The Composition of Descriptive Representation

JOHN GERRING   *The University of Texas at Austin, United States*
CONNOR T. JERZAK   *The University of Texas at Austin, United States*
ERZEN ÖNCEL   *Özyeğin University, Turkey*

How well do governments represent the societies they serve? A key aspect of this question concerns the extent to which leaders reflect the demographic features of the population they represent. To address this important issue in a systematic manner, we propose a unified approach for measuring descriptive representation. We apply this approach to newly collected data describing the ethnic, linguistic, religious, and gender identities of over fifty thousand leaders serving in 1,552 political bodies across 156 countries. Strikingly, no country represents social groups in rough proportion to their share of the population. To explain this shortfall, we focus on compositional factors—the size of political bodies as well as the number and relative size of social groups. We investigate these factors using a simple model based on random sampling and the original data described above. Our analyses demonstrate that roughly half of the variability in descriptive representation is attributable to compositional factors.

INTRODUCTION

How well do governments represent the societies they serve? One aspect of this question concerns whether the demographic features of a population are reflected in a country's political leadership. Scholars debate how much impact this "descriptive" aspect of representation has on legitimacy, trust, participation, political recruitment, social conflict, and public policy.[1] We bracket these issues in order to focus on the achievement of descriptive representation.

Extant work suggests a litany of reasons for why social groups might be over- or under-represented in government. Explanations center on constitutional structures (e.g., regime-type and federalism), electoral rules (e.g., officeholding laws, districting, electoral systems, quotas, and reserved seats), informal institutions (e.g., party recruitmentand the availability of candidates), and a host of societal factors (e.g., violence, social exclusion, spatial segregation, poverty, inequality, economic development, and political culture).[2]

Without dismissing these factors, we argue that much of the variability in representation is *compositional*—a byproduct of the size of political bodies and the number and size of groups eligible for representation in those bodies. Equitable representation is fostered by large bodies and homogeneous populations. Failures of representation are likely when political bodies are small or populations are heterogeneous.

We derive theoretical justification for this argument from the analytical expectation of representation under a random-sampling assumption. Empirically, we demonstrate that compositional factors account for roughly half of the variability in descriptive representation across political bodies and across countries (aggregating across bodies). We show that these effects hold across a variety of offices (executives, cabinets, parliamentary party groups, upper and lower chambers of parliament, and supreme courts), across major categories of identity (ethnic, linguistic, religious, gender, and the intersection of ethnicity and gender), and in a wide variety of settings—in rich and poor societies, in democracies and autocracies, in elective and appointive offices, and in all regions of the world.

A number of prior studies bear on various aspects of this question. Goodin (2004) addresses the issue from a theoretical perspective. A suite of studies analyzes the empirical association between local council size and female representation (Bingle 2016; Bullock III and MacManus 1991; Kellogg et al. 2019; Kjaer, Dittmar, and Carroll 2018; Kjaer and Elklit 2014). A few studies focus on the association between the size of the national legislature and female representation (Matland 1998; Oakes and Almquist 1993). One study addresses the

John Gerring (ID), Professor, Department of Government, The University of Texas at Austin, United States, jgerring@austin.utexas.edu.
Connor T. Jerzak (ID), Assistant Professor, Department of Government, The University of Texas at Austin, United States, connor.jerzak@austin.utexas.edu.
Erzen Öncel (ID), Assistant Professor, Department of International Relations, Özyeğin University, Turkey, erzen.oncel@ozyegin.edu.tr.

---

[1] Recent surveys of this vast literature focus on women (Celis and Erzeel 2020; Escobar-Lemmon and Taylor-Robinson 2014; Krook and O'Brien 2012; Paxton, Hughes, and Barnes 2020; Wängnerud 2009), minorities (Bird, Saalfeld, and Wüst 2011; Lublin 2014; Ruedin 2020), or—in a few cases—on multiple domains (Htun 2016; Hughes 2013; Ruedin 2013). Most studies indicate that descriptive representation is consequential (Wängnerud 2009), but see Homola (2019) and qualifications offered by Mackay (2008).

---

[2] See citations in Footnote 1.

size and distribution of ethnic groups and the impact of these factors on representation in national legislatures (Ruedin 2009).

Most of these studies suggest some role for compositional factors; however, results are not especially strong or consistent. This may be a product of limited empirical domains. Studies of local government center on one country (the United States). All studies (local or national) center on legislatures, leaving aside other offices. Most studies focus on women, leaving aside identities that are harder to measure and compare cross-nationally such as those grounded in race, ethnicity, religion, and language.

The present study offers an expansive approach to the study of descriptive representation. Original coding incorporates over fifty thousand political leaders serving in 1,552 political bodies across 156 countries. The identities of these leaders are compared with the population characteristics of each country, gathered from surveys and censuses. The resulting index provides a summary measure of how well each political body achieve its representational function, a score that may be aggregated across political bodies (e.g., executive, legislative, and judicial), across social groups (as defined, e.g., by gender, language, religion, ethnicity), and across countries. We believe that this empirical approach holds promise for unifying a fertile but fragmented field of study.

Section "Coordination Problems" lays out our theoretical framework. Section "Data Collection" introduce our data, section "An Index of Descriptive Representation" propose an all-purpose index for measuring descriptive representation and show how it applies to countries around the world, section "Main Tests" test compositional factors in a multivariate analysis, section "Inside the Box" explore implications of the theory, section "Dimensions of Identity" test compositional effects across various dimensions of identity (including the intersection of gender and ethnicity), section "Contexts" test compositional effects across varying contexts, section "Instrumental Variables" model treatment assignment with instrumental variables, section "Additional Robustness Tests" briefly discuss results from additional robustness tests, and section "Limitations" review empirical limitations of the study. A short conclusion summarizes the findings and reflects upon their implications.

COORDINATION PROBLEMS

Before diving into this much-debated subject, we must define several key terms.

A *political body* refers to any organizational entity that is granted a formal or informal role in governance. (We assume that organizations playing an important political role also play a representational role, regardless of how members are chosen [Rehfeld 2006].) *Leaders*, aka elites or representatives, are those chosen from a population of citizens to serve on these political bodies. *Selectors* appoint or elect leaders. (We leave aside questions about how the selectorate is formed.)

Diversity, or heterogeneity, refers to the number of social groups and their relative size. The more equal in size, the greater the *entropy* across groups. *Groups* are defined by intrinsic traits such as ethnicity, race, caste, religion, language, gender, sexuality, age, disability, occupation, social class, education, or place of origin. Some of these identities are descent-based (Chandra 2006); others are more open-ended, but all are assumed to inform a group's identity.

Descriptive representation refers to a representational capacity deriving from shared identity and shared life experience. It is who political representatives are that matters. Descriptive representation is achieved when the intrinsic traits of persons chosen for a political body mirror those found in the population they are intended to represent (Mansbridge 1999; Pitkin 1967).

Standing in the way of this achievement is a formidable coordination challenge.[3] This arises from the dynamics of identity politics. While identities are fluid in many social contexts, in political contexts, they are usually treated as fixed categories. Political leaders are viewed as one of "us" or one of "them" (Goodyear-Grant and Tolley 2019). Although leaders have a certain amount of wiggle room to frame and reframe their public identity, such subtleties are difficult to communicate and may not be accepted by others. If cues are too subtle, representational bonds may fray. Constituents may not feel represented by someone whose identity is difficult to grasp (Lemi 2021) or whose identity undergoes frequent changes, either of which may call into question vital leadership qualities such as authenticity and credibility.
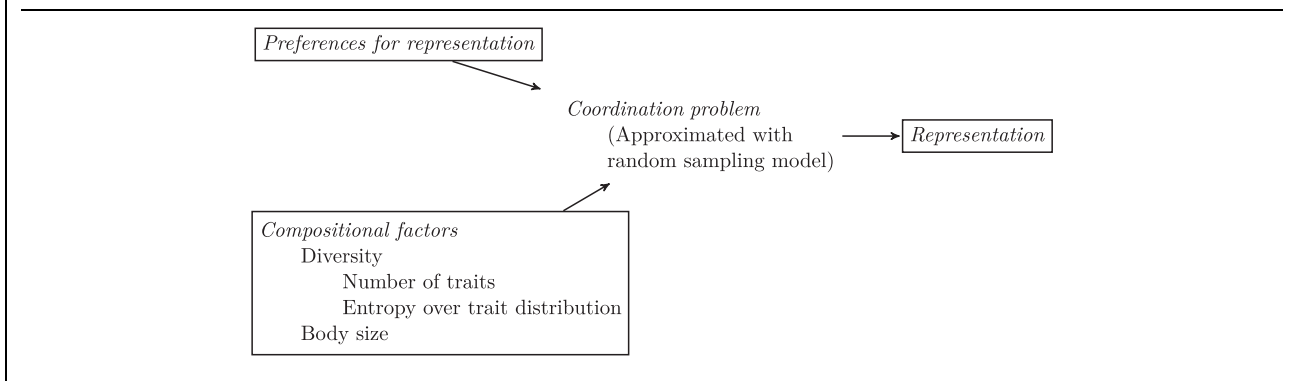
It follows that the task of descriptive representation centers on selection, for once leaders are selected, there is little room for adjustment. A leader's ideology can be dialed up or dialed down, framed in various ways, and adjusted over time in response to events, new information, and changes in public opinion. Not so their identity.

At the point of selection, a two-level coordination challenge arises. On one level are social groups demanding representation; on the other, political bodies providing representation. Fitting these pieces together in the right proportion is complicated, even if there is political will to do so.

Imagine a game of musical chairs with millions of citizens and a few dozen chairs. Seats must be apportioned so that traits among the seated reflect traits in the general population. If there are fewer chairs than traits, the coordination problem is clearly unsolvable. Even if chairs out-number traits, it is a challenging task to align traits in their proper proportion. Note that errors may arise in both directions: a trait may be under- or over-represented.

This fiddly task is rendered more difficult in a typical political setting. Here, we are likely to find multiple bodies with different sizes, constituencies, and rules of

---

[3] *Coordination* is invoked here in a general sense; we do not propose a particular game-theoretic model, though such a model might be constructed.

