congruence.
   - Estimated to be about 30-40% in this period.

2. **Relative communicative power**, or how much different opinions and social groups are over-represented in elite (legislative) speech relative to their numbers in the electorate.
   - Liberal and centrist opinions on immigration are consistently more over-represented than conservative ones.
     - Ages 18-29 most over-represented.
     - Leavers, non-graduates, the white working class under-represented.

3. **Communicative responsiveness**, or correlation between lagged Δ prop. of voters (t-2,t-1) and Δ prop. legislative speech (t-1,t) in each response category.
   - Perfect responsiveness $\implies \rho = 1$.
   - Estimated as 0.61 given data coded so far (p<0.05).

Additional results

1. **Communicative malapportionment**, based on measure of electoral malapportionment proposed by Samuels and Snyder (2001).
   - Intuition: prop. of speech delivered by legislators who would not have received floor time under perfect comm. congruence.
   - Estimated to be about 30-40% in this period.

2. **Relative communicative power**, or how much different opinions and social groups are over-represented in elite (legislative) speech relative to their numbers in the electorate.
   - Liberal and centrist opinions on immigration are consistently more over-represented than conservative ones.
     - Ages 18-29 most over-represented.
     - Leavers, non-graduates, the white working class under-represented.

3. **Communicative responsiveness**, or correlation between lagged Δ prop. of voters (t-2,t-1) and Δ prop. legislative speech (t-1,t) in each response category.
   - Perfect responsiveness $\implies \rho = 1$.
   - Estimated as 0.61 given data coded so far (p<0.05).

Additional results

1. **Communicative malapportionment**, based on measure of electoral malapportionment proposed by Samuels and Snyder (2001).
   - Intuition: prop. of speech delivered by legislators who would not have received floor time under perfect comm. congruence.
   - Estimated to be about 30-40% in this period.

2. **Relative communicative power**, or how much different opinions and social groups are over-represented in elite (legislative) speech relative to their numbers in the electorate.
   - Liberal and centrist opinions on immigration are consistently more over-represented than conservative ones.
     - Ages 18-29 most over-represented.
     - Leavers, non-graduates, the white working class under-represented.

3. **Communicative responsiveness**, or correlation between lagged Δ prop. of voters (t-2,t-1) and Δ prop. legislative speech (t-1,t) in each response category.
   - Perfect responsiveness $\implies \rho = 1$.
   - Estimated as 0.61 given data coded so far (p<0.05).

Additional results

1. **Communicative malapportionment**, based on measure of electoral malapportionment proposed by Samuels and Snyder (2001).
   - Intuition: prop. of speech delivered by legislators who would not have received floor time under perfect comm. congruence.
   - Estimated to be about 30-40% in this period.

2. **Relative communicative power**, or how much different opinions and social groups are over-represented in elite (legislative) speech relative to their numbers in the electorate.
   - Liberal and centrist opinions on immigration are consistently more over-represented than conservative ones.
     - Ages 18-29 most over-represented.
     - Leavers, non-graduates, the white working class under-represented.

3. Communicative responsiveness, or correlation between lagged Δ prop. of voters (t-2,t-1) and Δ prop. legislative speech (t-1,t) in each response category.
   - Perfect responsiveness $\implies \rho = 1$.
   - Estimated as 0.61 given data coded so far (p<0.05).

Additional results

1. **Communicative malapportionment**, based on measure of electoral malapportionment proposed by Samuels and Snyder (2001).
   - Intuition: prop. of speech delivered by legislators who would not have received floor time under perfect comm. congruence.
   - Estimated to be about 30-40% in this period.

2. **Relative communicative power**, or how much different opinions and social groups are over-represented in elite (legislative) speech relative to their numbers in the electorate.
   - Liberal and centrist opinions on immigration are consistently more over-represented than conservative ones.
     - Ages 18-29 most over-represented.
     - Leavers, non-graduates, the white working class under-represented.

3. **Communicative responsiveness**, or correlation between lagged Δ prop. of voters (t-2,t-1) and Δ prop. legislative speech (t-1,t) in each response category.
   - Perfect responsiveness $\implies \rho = 1$.
   - Estimated as 0.61 given data coded so far (p<0.05).

Additional results

1. **Communicative malapportionment**, based on measure of electoral malapportionment proposed by Samuels and Snyder (2001).
   - Intuition: prop. of speech delivered by legislators who would not have received floor time under perfect comm. congruence.
   - Estimated to be about 30-40% in this period.

