Descriptive Representation in Politics: A Measurement Proposal and Application for Brazil*

Sergio Firpo† Michael França‡ Leila Pereira § Alysson Portella¶

Rafael Tavares||

This version: March 2, 2023
(click here for the latest version)

Abstract

This paper introduces a new measure of descriptive representation in politics. It improves upon simple measures such as the share or the number of elected minorities by also taking into account the minority size among the represented. Moreover, contrary to disproportionality measures that are sometimes employed in the context of descriptive representation, our measure attains the upper bound in more usual situations. Based on the statistical properties of our index, we derive hypothesis tests that can be used to assess the likelihood of a representative body's observed composition being the result of an unbiased process. To illustrate its use, we apply the proposed index to evaluate the racial imbalances in the 2014 and 2018 Brazilian National and State legislative elections. This exercise shows that the inclusion of Non-whites in institutional politics has been partial and restricted only to candidacies. The imbalance among elected representatives is considerably larger than among candidates and we can reject the hypothesis that it was generated by chance in the majority of Brazilian states. We also investigate differences between left-wing and right-wing parties, observing lower imbalances among candidates in the left, but no difference when we restrict the analysis to elected representatives.

Keywords: Descriptive representation, Elections, Index, Race
JEL Classifications: D63, D72, J15

*We thank Lucas Novaes for useful comments on a previous version of this manuscript. All remaining errors are exclusive ours. We also thank the Open Society Foundation for their financial support for this research.

†Insper Instituto de Ensino e Pesquisa. E-mail: firpo@insper.edu.br
‡Insper Instituto de Ensino e Pesquisa. E-mail: michaeltulioramos@gmail.com
§Insper Instituto de Ensino e Pesquisa. E-mail: leilaarocha@gmail.com
¶Insper Instituto de Ensino e Pesquisa. E-mail: alyssonlp1@insper.edu.br
||Universidade de São Paulo. E-mail: rafaelaatavares@gmail.com

1 Introduction

Pitkin (1967) lists four different views of representation: formal, symbolic, descriptive, and substantive. Descriptive representation refers to the similarity between the representatives and the represented with respect to ethnicity, gender, nationality, social class, or any other socially relevant attribute. Some authors have argued that descriptive representation is important in and of itself (Mansbridge, 1999; Phillips, 1995; Dovi, 2002). Others have studied the determinants and consequences of descriptive representation, focusing mostly on gender and ethnic imbalances (Griffin and Newman, 2007; Grose, 2011; Wängnerud, 2009; Stockemer, 2015; Hänni and Saalfeld, 2020). Common issues dealt by this literature include whether descriptive representation is tied to substantive representation (Schwindt-Bayer and Mishler, 2005; Wallace, 2014; Boas and Smith, 2019) and how it affects turnout (Whitby, 2007; Rocha et al., 2010; Fraga, 2016a,b).

Generally, studies adopt the minority share or their number among representatives or candidates as a proxy for representation (Pantoja and Segura, 2003; Whitby, 2007; Rocha et al., 2010). Nevertheless, these are only valid measures when the population composition is homogeneous across the units of analysis (e.g., districts or countries). Using such measures in other contexts can potentially mislead empirical analyses. For example, as Fraga (2016a) points out, the estimated effect of representation on turnout may be biased if districts with larger minority shares are also more likely to have minority candidates. Comparative studies dealing with the representation of ethnic minorities are particularly sensitive to these limitations.[1] Hence, developing a more general and precise measure of descriptive representation is a challenge that the literature still needs to overcome.

Some studies have already provided advancements in that direction by using measures of disproportionality that consider both the minorities' shares among representatives and their shares among the represented (Monroe, 1994; Taagepera and Grofman, 2003; Koppel and Diskin, 2009; Ruedin, 2009; Gerring et al., 2022). However, these measures are limited as well. For instance, Duncan and Duncan's index only reaches its lowest value in the absence of representation of any domestic groups whatsoever. To better clarify the issue, suppose a case of a highly imbalanced representation where men correspond to 50% of a population but the elected congress has no women among the elected officials. In this case, Duncan and Duncan's index (Duncan and Duncan,

[1]For instance, Schwindt-Bayer and Mishler (2005) chose to focus their cross-country study on women precisely because they "constitute approximately 50% of the population virtually everywhere" (p. 408). Ethnic minorities, in contrast, usually differ both in number of groups and their relative size.

1955) takes the value of 1/2, not the extreme value of 1, as one would expect in a case of such extreme imbalance. This index only attains its upper bound in the unlikely situation of complete dominance by foreign powers not represented in the society. This is also true for the Gallagher index (Gallagher, 1991), for example. As we show in our empirical application, these situations are far from trivial, even in a context in which the political minority comprises a considerable share of the population.