FIGURE 1. A Compositional Model of Descriptive Representation

Preferences for representation

Coordination problem
(Approximated with
random sampling model)

Representation

Compositional factors
    Diversity
        Number of traits
        Entropy over trait distribution
    Body size

selection—some appointive, some elective, and among the latter, a variety of complex electoral rules. How should selectors coordinate their efforts? Should each advocate for their own group, or should they support other groups—or only certain groups (e.g., minorities)? Note also that group identities are likely to cross-cut, so there are multiple dimensions of descriptive representation to juggle—ethnicity, gender, religion, and so forth (Collins and Bilge 2020). Achieving representation along one dimension may involve sacrifices along another. Finally, descriptive traits are not the only criteria of relevance for a top political job. Selectors must also consider ideology, experience, integrity, charisma, resources, and past performance—any of which may take precedence (Fisher et al. 2018, Part 3).

Coordination challenges faced by descriptive representation are evidently immense. Yet, we argue that much of the action is contained in the core elements of the task: the composition of society and the composition of a political body. As a society becomes more diverse or a political body shrinks, obstacles to representation arise. It's hard to fit multiple categorical traits into a small container in just the right proportions.

In principle, *compositional* effects apply to any political body, any set of descriptive traits, and any political context, whether appointive or elective, democratic or autocratic. They are perhaps most familiar in the context of electoral rules, where larger districts are often thought to allow more space for representation (Lijphart 2004).[4] In this study, we explore a range of political bodies (executives, cabinets, political parties, supreme courts, and legislatures), four main descriptive traits (ethnicity, religion, language, and gender), and a global sample of countries.

Compositional effects are not universal, however. Note that because coordination presents an obstacle to the realization of a given preference set it is necessarily contingent upon that set of preferences. If preferences are weak or a trait is judged to be

inappropriate for public office, that trait is unlikely to be represented. For example, if there is little inclination to represent youth, and age is correlated with other valued traits such as experience, the young are unlikely to be represented (Stockemer and Sundström 2018). In situations like this, preferences override compositional factors, a point discussed at length in Supplementary Materials I.

The resulting framework, summarized in Figure 1, suggests that preferences interact with compositional factors to generate a coordination problem. The resolution of this problem determines the level of descriptive representation achieved for a given trait. We will show that successes and failures follow those encountered when sampling randomly from a population.

A Random Sampling Model

Although the inputs and outputs of the framework illustrated in Figure 1 are transparent and measurable, mechanisms are opaque. We cannot directly observe coordination. Helpfully, we can *approximate* it by examining how representation is affected by compositional changes under certain assumptions.

In our model, intrinsic traits are chosen randomly from a population subject to conditions imposed by compositional factors. This setup allows us to hold background features constant while zeroing in on factors of theoretical interest—the size of the political body and the number and size of groups eligible for representation—which can be manipulated independently. In this fashion, simulations function as ersatz experiments that can be repeated infinitely and fine-tuned to test varying doses and interactions among treatments (De Marchi 2005).

Of course, this depends upon the plausibility of assumptions. Modeling selection as a random draw from a population makes sense insofar as the items under consideration are understood as traits (carried by representatives), not the representatives themselves (who are selected randomly only in the rare case of sortition). The crucial assumption is that the relationship of theoretical interest—between compositional factors and representation—plays out under conditions of random selection much as it does in the real world. If

---

[4] Other researchers have pointed out that by lowering the threshold for smaller parties—which, since they are small, have less room to accommodate diversity—increased district magnitude may have countervailing effects at the party level (Lucardi and Micozzi 2022; Matland and Taylor 1997; Moser 2008).

the apparatus is realistic in this respect, it serves our purpose.

The logic of sampling implies that the larger a sample, and the more homogeneous a population, the more likely a given sample will accurately represent that population. Our model is designed to explore this dynamic by quantifying precisely the ways in which representation responds to exogenous changes in compositional factors.

The outcome of interest, R_b, is an index measuring how closely the distribution of a trait within a political body reflects the distribution of that trait in a population, the core idea behind descriptive representation (Pitkin 1967, Chapter 4). This is similar to the intuition behind measures of vote/seat proportionality such as the Rose Index of Proportionality (Rose 1984). Just as parties should receive representation in accordance with their votes, groups should receive representation in accordance with their numbers. Following this logic, and building on Ruedin (2009), our proposed index of representation takes the following form:

$$R_b = 1 - \frac{1}{2}\sum_{k=1}^{K}|g_{p_k} - G_{b_k}|, \qquad (1)$$

where R_b is the degree of representation present in a particular political body, g_{P_k} is group k's share of the population, G_{B_k} is group k's share of a political body, and K is the total number of groups in the population. The functional form of Equation 1 is chosen so that the representation index is bounded between 0 (no representation) and 1 (perfect alignment between the characteristics of a political body and a population).

The expected value of this representational index is probed as three compositional factors change—body size (n_b), the number of groups in society (K), and the relative size of those groups (g_{p_k}), which we will characterize using entropy. The expected representation under this random sampling process is described in Proposition I (see section S.II.1.12 in Supplementary Materials II for derivation).[5]

Proposition I. Under random sampling of citizens to bodies, expected representation is

$$\mathbb{E}[R_b] = 1 + \frac{1}{2}\sum_{k=1}^{K}\Bigg\{\left(1-g_{p_k}\right)^{n_b-\lfloor n_b g_{p_k}\rfloor}$$

$$\times g_{p_k}^{\lfloor n_b g_{p_k}\rfloor+1}\left(\lfloor n_b g_{p_k}\rfloor + 1\right)\binom{n_b}{\lfloor n_b g_{p_k}\rfloor + 1}\Bigg\},$$

$$(2)$$

where n_b denotes the size of body b, g_{p_k} denotes the fixed population share of group k, $\binom{a}{b}$ denotes the

---

[5] For confirmation of the formula for $\mathbb{E}[R_b]$ using Monte Carlo methods, see Figure S.II.6 in Supplementary Materials II.

binomial coefficient, $_aC_b$, and $\lfloor a \rfloor$ denotes the floor function (e.g., bringing 3.4 to 3).

Intuitively, $\mathbb{E}[R_b]$ represents the average representation score across myriad random draws with a body of size n_b and group population shares defined by the values of g_{p_k}—in effect, the degree of representation one would expect if representation was structured only by compositional factors.

This exercise is designed to capture overall representation, not the representation of particular groups. Note also that a group may be over-represented in some draws and under-represented in others. Our concern is with the average discrepancy across all of these draws (as opposed to the discrepancy of the average).

The complex dynamics of Equation 2 may be visualized by placing each dimension along one edge of a graph and observing how representation scores change as the parameters of these dimensions change. In these stylized scenarios, we assume that all members of the political body are drawn from a fixed population composed of one hundred persons. Since three explanatory elements are in play, two diagrams are required.

The left panel of Figure 2 focuses on body size and number of groups (of equivalent size). The bottom row is yellow, signaling perfect representation when society is composed of a single group. If all members of society are of Type A, all members of a political body drawn from that society must also be of Type A. The far-right column is also yellow, signaling perfect representation when the size of the political body equals the size of the population. If all members of society are included in a political body, that body (effectively, a popular assembly) must achieve perfect representation. Discrepancies increase as the number of groups increases or the body size shrinks, as depicted at the top-left of the diagram, where the representation score goes to 0.

The right panel of Figure 2 focuses on body size, as previously, along with group entropy, defined as

$$\text{Entropy}(g_{p_1}, g_{p_2}, ..., g_{p_K}) = -\sum_{k=1}^{K}g_{p_k}\log(g_{p_k}), \qquad (3)$$

where again g_{p_k} represents the population share for group k. Entropy is bounded theoretically at 0, where only one group is present in the population. Maximum entropy is achieved when all groups are of equal size. This value is unbounded as it depends on the total number of groups in play—more groups lead to higher values if all groups are of equal size.

For heuristic purposes, we limit our exercise to six groups (the median value of ethnic groups in our global dataset, described below). With the number of groups held constant, we can observe the impact of changes in their relative size. Entropy is low where there is extreme inequality, that is, one group encompasses nearly the entire population. We draw different group share values (which constitute a six-dimensional simplex) from a Dirichlet distribution with the α parameters all set to 1 so that all group share combinations are equally likely. We then

FIGURE 2. A Model of Representation
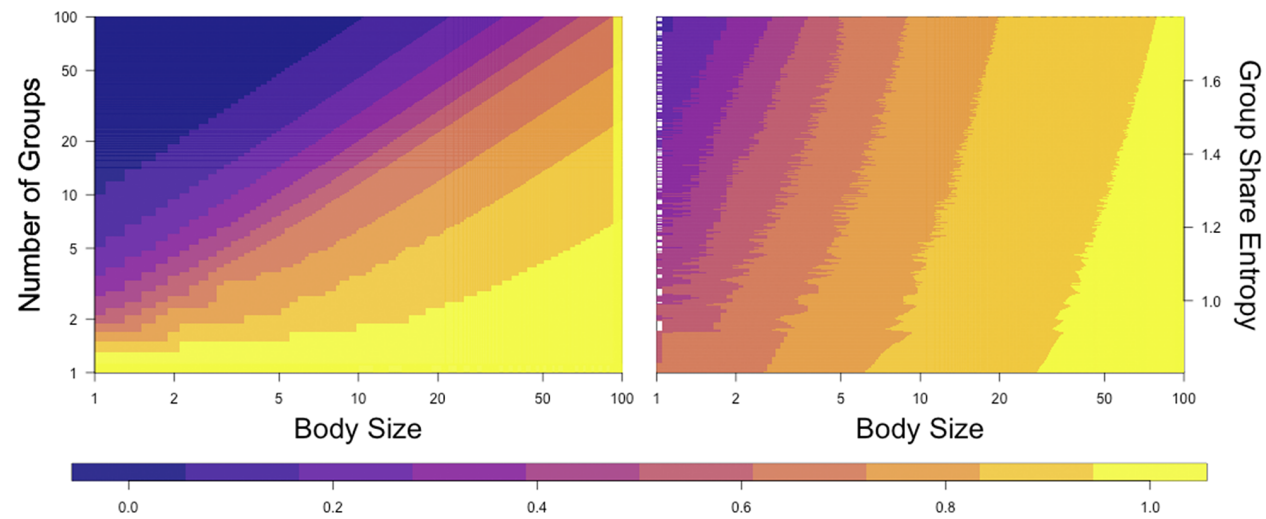
Note: Color values indicate different levels of the expected value of the representation index as outlined in Equation 2. Left: Values of expected representation decrease with the number of groups, but increase with body size. Right: Expected representation index also gets smaller as the population entropy grows.

average across the randomness inherent in this process of generating the group shares. For any body size (except where the size of the body equals the size of the population), increasing entropy decreases representation. Again, the upper left quadrant signals the point of worst representation—where body size is smallest and group entropy greatest.