2. **Relative communicative power**, or how much different opinions and social groups are over-represented in elite (legislative) speech relative to their numbers in the electorate.
   - Liberal and centrist opinions on immigration are consistently more over-represented than conservative ones.
     - Ages 18-29 most over-represented.
     - Leavers, non-graduates, the white working class under-represented.

3. **Communicative responsiveness**, or correlation between lagged Δ prop. of voters (t-2,t-1) and Δ prop. legislative speech (t-1,t) in each response category.
   - Perfect responsiveness $\implies \rho = 1$.
   - Estimated as 0.61 given data coded so far (p<0.05).

Additional results

1. **Communicative malapportionment**, based on measure of electoral malapportionment proposed by Samuels and Snyder (2001).
   - Intuition: prop. of speech delivered by legislators who would not have received floor time under perfect comm. congruence.
   - Estimated to be about 30-40% in this period.

2. **Relative communicative power**, or how much different opinions and social groups are over-represented in elite (legislative) speech relative to their numbers in the electorate.
   - Liberal and centrist opinions on immigration are consistently more over-represented than conservative ones.
     - Ages 18-29 most over-represented.
     - Leavers, non-graduates, the white working class under-represented.

3. **Communicative responsiveness**, or correlation between lagged Δ prop. of voters (t-2,t-1) and Δ prop. legislative speech (t-1,t) in each response category.
   - Perfect responsiveness $\implies \rho = 1$.
   - Estimated as 0.61 given data coded so far (p<0.05).

Conclusion

- How do we measure the distribution of effective political power if voter preferences are partly endogenous to elite communication?

- Propose a new measurable criterion for political inequality in such contexts: communicative representation, or how much speech by elite actors reflects, as well as responds to, the entire distribution of voter preferences.

- As proof of concept, I analyze communicative representation on immigration in the UK, 2015–2019.
  - Initial results suggest there is both communicative malapportionment *and* communicative responsiveness on this issue.

- Suggested approach is scaleable to a large number of countries, issues and time periods; potentially other domains of elite communication.

- A criterion to compare actually existing democracies against a benchmark of 'perfect' communicative representation, and explore correlates of variation in performance.
  - I hope, joins efforts to place democratic theory on more 'realistic' foundations (e.g. Knight and Johnson 2011, Bagg 2018).

Conclusion

- How do we measure the distribution of effective political power if voter preferences are partly endogenous to elite communication?

- Propose a new measurable criterion for political inequality in such contexts: communicative representation, or how much speech by elite actors reflects, as well as responds to, the entire distribution of voter preferences.

- As proof of concept, I analyze communicative representation on immigration in the UK, 2015–2019.
  - Initial results suggest there is both communicative malapportionment *and* communicative responsiveness on this issue.

- Suggested approach is scaleable to a large number of countries, issues and time periods; potentially other domains of elite communication.

- A criterion to compare actually existing democracies against a benchmark of 'perfect' communicative representation, and explore correlates of variation in performance.
  - I hope, joins efforts to place democratic theory on more 'realistic' foundations (e.g. Knight and Johnson 2011, Bagg 2018).

Conclusion

- How do we measure the distribution of effective political power if voter preferences are partly endogenous to elite communication?

- Propose a new measurable criterion for political inequality in such contexts: communicative representation, or how much speech by elite actors reflects, as well as responds to, the entire distribution of voter preferences.

- As proof of concept, I analyze communicative representation on immigration in the UK, 2015–2019.
    - Initial results suggest there is both communicative malapportionment *and* communicative responsiveness on this issue.

- Suggested approach is scaleable to a large number of countries, issues and time periods; potentially other domains of elite communication.

- A criterion to compare actually existing democracies against a benchmark of 'perfect' communicative representation, and explore correlates of variation in performance.
    - I hope, joins efforts to place democratic theory on more 'realistic' foundations (e.g. Knight and Johnson 2011, Bagg 2018).

Conclusion

- How do we measure the distribution of effective political power if voter preferences are partly endogenous to elite communication?

- Propose a new measurable criterion for political inequality in such contexts: communicative representation, or how much speech by elite actors reflects, as well as responds to, the entire distribution of voter preferences.

- As proof of concept, I analyze communicative representation on immigration in the UK, 2015–2019.
  - Initial results suggest there is both communicative malapportionment *and* communicative responsiveness on this issue.

- Suggested approach is scaleable to a large number of countries, issues and time periods; potentially other domains of elite communication.

- A criterion to compare actually existing democracies against a benchmark of 'perfect' communicative representation, and explore correlates of variation in performance.
  - I hope, joins efforts to place democratic theory on more 'realistic' foundations (e.g. Knight and Johnson 2011, Bagg 2018).