In this paper, we propose a new measure of descriptive representation that can deal with both of these problems: it can be applied to different contexts in which the share of minority groups vary and its range is designed in such a way that cases of extreme misrepresentation (dominance by one domestic group) result in the index reaching its largest value. The *Descriptive Representation Index (DRI)* takes as inputs only the share of the minority group in the represented population and among the representatives. Contrary to disproportionality indexes, DRI range lies between -1 and 1. It assumes the lowest (highest) value when the representatives belong exclusively to the majority (minority) group. In the case of perfect equality in descriptive representation, the index takes the value of zero. Nonetheless, by taking the DRI's absolute or squared value, one can easily standardize the index to vary between 0 and 1.

Apart from measuring bias in descriptive representation, the DRI can also be interpreted as the political elite's marginal cost of waiving power to minorities in exchange for more descriptive representation. When minorities' population share is low, it may be easier to improve descriptive representation. One extra chair in Parliament contributes more to improving the minority position when their size is small than when their size is large. In contrast, Duncan and Duncan's index considers that the marginal cost of waiving power is constant. It shows only how many chairs the elite would need to give up to obtain a more balanced representation.

We also build upon the statistical properties of our index to provide confidence bands around which imbalances in representation are likely the result of pure chance. Based on the binomial distribution, the minority share among the represented and the size of the representative body, we derive DRI's probability distribution under the null hypothesis of random sampling of representatives from the poll of represented. This probability distribution is then used to construct confidence intervals under the null hypothesis of no imbalance in representation of minorities. A similar procedure can be used to derive pseudo p-values under this same null hypothesis. This innovation allows researchers to verify whether the observed descriptive imbalance is statistically significant, in the sense of being statistically discernible from a pure random outcome.[2]

[2] It also connects with the idea of both composition and size of representative bodies being relevant

To illustrate the properties of our index, we propose an application to investigate imbalances of racial representation in Brazilian legislative elections. Once regarded as an example of "racial democracy", Brazil suffers from significant racial disparities in well-being (Telles, 2004). As some research has already demonstrated, racial inequalities are also present in political representation (Campos and Machado, 2015; Bueno and Dunning, 2017; Janusz, 2018). Apart from considerable inequalities in representation, Brazil is also characterized by large regional heterogeneity regarding the share of White and Non-white individuals. In the northeast region of Brazil, the state of Bahia is primarily comprised of Non-white individuals (only 22% of its population declares being White). The opposite occurs in the State of Santa Catarina, in the southern region of Brazil, where 84% of the population declares being White. This provides the perfect setting to apply our index. We evaluate imbalances in the representation of Non-whites among both candidates and elected representatives in the National and State legislative elections of 2014 and 2018.

The results show that there are large and statistically significant deficits in racial descriptive representation in Brazil. Differences in the representation of racial minorities are relatively small when we consider the pool of candidates, even though they favor Whites. At the national level, the index takes a value of around -0.20 and -0.30, but there is considerable variation across Brazilian States, with some displaying racial balance. However, the level of descriptive imbalance increases considerably when we restrict the analysis to the pool of elected representatives. At the national level, DRI is approximately -0.6, reaching -1 in some States, representing a situation of full White political dominance. In particular, four out of 27 Brazilian States did not elect any Non-White representative for the National Congress in 2014.

We also look at how descriptive representation varies based on party ideology. We find that although left-wing parties tend to have a more balanced racial descriptive representation when we consider candidates, this advantage vanishes when we look only at elected representatives.

Our contribution to the literature is threefold. First and foremost, we introduce a new way to assess descriptive representation that allows for statistical inference of potential imbalances. Our approach improves upon simple measures such as shares or the presence of minority representatives and candidates (Fraga, 2016a; Geese, 2022) by taking into account the share of the represented in the districts. It also improves upon disproportionality measures (Ruedin, 2009; Gerring et al., 2022) by allowing the index to achieve maximal value in more usual situations in contemporary politics.

in determining balance in the representation of minorities (Geese, 2022).

Our new indicator further contributes to the literature of descriptive representation measurement by including confidence intervals within which imbalances are likely the result of chance rather than actual disparities in representation.

Second, our empirical application to Brazilian politics complements the literature on descriptive representation in contemporary democracies, which have focused mostly on developed Western countries (Pantoja and Segura, 2003; Schwindt-Bayer and Mishler, 2005; Rocha et al., 2010; Wallace, 2014). We show that racial inequalities in political representation in Brazil are large, further debunking the myth of Brazil's "racial democracy".