This simulation exercise establishes a baseline for understanding the problem of coordination in descriptive representation. In subsequent sections of the paper, we test these theoretical expectations. It turns out that point estimates based on simulations are successful in explaining a good deal of the variance in representation across the world, reinforcing our theoretical expectation that failures in representation are often due to compositional factors.

DATA COLLECTION

Having examined our topic abstractly, we turn to the real world of representation. To do so, we extend the Global Leadership Project (Gerring et al. 2019) with new coding focused on the demographic characteristics of leaders and a second round of data collection. Details pertaining to the recruitment of experts and the collection of data are discussed in Supplementary Materials IV. In this section, we explain the coding and discuss challenges posed by missingness and measurement error.

In accordance with the scope conditions of our theory, we focus on four dimensions of identity where there appear to be strong preferences for descriptive representation. (Several additional traits including age and education are explored in Supplementary Materials I.) For each leader, country experts code *gender* (male or female),[6] *language* (mother tongue), *religion* (by birth), and *ethnicity*. The latter is understood as an *uber*-identity, describing the most important cleavage existing in a country at a particular point in time—which might be defined by race, religion, language, caste, region, cultural practices, or some combination of the foregoing.

In making coding decisions, country experts draw on a variety of sources including parliamentary websites, Wikipedia entries, country-specific sources, and clues implicit in a leader's name, place of birth, and so forth. Coding thus rests largely on each leader's presentation of self, their public persona—presumably crafted for political purposes. Since representation is a public act, it is appropriate to focus on how leaders present themselves. Fortuitously, this information is usually readily accessible.

After compiling the data, we arrive at a total of 2 genders (male and female), 11 religions, 280 languages, and 807 ethnicities. (A complete list of categories is posted in section S.IV.1.3 in Supplementary Materials IV.) To ascertain the size of each group in the general population, we consult censuses or surveys for each country—limiting our purview to legal citizens. (Since nonvoting aliens are generally disqualified from holding office, they are excluded from our analysis.) Only groups with more than one hundred thousand individuals or composing more than 1% of the population are included. (Groups falling short of this threshold are removed from the analysis even if

---

[6] At the time of data collection, few political leaders identified publicly as nonbinary.

TABLE 1. Descriptive Statistics

	Country coverage		Number of groups/bodies		Number of leaders coded	
	N	Rounds	Total	Per country	Total	Per country/round
Social groups	156	1,2	1,099	13	53,560	198
Ethnicity	152	1,2	807	5	44,363	172
Language	133	1	280	4	23,252	114
Religion	107	1	11	4	12,981	82
Gender	153	1,2	2	2	51,404	195
Political bodies	156	1,2	1,552	11	53,560	198
Executive	156	1,2	156	1	431	2
Supreme court	125	1,2	125	1	2,006	8
Party	114	1,2	939	10	38,008	232
Cabinet	150	1,2	150	1	6,025	27
Upper chamber	56	2	56	1	5,338	57
Lower chamber	126	1,2	126	1	41,376	225

they achieve representation, as happens on a few occasions.)

Leaders are classified into seven political bodies: (a) executive (including all persons who perform an executive function such as president or prime minister, but not those whose role is purely symbolic), (b) cabinet (with and without portfolio), (c) parliamentary party group, (d) upper chamber of the legislature (if bicameral), (e) lower (or unicameral) chamber of the legislature, (f) legislature at-large (both upper and lower chambers), and (g) supreme (or constitutional) court.

Since these categories are partially overlapping, leaders may belong to multiple political bodies, which means that units of analysis are not entirely independent. Supplementary analyses employ hierarchical bootstrap sampling to provide standard errors that account for the nonindependence of data points (see section "Additional Robustness Tests").

The resulting dataset, summarized in Table 1, incorporates 156 countries, two rounds of coding (2010–13 and 2017–19), 1,099 social groups, 1,552 political bodies, and 53,560 leaders. Coverage along these dimensions is more extensive than any comparable dataset, though limited to two points in time.

Even so, coverage is uneven. Political bodies, the main units of analysis in the tests that follow, are included only if 75% of their members are coded along the relevant social dimension. For a core group of 120 countries, coverage is fairly strong across all dimensions; for other countries, only one or two identities are successfully coded. There is also unevenness across rounds. Linguistic and religious identities are coded only in the first round, while upper chambers are coded only in the second round.

To cope with problems of missingness, we take several steps to ensure that the main results are robust. First, we replicate benchmark analyses with no threshold of inclusion: all political bodies for which *any* members are coded (for the relevant social dimension) are included. Second, we replicate benchmark analyses using datasets where missing values are

imputed. Third, we conduct analyses limited to a single round, or a single dimension of social identity. Results from these tests are very close to those obtained from the benchmark sample, as reported in tables and figures to follow.

In addition to problems of missingness, we must consider potential measurement error. Highly subjective features of human identity are sometimes difficult to define and to code. Reassuringly, an inter-coder reliability analysis shows a high level of agreement across expert coders (section S.IV.1.2 in Supplementary Materials IV). Moreover, when the same leader appears in rounds 1 and 2, their ethnicity is coded identically 97% of the time. This constitutes an informal inter-coder reliability test since most countries are coded by different experts across rounds.

Of course, one must also be concerned with the definition of social categories, which may be aggregated in different ways. For example, religion may be classified with coarse categories (e.g., Christian) or differentiated categories (e.g., Protestant, Catholic, and Orthodox). Additional complexities arise when one considers the intersection of multiple dimensions of identity, which may be interactive rather than simply additive (Celis and Erzeel 2017).

We cannot test all possible measures of identity—an essentially infinite set. However, we do explore different dimensions (gender, linguistic, religious, and ethnic), one intersectional identity (ethnic+gender), and regions of the world with different identity configurations. We also conduct analyses in which observed ethnic groups are combined in randomly chosen superordinate groups. Results from all of these tests (described below) are robust, offering reassurance that our findings are not hostage to arbitrary choices in categories or occasional measurement errors. Further reassurance is offered by a convergent validity test that compares our country-level index of ethnic representation with a comparable index from Ruedin (2009) (see Figure S.II.3 in Supplementary Materials II).

AN INDEX OF DESCRIPTIVE REPRESENTATION

Having introduced our data, let us consider how it might be aggregated into an overall index of representation, one suitable for comparisons across varied dimensions and contexts.

Traditionally, work on descriptive representation focuses on groups facing special discrimination or economic hardship, for which the term *minority* is loosely employed. Unfortunately, exclusion is not easy to measure, being a matter of degree and subject to differing perceptions. Moreover, discrimination and economic hardship do not always overlap (groups may face discrimination despite being relatively affluent) and are not always associated with relative size (complicating the idea that an oppressed group must necessarily be a statistical minority). Finally, one must reckon with societies where groups are not clearly ranked by social or socioeconomic status, a pattern common in sub-Saharan Africa. What does adequate representation mean in these contexts?

To sidestep these obstacles, we accept the reality that groups may be defined in any number of ways—by ethnicity, gender, and so forth—and that these categories may be understood in different ways and combined in various ways, generating intersectional categories.

However defined, we propose an encompassing approach to measurement. *All* members of society must fit somewhere within the chosen dimension; they must count as members of some group when calculating overall representation across that dimension. If ethnicity is the dimension of interest, for example, all persons must be placed within an ethnic category, not just those deemed "minorities." Helpfully, representation is a zero-sum outcome: one group's representation must be achieved at the expense of another's. If one ethnic group is over-represented, another must be under-represented. Accordingly, we need not agonize over whom to identify as minorities and majorities or attempt to ascertain which groups are subject to what degree of discrimination.

In aggregating across groups (along a single dimension of identity), we employ the representation index outlined in Equation 1, replacing theoretical quantities with data collected for this project:

$$\hat{R}_b = 1 - \frac{1}{2} \sum_{k=1}^{K} |\hat{g}_{p_k} - \hat{G}_{b_k}|. \tag{4}$$

Here, $\hat{R}_b$ is the estimated degree of representation present in a particular political body, $\hat{g}_{P_k}$ is an estimate of group k's share of the population from country-level sources, $\hat{G}_{B_k}$ is the estimate of group k's share of a political body from our data, and K is the total number of groups in the population.

To illustrate the dynamics of Equation 4, imagine a society with three equal-sized groups (33%, 33%, 33%). At one extreme, each group achieves equal representation (33%, 33%, 33%), rendering a representation score of 1.00, a perfectly proportional relationship between body and population group shares.

At the other extreme, all offices are controlled by individuals who are not members of the population (0%, 0%, 0%), for example, foreign colonizers or an occupying power. Here, the representation score is precisely 0. In a more typical scenario, one group achieves twice the number of offices as the other two (50%, 25%, 25%), generating a representation score of 0.83.

Country Scores

So measured, how does descriptive representation vary across countries? To answer this question, we must aggregate representation scores across offices within each country (executive, parliamentary, and judicial) for each chosen dimension of social identity (gender, language, religion, and ethnicity), and then aggregate across those dimensions to obtain a summary score.

Since missingness across countries may introduce bias, we first impute missing values for political bodies that are partially coded in rounds 1 or 2. Values are imputed by fitting a nonlinear prediction model for each variable and iteratively predicting missing values with that model until convergence (Stekhoven 2015). Background factors in the imputation stem include covariates from Model 4 in Table 2. (Imputations for *expected* representation in Figure 3 are conducted separately.)

Table S.III.1 in Supplementary Materials III provides a complete list of countries and their representation scores across all four dimensions, combining results for all (national) political bodies in each country. A summary score for each country is derived by averaging across the four dimensions—gender, language, religion, and ethnicity.