Conclusion

- How do we measure the distribution of effective political power if voter preferences are partly endogenous to elite communication?

- Propose a new measurable criterion for political inequality in such contexts: **communicative representation**, or how much speech by elite actors reflects, as well as responds to, the entire distribution of voter preferences.

- As proof of concept, I analyze communicative representation on immigration in the UK, 2015–2019.
  - Initial results suggest there is both communicative malapportionment *and* communicative responsiveness on this issue.

- Suggested approach is scaleable to a large number of countries, issues and time periods; potentially other domains of elite communication.

- A criterion to compare actually existing democracies against a benchmark of 'perfect' communicative representation, and explore correlates of variation in performance.
  - I hope, joins efforts to place democratic theory on more 'realistic' foundations (e.g. Knight and Johnson 2011, Bagg 2018).

Conclusion

- How do we measure the distribution of effective political power if voter preferences are partly endogenous to elite communication?

- Propose a new measurable criterion for political inequality in such contexts: **communicative representation**, or how much speech by elite actors reflects, as well as responds to, the entire distribution of voter preferences.

- As proof of concept, I analyze communicative representation on immigration in the UK, 2015–2019.
    - Initial results suggest there is both communicative malapportionment *and* communicative responsiveness on this issue.

- Suggested approach is scaleable to a large number of countries, issues and time periods; potentially other domains of elite communication.

- A criterion to compare actually existing democracies against a benchmark of 'perfect' communicative representation, and explore correlates of variation in performance.
    - I hope, joins efforts to place democratic theory on more 'realistic' foundations (e.g. Knight and Johnson 2011, Bagg 2018).

Conclusion

- How do we measure the distribution of effective political power if voter preferences are partly endogenous to elite communication?

- Propose a new measurable criterion for political inequality in such contexts: communicative representation, or how much speech by elite actors reflects, as well as responds to, the entire distribution of voter preferences.

- As proof of concept, I analyze communicative representation on immigration in the UK, 2015–2019.
    - Initial results suggest there is both communicative malapportionment *and* communicative responsiveness on this issue.

- Suggested approach is scaleable to a large number of countries, issues and time periods; potentially other domains of elite communication.

- A criterion to compare actually existing democracies against a benchmark of 'perfect' communicative representation, and explore correlates of variation in performance.
    - I hope, joins efforts to place democratic theory on more 'realistic' foundations (e.g. Knight and Johnson 2011, Bagg 2018).

Preference Endogeneity and the Problem of Electoral Representation

Chitralekha Basu

July 5, 2022

Cologne Center for Comparative Politics
University of Cologne

Empirical Strategy

Empirical strategy

Measuring communicative representation on an issue requires:

1. Identifying the distribution of voter preferences on that issue.
2. Identifying elite speech on that issue in a particular domain (e.g. legislative speech).
3. Estimating elite preferences on that issue using text scaling methods.
4. Rescaling elite and voter preferences so that they are on the same, comparable, scale.
5. Comparing the distribution of voter preferences and of elite *speech* on the issue.
6. Measuring the correlation between changes in the distribution of voter preferences and of elite speech.

I measure the preferences of British voters on immigration in 2016 and 2018 using data from the British Election Study (waves 7, 11 and 14).

- Question wording: Some people think that the UK should allow *many more* immigrants to come to the UK to live and others think that the UK should allow *many fewer* immigrants. Where would you place yourself and the parties on this (11 point) scale?

- Question asked (at least) annually from 2016 onwards.

I use the sampling weights provided to ensure the resulting distribution is representative of UK electorate.

Advantage: we can use voter placements of the parties to anchor legislators and voters on the same scale.

I measure the preferences of British voters on immigration in 2016 and 2018 using data from the British Election Study (waves 7, 11 and 14).

- Question wording: Some people think that the UK should allow *many more* immigrants to come to the UK to live and others think that the UK should allow *many fewer* immigrants. Where would you place yourself and the parties on this (11 point) scale?
- Question asked (at least) annually from 2016 onwards.

I use the sampling weights provided to ensure the resulting distribution is representative of UK electorate.

Advantage: we can use voter placements of the parties to anchor legislators and voters on the same scale.

Measuring voter preferences on immigration

I measure the preferences of British voters on immigration in 2016 and 2018 using data from the British Election Study (waves 7, 11 and 14).

- <u>Question wording:</u> Some people think that the UK should allow *many more* immigrants to come to the UK to live and others think that the UK should allow *many fewer* immigrants. Where would you place yourself and the parties on this (11 point) scale?
- Question asked (at least) annually from 2016 onwards.