So calculated, the mean representation score across all countries in our sample is 0.73, the standard deviation is 0.07, and the minimum/maximum values are 0.50 and 0.88. The highest levels of overall representation are achieved in Iceland, Poland, Norway, Denmark, and Finland, while the lowest levels are registered in Sierra Leone, Central African Republic, Solomon Islands, Indonesia, and Congo (DRC).

To demonstrate the contribution of compositional factors to these country-level results, Figure 3 plots observed values against values we would expect under conditions of random sampling given the size of each political body and the number and relative size of social groups, as described by Equation 4.

Evidently, random sampling offers a very reasonable approximation of the selection process if compositional factors are taken into account—a point confirmed in the first analysis of Table 2. As we shall see, the fit between them is not substantially improved when institutional, sociological, and economic factors are added to the model.

Note, however, that all data points fall below the diagonal line in Figure 3. This means that all countries are less representative than they would be if the selection of political leaders were entirely random. (The size of this representational gap is calculated for each country in Table S.III.1 in Supplementary Materials III.) Indeed, the maximum observed value (0.88) is only

TABLE 2.   Main Analysis

	Model 1	Model 2	Model 3	Model 4	Model 5
Expected representation	0.84 (87.72)*	0.85 (81.91)*			
Body size (log)			0.04 (13.34)*	0.05 (13.91)*	0.05 (12.28)*
Fractionalization			−0.69 (−24.92)*	−0.63 (−13.72)*	−0.66 (−21.11)*
Continuous covariates					
Lexical index					0.01 (2.64)*
Population (log)					−0.01 (−1.65)
GDP per capita (log)					0.01 (0.90)
Gini index					0.00 (−0.82)
Factor covariates					
Identity				✓	✓
Body type				✓	✓
Gender quota type				✓	✓
Ethnicity quota type				✓	✓
Selection rule				✓	✓
Round				✓	✓
Country				✓	
Intercept			✓	✓	✓
Other statistics					
Countries	156	156	156	156	156
No. of obs.	156	6,628	6,628	6,628	6,628
Adj. R^2	0.60	0.39	0.46	0.51	0.48
Unit of analysis	C	C-B-G	C-B-G	C-B-G	C-B-G
Dependent variable	Rep. index	Rep. index	Rep. index	Rep. index	Rep. index

Note: Outcome: representation index (where 1 = perfect representation), measured across various identities—ethnicity, religion, language, and gender. Estimator: ordinary least squares, *t*-statistics in parentheses, standard errors clustered by country. * denotes $p < 0.05$. Missing values were imputed in Model 1 to ensure compatibility across country and standard errors calculated via the country-level block bootstrap (with the imputation model re-fit on every bootstrap draw); see main text. In the unit of analysis row, "C" denotes country, "G" denotes group, and "B" denotes body. Full model results are given in Table S.V.1 in Supplementary Materials V.

FIGURE 3.   The Relationship between Observed and Expected Representation, Aggregated to the Country Level
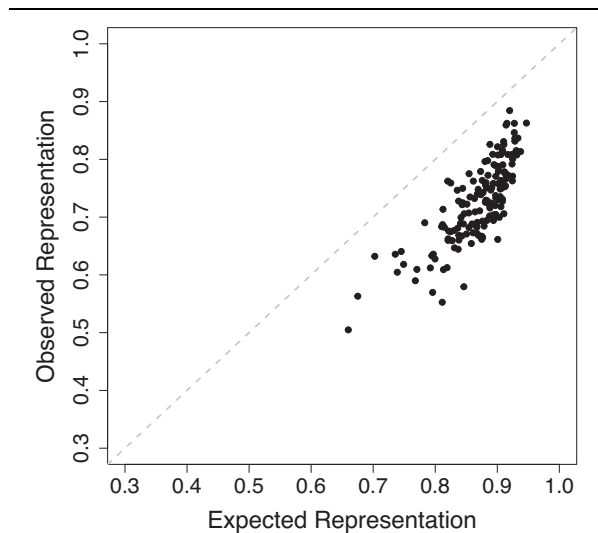
Note: Missing representation values have been imputed to ensure comparability across country as described in the main text. A regression model summarizing this relationship can be found in column 1 of Table 2 (see Table S.V.1 in Supplementary Materials V for full model specification).

slightly greater than the average value of expected representation (0.87).

MAIN TESTS

Having looked at representation in a descriptive fashion, we turn to a series of analyses that attempt to probe the causal effect of compositional factors. Initial tests are shown in Table 2 (for descriptive statistics, see Table S.II.1 in Supplementary Materials II).

Model 1 replicates the bivariate scatterplot shown in Figure 3. Here, country-level representation scores are regressed against those predicted by Equation 4 (in which leaders are chosen randomly, taking into account the size of political bodies and the distribution of social groups). More formally, we assume

$$\hat{R}_c = \gamma \times \mathbb{E}[\hat{R}_c] + \epsilon_c, \qquad (5)$$

where *c* is a country index, as described in the previous section. When the γ coefficient in this single-parameter model is greater than 1, observed representation will be on average higher than expected under random sampling; when below 1 it is lower than expected. This bivariate model explains over half of the variance in representation across countries.

In further analyses, the outcome is disaggregated. Rather than country-level scores, we look at the representation score for each political body across each identity dimension (gender, language, religion, and ethnicity) and each round of data collection. An individual observation is therefore composed of a political body, a social identity, and a coding round. For this purpose, we employ raw (unimputed) data, with the provision that a political body is included if 75% of its members are coded across a particular dimension of identity. The total size of the body is therefore defined as the number of members whose identity is known. For example, if 90 members of a one hundred-member legislature are coded for ethnicity, $n=90$.

While offering greater empirical leverage, this disaggregated approach is likely to increase stochastic error. Numerous factors may affect the level of representation achieved for a particular political body along a particular dimension of social identity at a particular point in time. These stochastic factors are minimized in country-level aggregate scores, as instances of over- and under-representation cancel each other out. In body-level analyses, they remain, and thus are likely to weaken the overall fit of the model.

Model 2 in Table 2 includes only the expected representation index, that is, the prediction issued by Equation 4 under conditions of random sampling and using the same structure as in Equation 5. This bivariate model (no intercept) accounts for over one-third of the variance. As expected, overall model fit is attenuated relative to Model 1, which we attribute to increased stochastic error. However, the point estimates for expected representation are nearly identical.

In further tests, we distinguish two compositional factors. The size of each political body is understood as its membership, transformed by the natural logarithm (to account for the diminishing marginal impact of larger membership). The dispersion of groups is measured with the Herfindahl index of fractionalization, which captures the probability that two randomly chosen individuals belong to the same social group (calculated separately for gender, linguistic, religious, and ethnic groups). Model 3, including only these variables, explains over two-fifths of the variability in representation.

Model 4 adds dummies for each social identity (ethnicity, religion, language, and gender), for the existence of a gender quota, for body type (executive, cabinet, et al.), for the selection rule applicable to that office (appointive, proportional, majoritarian, mixed, indirectly elected, and other), for each round of data collection (1 and 2), and for each country.

Model 5 drops country fixed effects in favor of country-level covariates. In choosing covariates, we rely on extant work and our own hunches about factors that could plausibly affect representation. Chosen covariates include the Lexical index of electoral democracy (Skaaning, Gerring, and Bartusevicius 2015), population (log), per capita GDP (log), and inequality, measured by the Gini coefficient of income inequality.

Estimates for the two compositional factors are extremely close across Models 3–5, despite dramatic changes in specification. Moreover, additional covariates scarcely improve model fit and also generally exhibit small t-statistics (see Table S.V.1 in Supplementary Materials V), suggesting that they are relatively minor influences on representation in this global sample.

As a further specification probe, we assess the out-of-sample predictive importance of right-side predictors from Model 5 via Lasso regression, which imposes a penalty on the absolute magnitude of coefficients, thereby setting some coefficients exactly to 0 unless they meaningfully improve predictive out-of-sample performance (approximated by cross-validation). We find that the regularizing Lasso model sets many of the model coefficients to exactly 0 but leaves as nonzero both body size and fractionalization, indicating that these quantities meaningfully improve out-of-sample predictions (see Table S.II.10 in Supplementary Materials II).

All of the available evidence suggests that a principal driver of representation is compositional: larger bodies generate better representation, and more heterogeneous countries generate worse representation. To get a sense of these effects, Figure 4 plots the expected values generated by the benchmark specification (Model 3, Table 2). Across our sample, fractionalization has a slightly steeper curve; nonetheless, a shift along the x-axis from minimum to maximum values translates into a substantial shift in representation for both regressors. Estimates are also precise, signaled by the extremely tight confidence bounds around these estimates.

INSIDE THE BOX

Having offered a parsimonious account of compositional effects, we are now in a position to disaggregate the treatment, thereby shedding light on potential mechanisms and also on further implications of our theory.

Based on the idea that larger bodies provide better representation, we infer that the representational capacity of various bodies follows their size. In our sample, average membership is as follows: executive ($n = 1$–2), supreme court ($n = 9$), parliamentary party ($n = 18$), cabinet ($n = 20$), upper house ($n = 80$), and lower house ($n = 232$). Accordingly, we expect the degree of representation to increase in a monotonic fashion from the smallest body to the largest using this indirect measure of body size.

To test this expectation, the first model in Table 3 regresses the representation index against these body types (with executive as the excluded category) along with fixed effects for identity, quota type, selection rule, round, and country. Results accord with theoretical expectations insofar as larger bodies are generally more representative. The notable exception is parliamentary parties, which may reflect the concentrated support that some parties obtain from specific identity groups.