I use the sampling weights provided to ensure the resulting distribution is representative of UK electorate.

Advantage: we can use voter placements of the parties to anchor legislators and voters on the same scale.

I measure the preferences of British voters on immigration in 2016 and 2018 using data from the British Election Study (waves 7, 11 and 14).

- Question wording: Some people think that the UK should allow *many more* immigrants to come to the UK to live and others think that the UK should allow *many fewer* immigrants. Where would you place yourself and the parties on this (11 point) scale?
- Question asked (at least) annually from 2016 onwards.

I use the sampling weights provided to ensure the resulting distribution is representative of UK electorate.

Advantage: we can use voter placements of the parties to anchor legislators and voters on the same scale.

Measuring voter preferences on immigration

I measure the preferences of British voters on immigration in 2016 and 2018 using data from the British Election Study (waves 7, 11 and 14).

- Question wording: Some people think that the UK should allow *many more* immigrants to come to the UK to live and others think that the UK should allow *many fewer* immigrants. Where would you place yourself and the parties on this (11 point) scale?
- Question asked (at least) annually from 2016 onwards.

I use the sampling weights provided to ensure the resulting distribution is representative of UK electorate.

Advantage: we can use voter placements of the parties to anchor legislators and voters on the same scale.

Identifying legislative speech on immigration

I source parliamentary speeches from the UK House of Commons that were delivered within the relevant timeframe from the ParlSpeech database (Rauh and Schwalbach 2020).

- To identify speeches dealing specifically with *immigration*, I apply the cross-domain supervised topic classifier proposed by Osnabrüegge, Ash and Morelli (2021).
    - Classifier trained using labelled UK party platforms coded by the Comparative Agendas Project (Froio, Bevan and Jennings 2017).
    - Applied to target corpus of parliamentary speeches delivered between May 2015 and July 2019.
- Lower data collection costs than a dictionary-based approach or within-domain topic classifiers (e.g. Barberá et al. 2021).
- Less subjective than a unsupervised topic classification approach like structural topic modeling (Roberts et al. 2013).
- Easily scaled to other issues and countries/periods for which labelled manifesto data and digitized speech data is available.

Identifying legislative speech on immigration

I source parliamentary speeches from the UK House of Commons that were delivered within the relevant timeframe from the ParlSpeech database (Rauh and Schwalbach 2020).

- To identify speeches dealing specifically with *immigration*, I apply the cross-domain supervised topic classifier proposed by Osnabrüegge, Ash and Morelli (2021).
  - Classifier trained using labelled UK party platforms coded by the Comparative Agendas Project (Froio, Bevan and Jennings 2017).
  - Applied to target corpus of parliamentary speeches delivered between May 2015 and July 2019.
- Lower data collection costs than a dictionary-based approach or within-domain topic classifiers (e.g. Barberá et al. 2021).
- Less subjective than a unsupervised topic classification approach like structural topic modeling (Roberts et al. 2013).
- Easily scaled to other issues and countries/periods for which labelled manifesto data and digitized speech data is available.

Identifying legislative speech on immigration

I source parliamentary speeches from the UK House of Commons that were delivered within the relevant timeframe from the ParlSpeech database (Rauh and Schwalbach 2020).

- To identify speeches dealing specifically with *immigration*, I apply the cross-domain supervised topic classifier proposed by Osnabrüegge, Ash and Morelli (2021).
    - Classifier trained using labelled UK party platforms coded by the Comparative Agendas Project (Froio, Bevan and Jennings 2017).
    - Applied to target corpus of parliamentary speeches delivered between May 2015 and July 2019.
- Lower data collection costs than a dictionary-based approach or within-domain topic classifiers (e.g. Barberá et al. 2021).
- Less subjective than a unsupervised topic classification approach like structural topic modeling (Roberts et al. 2013).
- Easily scaled to other issues and countries/periods for which labelled manifesto data and digitized speech data is available.

Identifying legislative speech on immigration

I source parliamentary speeches from the UK House of Commons that were delivered within the relevant timeframe from the ParlSpeech database (Rauh and Schwalbach 2020).

- To identify speeches dealing specifically with *immigration*, I apply the cross-domain supervised topic classifier proposed by Osnabrüegge, Ash and Morelli (2021).
    - Classifier trained using labelled UK party platforms coded by the Comparative Agendas Project (Froio, Bevan and Jennings 2017).
    - Applied to target corpus of parliamentary speeches delivered between May 2015 and July 2019.