FIGURE 4. Predicted Representation Index Values Based on Model 3 in Table 2 with 95% Confidence Intervals
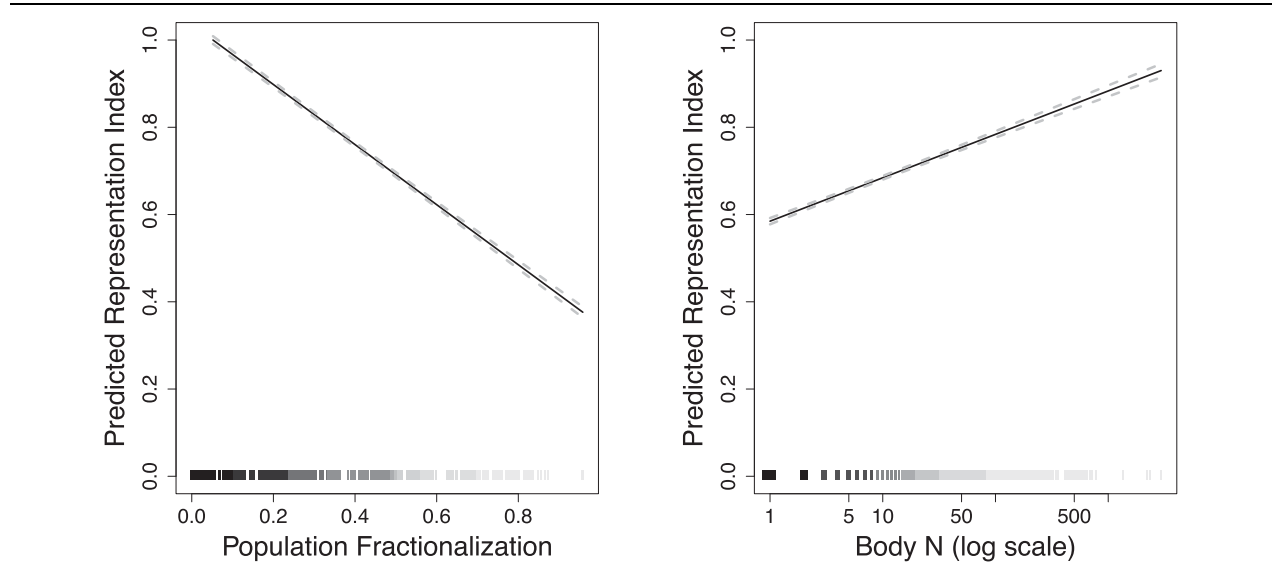
Note: Mean/median/SD values across the sample: body size (35/6/123), fractionalization (0.43/0.50/0.21). Above the *x*-axis labels for both plots, we display rug plots illustrating the empirical density of data points in our sample.

TABLE 3. Implications of the Main Analysis

	Entire (1)	Parties only (2)	Entire (3)	Entire (4)
Body size (log)		0.05 (12.83)*	0.05 (13.96)*	0.05 (14.05)*
No. of groups (log)				−0.20 (−9.32)*
Group entropy			−0.19 (−11.29)*	
Fractionalization	−0.64 (−14.56)*	−0.68 (−12.46)*		
Body indicators				
Executive baseline (1–2)	✓			
Supreme court (9)	−0.01 (−0.33)			
Party (18)	−0.10 (−3.40)*			
Cabinet (22)	0.01 (0.15)			
Upper house (80)	0.00 (−0.07)			
Lower house (232)	0.02 (0.85)			
Factor covariates				
Identity	✓	✓	✓	✓
Body type			✓	✓
Gender quota type	✓	✓	✓	✓
Selection rule	✓		✓	✓
Round	✓	✓	✓	✓
Country	✓	✓	✓	✓
Other statistics				
Countries	156	114	156	156
No. of obs.	6,628	4,274	6,628	6,628
Adj. R^2	0.46	0.52	0.50	0.45
Dependent variable	Rep. index	Rep. index	Rep. index	Rep. index

Note: Outcome: representation, measured for each identity—ethnicity, religion, language, and gender. Higher values indicate better representation. Estimator: ordinary least squares, *t*-statistics in parentheses, standard errors clustered by country. * denotes $p < 0.05$. Full model results are given in Table S.V.2 in Supplementary Materials V.

A better research design compares parties of different sizes to each other, thereby sidelining a great many background factors that might serve as confounders. Following our theory, larger parties should be more representative than smaller parties. Model 2 is therefore limited to parliamentary party groups, along with the usual vector of controls (excluding selection rule, which is collinear with country fixed effects).

Coefficients for the compositional factors of theoretical interest are nearly identical to that of the benchmark model in Table 2, confirming that parties follow the pattern established for other political bodies.

Finally, our theory suggests that both the number of groups (log) and group entropy exert independent effects on representation. In our benchmark specification, these are combined into a single fractionalization measure. In Models 3 and 4, we differentiate these factors, tested separately by virtue of their collinearity. As expected, the number of groups and their entropy (similarity in size) both reduce representation.

DIMENSIONS OF IDENTITY

In this section, we explore compositional effects across different identity categories. To the four familiar dimensions—ethnicity, religion, language, and gender—we add a measure of *intersectionality*, formed from the intersection of ethnicity and gender (following Weldon 2008). There are of course many other potential intersectionalities one might explore (Collins and Bilge 2020). However, since our construction of ethnicity aims to represent the most important cleavage in each society, and gender is orthogonal to ethnicity, this seems a logical choice.

Density plots for these five dimensions of identity, displayed in Figure 5, show the empirical distribution of representation scores across each dimension. The curves for language, ethnicity, and religion are similar,

with modes just below 1 (perfect representation) and long left tails. Evidently, many political bodies achieve decent representation along these dimensions, while some are horribly askew.

Gender has an accentuated mode at 0.5, marking the point where political bodies are dominated by a single gender (male). The truncated left tail is a product of the distribution of gender in populations across the world. Because men and women compose roughly half of the population everywhere, the greatest possible violation of equal representation—that is, the total exclusion of women from public office—is not as extreme a violation as the total exclusion of a linguistic, religious, or ethnic group comprising a super-majority. For example, the exclusion of Blacks from representation in Apartheid South Africa, where they composed roughly eighty percent of the population, would render a lower representation score than the exclusion of women.

Among the four core dimensions of identity—gender, religion, language, and ethnicity—the mean values of representation, illustrated across the x-axis of Figure 5, are remarkably close. This is surprising given the disparate nature of these identities, their disparate histograms, and their weak intercorrelations (see Figure S.I.1 in Supplementary Materials I and also Ruedin 2010). One would have thought that some identities would be better represented than others. Yet, we find little variation across sample means.

By contrast, the intersectional index has a lower mean, and wider dispersion, than other dimensions of representation. A plausible explanation, consistent

FIGURE 5. The Shape of Descriptive Representation
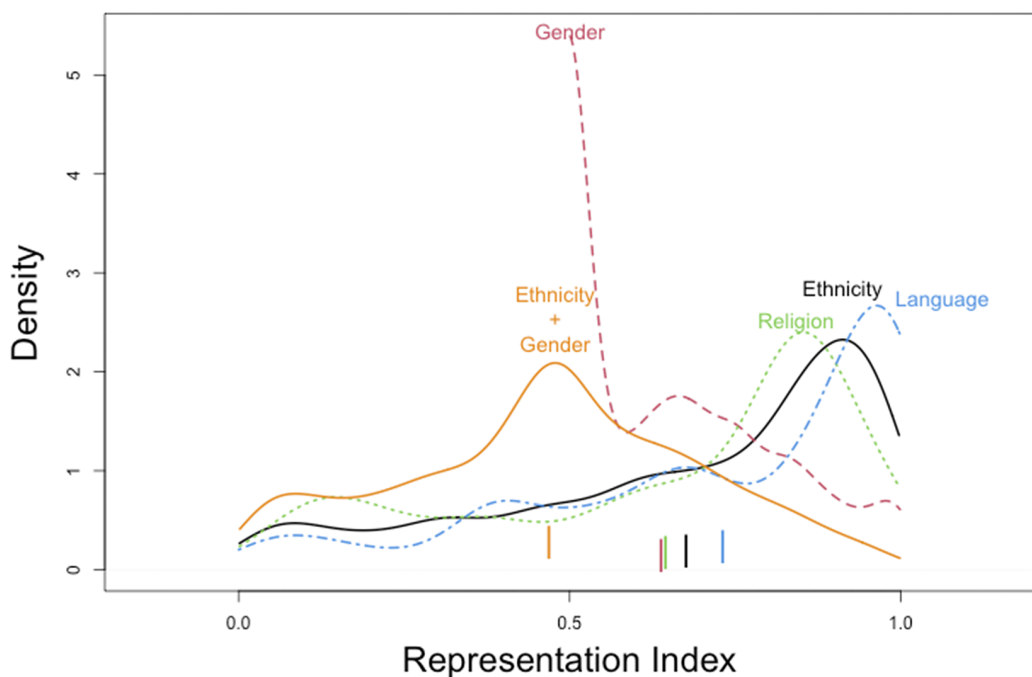
Note: Descriptive statistics (mean/median/SD): gender (0.64/0.60/0.16), ethnicity (0.68/0.77/0.28), language (0.73/0.85/0.28), religion (0.65/0.77/0.28), ethnicity–gender intersection (0.47/0.48/0.23). Group-level means are represented as tick marks at the bottom of the figure (with random jitter added along the *y*-axis to make the lines distinguishable).

TABLE 4.   Analysis by Group Identity

	Ethnicity (1)	Religion (2)	Language (3)	Gender (4)	Ethnicity+Gender (5)
Body size (log)	0.06 (8.30)*	0.02 (1.51)	0.05 (7.23)*	0.05 (14.36)*	0.08 (16.06)*
Fractionalization	−0.69 (−421.67)*	−0.72 (−6.14)*	−0.64 (−16.94)*	−0.95 (−3.22)*	−0.82 (−15.82)*
Continuous covariates					
Lexical index	✓	✓	✓	✓	✓
Population (log)	✓	✓	✓	✓	✓
GDP per capita (log)	✓	✓	✓	✓	✓
Gini index	✓	✓	✓	✓	✓
Factor covariates					
Body type	✓	✓	✓	✓	✓
Gender quota type				✓	✓
Ethnicity quota type	✓				✓
Selection rule	✓	✓	✓	✓	✓
Round	✓			✓	✓
Other statistics					
Countries	152	107	133	153	148
No. of obs.	2,076	871	1,377	2,304	1,913
Adj. R^2	0.57	0.26	0.60	0.27	0.61
Dependent variable	Rep. index	Rep. index	Rep. index	Rep. index	Rep. index

Note: Outcome: representation index. Higher values indicate better representation. Estimator: ordinary least squares, *t*-statistics in parentheses, standard errors clustered by country. * denotes $p < 0.05$. Full model results are given in Table S.V.3 in Supplementary Materials V.

with our theory, is that the multiplication of categories introduces greater social diversity, and with it additional coordination problems that translate into lower overall representation.