- Lower data collection costs than a dictionary-based approach or within-domain topic classifiers (e.g. Barberá et al. 2021).

- Less subjective than a unsupervised topic classification approach like structural topic modeling (Roberts et al. 2013).

- Easily scaled to other issues and countries/periods for which labelled manifesto data and digitized speech data is available.

Identifying legislative speech on immigration

I source parliamentary speeches from the UK House of Commons that were delivered within the relevant timeframe from the ParlSpeech database (Rauh and Schwalbach 2020).

- To identify speeches dealing specifically with *immigration*, I apply the cross-domain supervised topic classifier proposed by Osnabrüegge, Ash and Morelli (2021).
    - Classifier trained using labelled UK party platforms coded by the Comparative Agendas Project (Froio, Bevan and Jennings 2017).
    - Applied to target corpus of parliamentary speeches delivered between May 2015 and July 2019.
- Lower data collection costs than a dictionary-based approach or within-domain topic classifiers (e.g. Barberá et al. 2021).
- Less subjective than a unsupervised topic classification approach like structural topic modeling (Roberts et al. 2013).
- Easily scaled to other issues and countries/periods for which labelled manifesto data and digitized speech data is available.

Identifying legislative speech on immigration

I source parliamentary speeches from the UK House of Commons that were delivered within the relevant timeframe from the ParlSpeech database (Rauh and Schwalbach 2020).

- To identify speeches dealing specifically with *immigration*, I apply the cross-domain supervised topic classifier proposed by Osnabrüegge, Ash and Morelli (2021).
    - Classifier trained using labelled UK party platforms coded by the Comparative Agendas Project (Froio, Bevan and Jennings 2017).
    - Applied to target corpus of parliamentary speeches delivered between May 2015 and July 2019.
- Lower data collection costs than a dictionary-based approach or within-domain topic classifiers (e.g. Barberá et al. 2021).
- Less subjective than a unsupervised topic classification approach like structural topic modeling (Roberts et al. 2013).
- Easily scaled to other issues and countries/periods for which labelled manifesto data and digitized speech data is available.

Identifying legislative speech on immigration

I source parliamentary speeches from the UK House of Commons that were delivered within the relevant timeframe from the ParlSpeech database (Rauh and Schwalbach 2020).

- To identify speeches dealing specifically with *immigration*, I apply the cross-domain supervised topic classifier proposed by Osnabrüegge, Ash and Morelli (2021).
    - Classifier trained using labelled UK party platforms coded by the Comparative Agendas Project (Froio, Bevan and Jennings 2017).
    - Applied to target corpus of parliamentary speeches delivered between May 2015 and July 2019.
- Lower data collection costs than a dictionary-based approach or within-domain topic classifiers (e.g. Barberá et al. 2021).
- Less subjective than a unsupervised topic classification approach like structural topic modeling (Roberts et al. 2013).
- Easily scaled to other issues and countries/periods for which labelled manifesto data and digitized speech data is available.

Estimating legislator preferences on immigration

- To further reduce measurement error, I restrict attention to references to immigration within speeches.

  - Specifically, I extract ten word windows surrounding mentions of '*migr*', 'asylum', 'visa', 'refugee*' and 'border*'.

- I then follow O'Grady (2019) in using Wordscores to scale legislator preferences on immigration based on their speech.

  - To anchor the left extreme of the scale, I use the legislators who voted against the 2014 Immigration Act (for 2015-16), and those voting for a 'second referendum' in the indicative votes of Apr 2019 (for 2016-19).

  - To anchor the right extreme of the scale, I use legislators who supported a 2014 amendment to the aforementioned Immigration Act to extend transitional controls on Romanian and Bulgarian migrants to Britain (i.e. right-wing rebels against the governing Conservatives).

- Unsupervised scaling techniques like Wordfish and Wordshoal produced a less plausible ordering of legislators on the immigration issue.

Estimating legislator preferences on immigration

- To further reduce measurement error, I restrict attention to references to immigration within speeches.

  - Specifically, I extract ten word windows surrounding mentions of '*migr*', 'asylum', 'visa', 'refugee*' and 'border*'.

- I then follow O'Grady (2019) in using Wordscores to scale legislator preferences on immigration based on their speech.

  - To anchor the left extreme of the scale, I use the legislators who voted against the 2014 Immigration Act (for 2015-16), and those voting for a 'second referendum' in the indicative votes of Apr 2019 (for 2016-19).

  - To anchor the right extreme of the scale, I use legislators who supported a 2014 amendment to the aforementioned Immigration Act to extend transitional controls on Romanian and Bulgarian migrants to Britain (i.e. right-wing rebels against the governing Conservatives).

- Unsupervised scaling techniques like Wordfish and Wordshoal produced a less plausible ordering of legislators on the immigration issue.

Estimating legislator preferences on immigration

- To further reduce measurement error, I restrict attention to references to immigration within speeches.

    - Specifically, I extract ten word windows surrounding mentions of '*migr*', 'asylum', 'visa', 'refugee*' and 'border*'.

- I then follow O'Grady (2019) in using Wordscores to scale legislator preferences on immigration based on their speech.

    - To anchor the left extreme of the scale, I use the legislators who voted against the 2014 Immigration Act (for 2015-16), and those voting for a 'second referendum' in the indicative votes of Apr 2019 (for 2016-19).

    - To anchor the right extreme of the scale, I use legislators who supported a 2014 amendment to the aforementioned Immigration Act to extend transitional controls on Romanian and Bulgarian migrants to Britain (i.e. right-wing rebels against the governing Conservatives).

- Unsupervised scaling techniques like Wordfish and Wordshoal produced a less plausible ordering of legislators on the immigration issue.

Estimating legislator preferences on immigration

- To further reduce measurement error, I restrict attention to references to immigration within speeches.

    - Specifically, I extract ten word windows surrounding mentions of '*migr*', 'asylum', 'visa', 'refugee*' and 'border*'.

- I then follow O'Grady (2019) in using Wordscores to scale legislator preferences on immigration based on their speech.

    - To anchor the left extreme of the scale, I use the legislators who voted against the 2014 Immigration Act (for 2015-16), and those voting for a 'second referendum' in the indicative votes of Apr 2019 (for 2016-19).

    - To anchor the right extreme of the scale, I use legislators who supported a 2014 amendment to the aforementioned Immigration Act to extend transitional controls on Romanian and Bulgarian migrants to Britain (i.e. right-wing rebels against the governing Conservatives).

- Unsupervised scaling techniques like Wordfish and Wordshoal produced a less plausible ordering of legislators on the immigration issue.

Estimating legislator preferences on immigration

- To further reduce measurement error, I restrict attention to references to immigration within speeches.

    - Specifically, I extract ten word windows surrounding mentions of '*migr*', 'asylum', 'visa', 'refugee*' and 'border*'.

- I then follow O'Grady (2019) in using Wordscores to scale legislator preferences on immigration based on their speech.

    - To anchor the left extreme of the scale, I use the legislators who voted against the 2014 Immigration Act (for 2015-16), and those voting for a 'second referendum' in the indicative votes of Apr 2019 (for 2016-19).

    - To anchor the right extreme of the scale, I use legislators who supported a 2014 amendment to the aforementioned Immigration Act to extend transitional controls on Romanian and Bulgarian migrants to Britain (i.e. right-wing rebels against the governing Conservatives).

- Unsupervised scaling techniques like Wordfish and Wordshoal produced a less plausible ordering of legislators on the immigration issue.

Estimating legislator preferences on immigration

- To further reduce measurement error, I restrict attention to references to immigration within speeches.

  - Specifically, I extract ten word windows surrounding mentions of '*migr*', 'asylum', 'visa', 'refugee*' and 'border*'.

- I then follow O'Grady (2019) in using Wordscores to scale legislator preferences on immigration based on their speech.

  - To anchor the left extreme of the scale, I use the legislators who voted against the 2014 Immigration Act (for 2015-16), and those voting for a 'second referendum' in the indicative votes of Apr 2019 (for 2016-19).

  - To anchor the right extreme of the scale, I use legislators who supported a 2014 amendment to the aforementioned Immigration Act to extend transitional controls on Romanian and Bulgarian migrants to Britain (i.e. right-wing rebels against the governing Conservatives).

- Unsupervised scaling techniques like Wordfish and Wordshoal produced a less plausible ordering of legislators on the immigration issue.

Placing legislators and voters on a common scale

- In order to compare the distributions of legislator and voter preferences on immigration, we need to place them on a common scale.

- Extant joint scaling methods have used common survey items to connect two disjoint sets of survey or roll call data (e.g. Bakker et al. 2014, Tausanovitch and Warshaw 2013).

  - Not very scalable: can only be implemented where elites have already been surveyed on an issue and using the same question as voters.

- Proposed solution: use a subset of legislators as anchors, placing them on the BES scale by using (attentive) respondents' placements of their parties.

  - Anchor legislators: for Labour and the Conservatives, serving immigration and home affairs spokespeople; for the Liberal Democrats, the party leader.