For our purposes, the most important issue is whether compositional effects vary across different dimensions of social identity. To assess this question, we replicate the benchmark specification (from Table 2) across each dimension in Table 4. Samples are focused on ethnicity (Model 1), religion (Model 2), language (Model 3), gender (Model 4), and intersectionality (Model 5).[7]

There is some evidence of causal heterogeneity across the five measures of identity. Of course, these may be stochastic or a product of varying samples. One must resist the temptation to over-interpret small differences.

The varying *precision* of these estimates can be probed in our random sampling framework, following section "A Random Sampling Model." This analysis, visualized in Figure S.II.1 in Supplementary Materials II, explores the residual standard deviation of the theoretical model, $\sqrt{\mathbb{E}[(R_b - \mathbb{E}[R_b])^2]}$, as we vary body size, group number, and group entropy. We find that the residual standard deviation shrinks as the number of groups increases. This could explain why compositional features are more precisely estimated for intersectional identities than for other social identities (as captured by *t*-statistics across these models).

In any case, the main takeaway is that body size is always associated with increased representation, while fractionalization is always associated with reduced representation. Most estimates are similar to the benchmark model in Table 2.

Of particular note are results contained in Model 1. Recall that ethnicity represents the most salient cleavage in a country at the time the data were collected. This is of course a judgment, and we must rely on our expert coders to perceive which dimensions are most important in a particular context. Nonetheless, it suggests that compositional factors matter as much for the most important cleavage as they do for other cleavages. These relationships are probed at greater length in section S.II.1.5 in Supplementary Materials II.

CONTEXTS

Having examined compositional effects across different identity categories, we now explore background factors that might impact compositional effects on representation. To test potential moderators, the full sample is divided into sub-samples according to the background factor of interest. Results of these paired tests are shown in Tables 5 and 6.

First, we explore the nature of the office, categorized as elective (Model 1, Table 5) or unelected (Model 2). The latter category includes supreme courts, cabinets, and a few parliamentary parties where there is no apparent elective process. We find that effects persist across both samples. The impact of body size and fractionalization is somewhat weaker across unelective bodies, though this may simply

---

[7] To generate a measure of population fractionalization for the intersection of ethnicity and gender, we adopt the simplifying assumption that all ethnic groups are composed equally of men and women, allowing us to generate a fractionalization index comparable to those for unidimensional identities.

TABLE 5. Analysis in Varying Contexts

	Elected (1)	Unelected (2)	Dem. (3)	Non-dem. (4)	R. 1 (5)	R. 2 (6)
Body size (log)	0.05 (10.24)*	0.03 (3.46)*	0.05 (10.47)*	0.05 (8.60)*	0.05 (11.89)*	0.06 (12.48)*
Fractionalization	−0.71 (−19.41)*	−0.57 (−18.07)*	−0.65 (−15.38)*	−0.71 (−22.86)*	−0.72 (-21.89)*	−0.67 (−19.96)*
Continuous covariates						
Lexical index	✓	✓			✓	✓
Population (log)	✓	✓	✓	✓	✓	✓
GDP per capita (log)	✓	✓	✓	✓	✓	✓
Gini index	✓	✓	✓	✓	✓	✓
Factor covariates						
Identity	✓	✓	✓	✓	✓	✓
Body type	✓	✓	✓	✓	✓	✓
Gender quota type	✓	✓	✓	✓	✓	✓
Ethnicity quota type	✓	✓	✓	✓	✓	✓
Selection rule	✓	✓	✓	✓	✓	✓
Round	✓	✓	✓	✓		
Other statistics						
Countries	156	155	102	54	152	150
No. of obs.	5,342	1,286	5,049	1,579	2,371	2,009
Adj. R	0.47	0.43	0.46	0.55	0.49	0.47
Dependent variable	Rep. index	Rep. index	Rep. index	Rep. index	Rep. index	Rep. index

Note: Outcome: representation, measured for each identity—ethnicity, religion, language, and gender. Higher values indicate better representation. Estimator: ordinary least squares, *t*-statistics in parentheses, standard errors clustered by country. * denotes $p < 0.05$. Full model results are given in Table S.V.4 in Supplementary Materials V.

TABLE 6. Heterogeneity Analysis by Region

	OECD (1)	Non-OECD (2)	Americas (3)	Asia (4)	Europe (5)	MENA (6)
Body size (log)	0.05 (6.63)*	0.05 (12.24)*	0.05 (5.96)*	0.04 (6.52)*	0.05 (5.73)*	0.05 (8.94)*
Fractionalization	−0.70 (−13.47)*	−0.67 (−20.11)*	−0.74 (−13.50)*	−0.64 (−9.44)*	−0.55 (−14.04)*	−0.72 (−15.42)*
Continuous covariates						
Lexical index	✓	✓	✓	✓	✓	✓
Population (log)	✓	✓	✓	✓	✓	✓
GDP per capita (log)	✓	✓	✓	✓	✓	✓
Gini index	✓	✓	✓	✓	✓	✓
Factor covariates						
Identity	✓	✓	✓	✓	✓	✓
Body type	✓	✓	✓	✓	✓	✓
Gender quota type	✓	✓	✓	✓	✓	✓
Ethnicity quota type	✓	✓	✓	✓	✓	✓
Selection rule	✓	✓	✓	✓	✓	✓
Round	✓	✓	✓	✓	✓	✓
Other statistics						
Countries	36	120	25	29	38	58
No. of obs.	1,589	5,039	1,069	1,456	1,766	2,097
Adj. R^2	0.53	0.46	0.47	0.48	0.43	0.54
Dependent variable	Rep. index	Rep. index	Rep. index	Rep. index	Rep. index	Rep. index

Note: Outcome: representation, measured for each identity—ethnicity, religion, language, and gender. Higher values indicate better representation. Estimator: ordinary least squares, *t*-statistics in parentheses, standard errors clustered by country. * denotes $p < 0.05$. Full model results are given in Table S.V.5 in Supplementary Materials V.

reflect the limited variability of our sample, composed largely of small bodies.

Second, we differentiate democracies (Model 3) and autocracies (Model 4). Democracies are understood as polities with minimally competitive multiparty elections for the legislature and the executive, operationalized as a score of 4–6 on the Lexical index (Skaaning, Gerring, and Bertusevicius 2015). We find virtually no difference in estimates for our two variables of theoretical interest across these sub-samples; compositional effects are equally strong in autocracies and democracies.

Third, we compare two periods in time corresponding to Round 1 (2010–13) and Round 2 (2017–19) of the data collection process. To make this comparison exact, we include only representation by ethnicity and gender, which were coded for both rounds. Changes in leadership and representation between rounds scarcely affect estimates of compositional effects across the two samples, as shown in Model 5 (Round 1) and Model 6 (Round 2).

Table 6 continues the exercise with another set of comparisons. First, we compare rich, industrialized countries with poorer, less developed countries. To differentiate the two groups, the sample is divided into OECD countries (Model 1) and non-OECD countries (Model 2). (The Lexical index is excluded from Model 1 as there is no variability in regime type within this sub-sample.) There is little difference across these sub-samples, suggesting that compositional effects are not moderated by economic development.

Next, we compare various regions of the world: the Americas (Model 3), Asia (Model 4), Europe (Model 5), and the Middle East and North Africa (Model 6). Again, coefficient estimates for the variables of theoretical interest are stable.

INSTRUMENTAL VARIABLES

Neither of the key variables in this study is randomly assigned, so one must consider whether the data-generating process might in some way confound estimates reported in previous tables. It is possible, for example, that the degree of diversity in a country—or the degree of representation achieved in a country—affects its institutions, including the size of political bodies. It is possible that the degree of representation and recognition achieved by social groups affects their self-definition (Liu 2011). It is even possible that deep-seated social norms affect the shape of society, the shape of institutions, and the representation of social groups.

We are at pains to work out all the possible ways in which confounders might affect the analyses presented in previous tables—though we take some comfort in the stability of the results across different specifications (which include controls for economic development, democracy, inequality, and country fixed effects) as well as the theoretical derivations presented in section "Coordination Problems." In this section, we approach the challenge of causal identification with instruments.

In order to serve their intended function, the chosen instruments must be exogenous and must affect the outcome only through the treatment variable. Because these assumptions are impossible to prove, we regard estimates posted in Table 6 as robustness tests rather than baseline models.

As instruments for the size of political bodies, we employ political body types, categorized as executive, upper house, or lower house in countries where membership to those bodies is determined by voter input. (Other types are excluded.) The assumption is that these body types are predictors of body size (as shown in Table 3) but are not for other reasons likely to be more or less representative. (This assumption would be violated if selectors relate to the identity dimension of candidates differently across political body types.)

To instrument for ethnic fractionalization, we employ a geographic feature—dispersion in elevation across regions of a country—grounded in work on the long-run sources of ethnic diversity (Michalopoulos 2012). (Other dimensions of identity are here excluded.) Geography is assumed to be exogenous and unrelated to representation, except through its influence on diversity. Figure S.II.8 in Supplementary Materials II visualizes the assumptions of these IVs.

The first-stage analyses in Table 7 offer a reasonably good fit to the data, and the IV diagnostics are favorable. (For example, we reject the null hypothesis in the weak instruments test; we also reject the null in the Wu–Hausman test, indicating that the use of instruments is helpful in accounting for endogeneity.) The second-stage analyses report estimates for the key variables that are comparable to those reported in Table 2, though slightly stronger for fractionalization. These analyses offer some reassurance against threats to inference stemming from nonrandom assignment.

ADDITIONAL ROBUSTNESS TESTS

In this section, we briefly discuss additional robustness tests whose full results are posted in Supplementary Materials II.

First, in section S.II.1.6 in Supplementary Materials II, we assess the extent to which our results are robust to the potentially arbitrary aggregation of ethnic categories, an issue discussed in section "Data Collection." To do so, we take the ethnic groups for a given country, randomly aggregate them into higher-order categories, and replicate our main analyses.