  - Future versions should adjust for differential item functioning as well (Aldrich and McKelvey 1977).

- As in Basu et al. (2022), I use a generalized additive model to generate a mapping between the two scales.

  - Extremes are bridged using left and right-most legislators based on Wordscores.

- Allows us to use historical survey data to locate legislators and voters on a common scale across many countries and years.

Additional Results

I suggest a measure of 'communicative malapportionment' based on the measure of electoral malapportionment proposed by Samuels and Snyder (2001):

$$M_{jt} = \frac{1}{2} \sum_{i=1}^{N} |l_{ijt} - v_{ijt}|$$

- Intuition: the proportion of legislative speech on an issue that is delivered by legislators who would have not received any floor time if there were perfect communicative congruence on that issue.

  - More interpretable than standard measures of distributional overlap (e.g. the Bhattacharya coefficient).

- Estimated to be 30-40% in this period.

Measuring communicative malapportionment

I suggest a measure of 'communicative malapportionment' based on the measure of electoral malapportionment proposed by Samuels and Snyder (2001):

$$M_{jt} = \frac{1}{2} \sum_{i=1}^{N} |l_{ijt} - v_{ijt}|$$

- Intuition: the proportion of legislative speech on an issue that is delivered by legislators who would have not received any floor time if there were perfect communicative congruence on that issue.

    - More interpretable than standard measures of distributional overlap (e.g. the Bhattacharya coefficient).

- Estimated to be 30-40% in this period.

I suggest a measure of 'communicative malapportionment' based on the measure of electoral malapportionment proposed by Samuels and Snyder (2001):

$$M_{jt} = \frac{1}{2} \sum_{i=1}^{N} |l_{ijt} - v_{ijt}|$$

- Intuition: the proportion of legislative speech on an issue that is delivered by legislators who would have not received any floor time if there were perfect communicative congruence on that issue.

  - More interpretable than standard measures of distributional overlap (e.g. the Bhattacharya coefficient).

- Estimated to be 30-40% in this period.

I suggest a measure of 'communicative malapportionment' based on the measure of electoral malapportionment proposed by Samuels and Snyder (2001):

$$M_{jt} = \frac{1}{2} \sum_{i=1}^{N} |l_{ijt} - v_{ijt}|$$

- Intuition: the proportion of legislative speech on an issue that is delivered by legislators who would have not received any floor time if there were perfect communicative congruence on that issue.
    - More interpretable than standard measures of distributional overlap (e.g. the Bhattacharya coefficient).
- Estimated to be 30-40% in this period.

Measuring relative communicative power

We can also measure the extent to which different opinions and sub-electorates are over-represented in legislative speech relative to their numbers as a whole.

- The relative communicative power p_i of an opinion group i is given by:

$$p_i = \frac{\text{fraction of legislators with opinion } i}{\text{fraction of voters with opinion } i}$$

- The relative communicative power P_j of a social group j is then:

$$P_j = \sum_i p_i \times \text{ fraction of social group } j \text{ with opinion } i$$

- I find that liberal and centrist opinions on immigration are consistently more over-represented than conservative ones.

  - Ages 18-29 most over-represented.
  - Leavers, non-graduates, the white working class under-represented.

Measuring relative communicative power

We can also measure the extent to which different opinions and sub-electorates are over-represented in legislative speech relative to their numbers as a whole.

- The relative communicative power p_i of an opinion group i is given by:

$$p_i = \frac{\text{fraction of legislators with opinion } i}{\text{fraction of voters with opinion } i}$$

- The relative communicative power P_j of a social group j is then:

$$P_j = \sum_i p_i \times \text{ fraction of social group } j \text{ with opinion } i$$

- I find that liberal and centrist opinions on immigration are consistently more over-represented than conservative ones.

  - Ages 18-29 most over-represented.
  - Leavers, non-graduates, the white working class under-represented.

Measuring relative communicative power

We can also measure the extent to which different opinions and sub-electorates are over-represented in legislative speech relative to their numbers as a whole.

- The relative communicative power p_i of an opinion group i is given by:

$$p_i = \frac{\text{fraction of legislators with opinion } i}{\text{fraction of voters with opinion } i}$$

- The relative communicative power P_j of a social group j is then:

$$P_j = \sum_i p_i \times \text{ fraction of social group } j \text{ with opinion } i$$

- I find that liberal and centrist opinions on immigration are consistently more over-represented than conservative ones.

  - Ages 18-29 most over-represented.
  - Leavers, non-graduates, the white working class under-represented.

Measuring relative communicative power

We can also measure the extent to which different opinions and sub-electorates are over-represented in legislative speech relative to their numbers as a whole.