Second, we return to the problem that our use of clustered standard errors can account for error covariances within countries but not the hierarchical structure of our data, where information about a single leader can contribute to multiple observations. In section S.II.1.9.2 in Supplementary Materials II, we replicate the main analysis employing a hierarchical bootstrap procedure described in section S.II.1.7 in Supplementary Materials II that accounts for multiple levels of uncertainty.

TABLE 7. IV Analysis

	S1: log(Body N) (1)	S2: Rep. index (2)	S1: Frac (3)	S2: Rep. index (4)
Body size (log)		**0.03 (2.58)***	0.00 (0.04)	0.06 (9.28)*
Fractionalization	−0.16 (−1.51)	−0.57 (−16.01)*		**−1.32 (−3.04)***
Instruments				
Executive baseline	✓			
Upper house	3.25 (6.04)*			
Lower house	4.25 (7.99)*			
SD (elevation)			0.09 (1.70)	
Continuous covariates				
Lexical index	✓	✓	✓	✓
Population (log)	✓	✓	✓	✓
GDP per capita (log)	✓	✓	✓	✓
Gini index	✓	✓	✓	✓
Factor covariates				
Gender quota type	✓	✓		
Ethnicity quota type	✓	✓	✓	✓
Selection rule	✓	✓	✓	✓
Round	✓	✓	✓	✓
Other statistics				
Countries	156	156	134	134
No. of obs.	1,084	1,084	1,859	1,859
Adj. R	0.96	0.46	0.30	0.34
Weak instruments	453.82*		33.18*	
Wu–Hausman	0.27		17.16*	

Note: First stage outcomes: log(body size) and fractionalization. Second stage outcome: levels of representation (where 1 = perfect representation). Estimator: two-stage least squares, *t*-statistics in parentheses, standard errors clustered by country. * denotes $p < 0.05$. Full model results are given in Table S.V.6 in Supplementary Materials V. Bold values are those of theoretical relevance and interpreted in the main text.

Third, we evaluate the degree to which our results may be affected by post-treatment bias due to the inclusion of parliamentary parties in the analysis. These bodies are post-treatment in the sense that the composition of a party could affect its popularity and, thus, its representation in the legislature. To obviate this issue, we replicate the main analysis excluding parties. Results are presented with clustered standard errors (section S.II.1.10.1 in Supplementary Materials II) and with significance assessed via the hierarchical bootstrap described above (section S.II.1.10.2 in Supplementary Materials II).

Finally, we assess whether results might be affected by our inclusion criterion, whereby political bodies are included only if we are able to gather information on the identity of more than 75% of its members along a particular dimension of identity. Relaxing this criterion, we include all political bodies for which *any* member can be coded. Results are presented with clustered standard errors (section S.II.1.11.1 in Supplementary Materials II) and a hierarchical bootstrap (section S.II.1.11.2 in Supplementary Materials II). In Figure S.II.4 in Supplementary Materials II, we plot the main regression coefficients using Model 4 in Table 2 as we vary the coverage threshold between 0 and 1. Section S.II.1.8 in Supplementary Materials II replicates the main analysis when randomly perturbing ethnic identities for which coders were uncertain.

These tests offer some reassurance with respect to the robustness of our main findings, as deviations from the benchmark model are generally small.

LIMITATIONS

Although capacious, the present study encounters several empirical limitations.

First, many traits—including sexual orientation, disability, and social class—are unexamined due to the difficulty of collecting data on these subjects on a global scale. Our theory suggests that these dimensions of identity are subject to compositional effects insofar as they are valued by selectors, a point that we develop in Supplementary Materials I but do not have an opportunity to test.

Second, our data sample the world at two points in the contemporary era. Accordingly, we are unable to directly address longer-term historical patterns. We theorize that as views about social groups change these changes should be reflected in compositional effects. Systematic tests must await better evidence.

A third limitation concerns the national focus of our data. We see no reason to suppose that subnational bodies are exempt from compositional effects, and a smattering of studies focused on local councils in the United States (cited at the outset) mostly supports that interpretation. Further research is needed, particularly outside the United States.

A fourth limitation stems from our global approach, which is unable to capture nuances of descriptive representation—arising, for example, from laws, norms, history, and geography—that pertain to specific countries or social groups. Helpfully, a rich body of research focuses on these issues in specific contexts, especially in

Europe and the United States.[8] The present study should be viewed as a complement, not a replacement, for these focused studies.

Finally, we have little ground for speculating upon compositional effects in *nonpolitical* bodies such as firms, labor unions, and other nongovernmental organizations. Insofar as there is growing pressure to represent society in these bodies, we would not be surprised if a similar dynamic applies. So far as we know, the impact of compositional factors on representation in these venues has not been studied; nor is it entirely clear how relevant constituencies should be defined in these contexts. We leave these matters for future research.

CONCLUSIONS

Nation-states are premised on the existence of a political community. Yet, they are often melded together in an arbitrary fashion, including disparate peoples with little sense of common identity (Anderson 2006). Democracy, in conjunction with new social movements, may encourage the efflorescence of distinctions (Benhabib 1996). Under the circumstances, one should not be surprised that the task of descriptive representation has proven to be a challenging one in the twenty-first century.

To better understand this phenomenon, we proposed an approach to measuring descriptive representation that is general in purview—applying to any society, any political body, and any dimension of identity. We then showed how this representation index maps onto newly gathered data covering 4 dimensions of identity, 156 countries, 1,552 political bodies, 2,052 social groups, and 53,560 political leaders.

Across this global sample, we find that descriptive representation falls considerably short of the ideal. In no country are all social groups represented in rough proportion to their share of the population. In no country does descriptive representation even reach the level that would be expected if representatives were chosen randomly (Figure 3). Moreover, aggregate shortfalls in representation are very similar across the four major dimensions of identity explored in this study —gender, religion, language, and ethnicity (Figure 5). The consistency of these patterns suggests that the problem of descriptive representation may be subject to generic features that are not captured by studies focused on a single country or a single dimension of identity.

On the basis of theory and a random sampling model, we argue that efforts to achieve descriptive representation encounter a formidable coordination problem centered on the size of political bodies and the configuration of social identities. Specifically, larger bodies are more representative than smaller bodies, and heterogeneous polities are less representative than homogeneous polities.

These compositional factors account for roughly half of the variability in descriptive representation across bodies and countries throughout the world today—a little more than half if aggregated by country, a little less if disaggregated by political body (Table 2). By contrast, other factors such as the type of office (executive, cabinet, et al.), selection rules (appointive, PR, et al.), regimes (democracy/autocracy), levels of economic development (per capita GDP), and inequality (the Gini index) appear to have only marginal impact on representation.

So far as we can tell, compositional effects are ubiquitous for traits whose representation is valued. Although we cannot claim to have tested every plausible context, we find evidence of compositional effects for gender, religion, language, and ethnicity across elective and nonelective offices, across democracies and autocracies, across rich and poor countries, across different regions of the world, across time, and across various traits, including intersectional identities Tables (4–6). By contrast, where traits are not highly valued, for example, for youth and for those with little education, compositional effects are much weaker, as shown in Supplementary Materials I.

Several implications follow from these findings.

First, countries with greater heterogeneity achieve worse representation. Iceland achieves higher levels of representation than India, to take two extreme cases.

Second, dimensions of identity exhibiting greater heterogeneity achieve worse representation than dimensions exhibiting greater homogeneity. In the United States, for example, the largest ethnic group (white) encompasses 62% of the population while the largest linguistic group (English) encompasses 88% of the population. Predictably, linguistic groups are better represented than ethnic groups.

Third, the entropy effect means that it is more difficult to achieve representation with a small number of equal-sized groups than with a large number of groups among which one predominates. Although Bahrain and China have a similar number of ethnic groups, in Bahrain they are roughly equal in size, while in China the Han compose over 90%. Predictably, China does a better job of representing ethnic diversity than Bahrain.

Fourth, whenever identities are renegotiated, the way in which categories are redefined has important repercussions for the degree to which traits are represented. Since the general trend seems to be toward greater differentiation, we should anticipate that acts of reclassification (e.g., by the decennial U.S. Census) will generate less faithful representation overall. Likewise, refashioning identity through the intersection of orthogonal categories (intersectionality) should also weaken the fit between group characteristics and leader characteristics.

Any multiplication of categories complicates the coordination challenge inherent in descriptive representation. Of course, this does not mean it is wrong for people to adopt more specific identities—to consider

---

[8] See, for example, Dancygier et al. (2015) and Reingold, Haynie, and Widner (2020).

themselves Chinese-American rather than Asian-American, for example. It simply means that as categories become more nuanced, representational demands placed upon the political system grow. Ceteris paribus, one can expect greater shortfalls. This is the tragic irony of identity politics: as identity becomes more differentiated—and, arguably, truer to lived experience—it becomes harder to represent politically.

Institutional fixes that promise to enhance descriptive representation are worth pondering. However, one must also ponder their impact on other political objectives. Tradeoffs are to be expected.

Our simulations demonstrate that random draws from the population do a better job of achieving descriptive representation than do existing institutions. In this light, one might consider the virtues of that ancient method of selection known as sortition or lot (Delannoi and Dowlen 2016), sometimes adopted for deliberative assemblies in the present era (Fishkin 2018). Of course, these methods of governance are not without their difficulties (Mansbridge 2010).

As a second example, let us turn to the matter of body size. We have shown that larger bodies are generally more representative, suggesting that descriptive representation may be improved by increasing the size of political bodies. Executives, cabinets, supreme courts, and legislatures can all be enlarged by statute or constitutional reform, and small political parties can be eliminated by imposing thresholds on party representation.

However, increasing the size of political bodies is not costless. The suppression of small parties would presumably reduce ideological diversity in a polity.[9] Increasing the size of a legislature would reduce its cohesion or reduce the power of backbenchers. Increasing the size of an executive and insisting that it operate in a collective fashion may reduce its capacity to perform executive functions. For any political body, enlargement presumably render its operation less efficient as there are more voices to be heard.[10]

In this light, failures of representation are not simply failures of will. They are also products of a formidable coordination problem stemming from the categorical and particularistic nature of descriptive representation. Leaders can stand for lots of things, and these things can change over time, easing the task of substantive representation. But they can *be* only a few things and these identities must remain fairly stable over time lest leaders lose credibility in the eyes of constituents. There is no easy way to solve this coordination problem because identity is complex and leadership space is limited, by design.