- The relative communicative power p_i of an opinion group i is given by:

$$p_i = \frac{\text{fraction of legislators with opinion } i}{\text{fraction of voters with opinion } i}$$

- The relative communicative power P_j of a social group j is then:

$$P_j = \sum_i p_i \times \text{ fraction of social group } j \text{ with opinion } i$$

- I find that liberal and centrist opinions on immigration are consistently more over-represented than conservative ones.

  - Ages 18-29 most over-represented.
  - Leavers, non-graduates, the white working class under-represented.

Measuring relative communicative power

We can also measure the extent to which different opinions and sub-electorates are over-represented in legislative speech relative to their numbers as a whole.

- The relative communicative power p_i of an opinion group i is given by:

$$p_i = \frac{\text{fraction of legislators with opinion } i}{\text{fraction of voters with opinion } i}$$
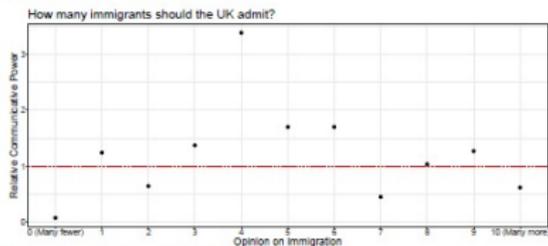
- The relative communicative power P_j of a social group j is then:

$$P_j = \sum_i p_i \times \text{ fraction of social group } j \text{ with opinion } i$$
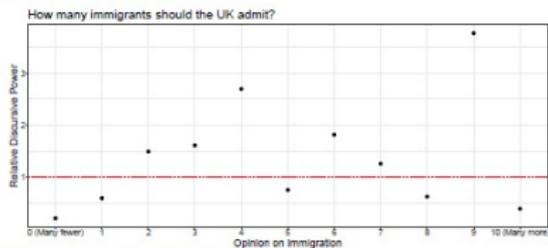
- I find that liberal and centrist opinions on immigration are consistently more over-represented than conservative ones.

    - Ages 18-29 most over-represented.
    - Leavers, non-graduates, the white working class under-represented.

Measuring relative communicative power

We can also measure the extent to which different opinions and sub-electorates are over-represented in legislative speech relative to their numbers as a whole.

- The relative communicative power p_i of an opinion group i is given by:

$$p_i = \frac{\text{fraction of legislators with opinion } i}{\text{fraction of voters with opinion } i}$$

- The relative communicative power P_j of a social group j is then:

$$P_j = \sum_i p_i \times \text{ fraction of social group } j \text{ with opinion } i$$

- I find that liberal and centrist opinions on immigration are consistently more over-represented than conservative ones.
  - Ages 18-29 most over-represented.
  - Leavers, non-graduates, the white working class under-represented.

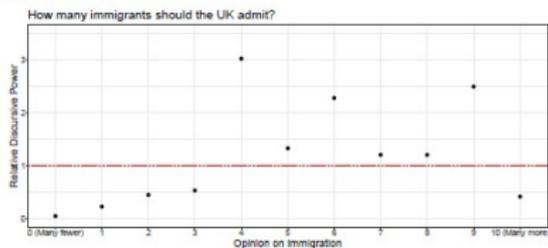# The relative communicative power of opinions on immigration
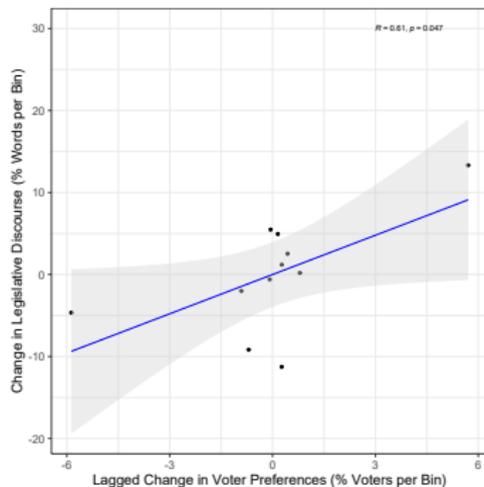

(a) May 2015 – June 2016

How many immigrants should the UK admit?

(b) June 2016 – June 2017

How many immigrants should the UK admit?

(c) June 2017 – July 2019

How many immigrants should the UK admit?

- Correlation between lagged Δ prop. of voters (t-2,t-1) and Δ prop. legislative speech (t-1,t) in each response category is 0.61 (p<0.05).
  - Perfect responsiveness $\implies \rho = 1$.