---

[9] Eliminating small parties would not have a great impact on descriptive representation in legislatures, according to our data.
[10] These voices may be muffled by limiting access to the podium and by other centralizing initiatives, as is common in large legislatures. However, if backbenchers have little power this may undermine the goals envisioned by descriptive representation.

SUPPLEMENTARY MATERIAL

To view supplementary material for this article, please visit https://doi.org/10.1017/S0003055423000680.

DATA AVAILABILITY STATEMENT

Research documentation and data that support the findings of this study are openly available at the American Political Science Review Dataverse: https://doi.org/10.7910/DVN/BIQZNT.

CONFLICT OF INTERESTS

The authors declare no ethical issues or conflicts of interests in this research.

ETHICAL STANDARDS

The authors affirm this research did not involve human subjects.

REFERENCES

Anderson, Benedict. 2006. *Imagined Communities: Reflections on the Origin and Spread of Nationalism*. London: Verso Books.
Benhabib, Seyla. 1996. *Democracy and Difference: Contesting the Boundaries of the Political*. Princeton, NJ: Princeton University Press.

Bingle, Benjamin. 2016. "A Matter of Size: Examining Representation and Responsiveness in State Legislatures and City Councils." Phd diss. Northern Illinois University.

Bird, Karen, Thomas Saalfeld, and Andreas M. Wüst. 2011. *The Political Representation of Immigrants and Minorities: Voters, Parties and Parliaments in Liberal Democracies*. New York: Routledge.

Bullock, Charles S. III, and Susan A. MacManus. 1991. "Municipal Electoral Structure and the Election of Councilwomen." *Journal of Politics* 53 (1): 75–89.

Celis, Karen, and Silvia Erzeel. 2017. "The Complementarity Advantage: Parties, Representativeness and Newcomers' Access to Power." *Parliamentary Affairs* 70 (1): 43–61.

Celis, Karen, and Silvia Erzeel. 2020. "Gender Equality." In *The Oxford Handbook of Political Representation in Liberal Democracies*, eds. Robert Rohrschneider and Jacques Thomassen, 192–210. Oxford: Oxford University Press.

Chandra, Kanchan. 2006. "What is Ethnic Identity and Does it Matter?" *Annual Review of Political Science* 9: 397–424.

Collins, Patricia Hill, and Sirma Bilge. 2020. *Intersectionality*. Hoboken, NJ: John Wiley & Sons.

Dancygier, Rafaela M., Karl-Oskar Lindgren, Sven Oskarsson, and Kåre Vernby. 2015. "Why are Immigrants Underrepresented in Politics? Evidence from Sweden." *American Political Science Review* 109 (4): 703–24.

De Marchi, Scott. 2005. *Computational and Mathematical Modeling in the Social Sciences*. Cambridge: Cambridge University Press.

Delannoi, Gil, and Oliver Dowlen. 2016. *Sortition: Theory and Practice*, Vol. 3. Luton: Andrews UK Limited.

Escobar-Lemmon, Maria C., and Michelle M. Taylor-Robinson. 2014. *Representation: The Case of Women*. Oxford: Oxford University Press.

Fisher, Justin, Edward Fieldhouse, Mark N. Franklin, Rachel Gibson, Marta Cantijoch, and Christopher Wlezien. 2018. *The Routledge Handbook of Elections, Voting Behavior and Public Opinion*. London: Routledge.

Fishkin, James S. 2018. "Random Assemblies for Lawmaking? Prospects and Limits." *Politics & Society* 46 (3): 359–79.

Gerring, John, Erzen Oncel, Kevin Morrison, and Daniel Pemstein. 2019. "Who Rules the World? A Portrait of the Global Leadership Class." *Perspectives on Politics* 17 (4): 1079–97.

Gerring, John, Connor T, Jerzak, and Erzen, Öncel. 2023. "The Composition of Descriptive Representation." Harvard Dataverse. Dataset. https://doi.org/10.7910/DVN/BIQZNT.

Goodin, Robert E. 2004. "Representing Diversity." *British Journal of Political Science* 34 (3): 453–68.

Goodyear-Grant, Elizabeth, and Erin Tolley. 2019. "Voting for One's Own: Racial Group Identification and Candidate Preferences." *Politics, Groups, and Identities* 7 (1): 131–47.

Homola, Jonathan. 2019. "Are Parties Equally Responsive to Women and Men?" *British Journal of Political Science* 49 (3): 957–75.

Htun, Mala. 2016. *Inclusion without Representation in Latin America: Gender Quotas and Ethnic Reservations*. Cambridge: Cambridge University Press.

Hughes, Melanie M. 2013. "Diversity in National Legislatures around the World." *Sociology Compass* 7 (1): 23–33.

Kellogg, Leander D., Al G. Gourrier, E. Lee Bernick, and Katheryn Brekken. 2019. "County Governing Boards: Where Are All the Women?". *Politics, Groups, and Identities* 7 (1): 39–51.

Kjaer, Ulrik, Kelly Dittmar, and Susan J. Carroll. 2018. "Council Size Matters: Filling Blanks in Women's Municipal Representation in New Jersey." *State and Local Government Review* 50 (4): 215–29.

Kjaer, Ulrik, and Jorgen Elklit. 2014. "The Impact of Assembly Size on Representativeness." *Journal of Legislative Studies* 20 (2): 156–73.

Krook, Mona Lena, and Diana Z. O'Brien. 2012. "All the President's Men? The Appointment of Female Cabinet Ministers Worldwide." *Journal of Politics* 74 (3): 840–55.

Lemi, Danielle Casarez. 2021. "Do Voters Prefer Just Any Descriptive Representative? The Case of Multiracial Candidates." *Perspectives on Politics* 19 (4): 1061–81.

Lijphart, Arend. 2004. "Constitutional Design for Divided Societies." *Journal of Democracy* 15 (2): 96–109.

Liu, Amy H. 2011. "Linguistic Effects of Political Institutions." *Journal of Politics* 73 (1): 125–39.

Lublin, David. 2014. *Minority Rules: Electoral Systems, Decentralization, and Ethnoregional Party Success*. Oxford: Oxford University Press.

Lucardi, Adrián, and Juan Pablo Micozzi. 2022. "District Magnitude and Female Representation: Evidence from Argentina and Latin America." *American Journal of Political Science* 66 (2): 318–36.

Mackay, Fiona. 2008. ""Thick" Conceptions of Substantive Representation: Women, Gender and Political Institutions." *Representation* 44 (2): 125–39.

Mansbridge, Jane. 1999. "Should Blacks Represent Blacks and Women Represent Women? A Contingent 'Yes'." *Journal of Politics* 61 (3): 628–57.

Mansbridge, Jane. 2010. "Deliberative Polling as the Gold Standard." *Good Society* 19 (1): 55–62.

Matland, Richard E. 1998. "Women's Representation in National Legislatures: Developed and Developing Countries." *Legislative Studies Quarterly* 23(1): 109–25.

Matland, Richard E., and Michelle M. Taylor. 1997. "Electoral System Effects on Women's Representation: Theoretical Arguments and Evidence from Costa Rica." *Comparative Political Studies* 30 (2): 186–210.

Michalopoulos, Stelios. 2012. "The Origins of Ethnolinguistic Diversity." *American Economic Review* 102 (4): 1508–39.

Moser, Robert G. 2008. "Electoral Systems and the Representation of Ethnic Minorities: Evidence from Russia." *Comparative Politics* 40 (3): 273–92.

Oakes, Anne, and Elizabeth Almquist. 1993. "Women in National Legislatures: A Cross-National Test of Macrostructural Gender Theories." *Population Research Policy and Review* 12(1): 71–81.

Paxton, Pamela, Melanie M. Hughes, and Tiffany D. Barnes. 2020. *Women, Politics, and Power: A Global Perspective*, 4th ed. Lanham, MD: Rowman & Littlefield Press.

Pitkin, Hanna F. 1967. *The Concept of Representation*. Berkeley: University of California Press.

Rehfeld, Andrew. 2006. "Towards a General Theory of Political Representation." *Journal of Politics* 68 (1): 1–21.

Reingold, Beth, Kerry L Haynie, and Kirsten Widner. 2020. *Race, Gender, and Political Representation: Toward a More Intersectional Approach*. Oxford: Oxford University Press.

Rose, Richard. 1984. "Electoral Systems: A Question of Degree or of Principle?" In *Choosing an Electoral System. Issues and Alternatives*, eds. Arend Lijphart and Bernard Grofman, 73–81. New York: Praeger.

Ruedin, Didier. 2009. "Ethnic Group Representation: A Cross-National Comparison." *Journal of Legislative Studies* 15 (4): 335–54.

Ruedin, Didier. 2010. "The Relationship between Levels of Gender and Ethnic Group Representation." *Studies in Ethnicity and Nationalism* 10 (1): 92–106.

Ruedin, Didier. 2013. *Why Aren't They There?: The Political Representation of Women, Ethnic Groups and Issue Positions in Legislatures*. Colchester, UK: ECPR Press.

Ruedin, Didier. 2020. "Regional and Ethnic Minorities." In *The Oxford Handbook of Political Representation in Liberal Democracies*, eds. Robert Rohrschneider and Jacques Thomassen, 211–27. Oxford: Oxford University Press.

Skaaning, Svend-Erik, John Gerring, and Henrikas Bartusevicius. 2015. "A Lexical Index of Electoral Democracy." *Comparative Political Studies* 48 (12): 1491–525.

Stekhoven, Daniel J. 2015. "MissForest: Nonparametric Missing Value Imputation Using Random Forest." *SAO/NASA Astrophysics Data System (ADS),* Smithsonian Astrophysical Observatory: Cambridge, MA.

Stockemer, Daniel, and Aksel Sundström. 2018. "Age Representation in Parliaments: Can Institutions Pave the Way for the Young?" *European Political Science Review* 10 (3): 467–90.

Wängnerud, Lena. 2009. "Women in Parliaments: Descriptive and Substantive Representation." *Annual Review of Political Science* 12: 51–6.

Weldon, S. Laurel. 2008. "Intersectionality." In *Politics, Gender, and Concepts*, eds. Gary Goertz and Amy G. Mazur, 193–218. Cambridge: Cambridge University Press.