Finally, we also contribute to the literature on descriptive representation across the ideological spectrum (Sobolewska, 2013; Zapata-Barrero, 2017; Farrer and Zingher, 2018). Our results indicate that although left-wing parties nominate more Non-white candidates, they are unable to translate this into better descriptive representation if we only consider elected representatives.

The rest of the paper proceeds as follows. Section 2 introduces the Descriptive Representation Index and discusses its properties. Then, we present the Brazilian setting and our data in Section 3. Results of our application are analyzed and discussed in Section 4. Section 5 concludes.

2 Measuring Descriptive Representation

2.1 Descriptive Representation Index (DRI)

In their influential work, Ransom and Sutch (2001) investigate the distribution of Black and White US workers across occupations in five Southern states after Emancipation. In order to assess how balanced racial participation was in each of these occupations, they derive an index from the binomial distribution.[3] This index varies from -1 to 1, with zero representing perfect equilibrium.

Ranson and Sutch's index can be expressed in the following way. Let b_j be the proportion of Black individuals in occupation j, and p be the share of Black individuals in the population of reference. The Ransom and Sutch's Racial Balance Index (RBI) is given by:

$$\mathrm{RBI}(b_j, p) = \left(\frac{b_j - p}{p}\right)\left(\frac{p}{1-p}\right)^{b_j} \tag{1}$$

[3]A full derivation of this index can be found in Appendix A.

RBI takes the value of 0 when the share of Black workers in occupation j is the same as in the reference population, that is, when $b_j = p$. When there are no Black workers in a given occupation, i.e., when $b_j = 0$, the index takes the value of -1, and when there are only Black workers in an occupation, with $b_j = 1$, the index takes the value of 1. Apart from Ransom and Sutch (2001), this index was also used to study occupational segregation in Brazil after slavery (Rodrigues, 2021) and contemporarily (Firpo et al., 2020).

It should be noted, however, that in some cases the index lies out of the range of -1 and 1, the alleged extreme values of the RBI. This can happen when the share of Black workers in a reference population is smaller than 21% or above 79%. For instance, when $p = 0.9$ and $b_j = 0.1$, the RBI takes the value of -1.11.

To correct this flaw, we propose a modification to Ransom and Sutch (2001) original index. For a detailed description on how the proposed index is derived, we refer readers to Appendix A. The updated index used in this paper, which we call *Descriptive Representation Index* (DRI), is defined as following:

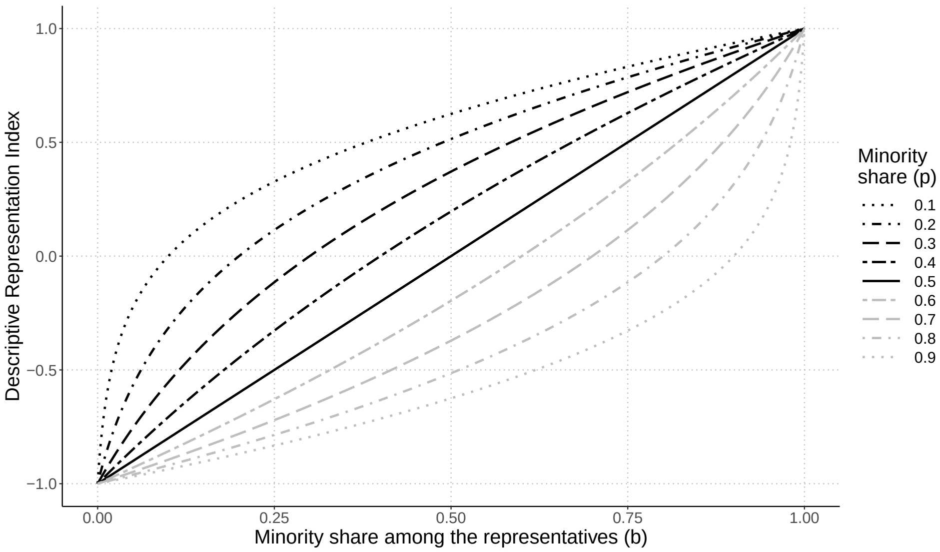
$$\mathrm{DRI}(b_j, p) = \frac{b_j - p}{\sqrt{(b_j - p)^2 + b_j(1 - b_j)}} \tag{2}$$

As the original RBI, DRI also reaches -1 and 1 when b_j takes the values of zero or one, and the value of zero is only reached when b_j is equal to p. However, in contrast to RBI, DRI is never below -1 or above 1. For instance, in the case when $p = 0.9$ and $b_j = 0.1$, DRI takes the value of -0.94. In practice, the values given by the RBI and DRI are similar, apart from extreme cases.

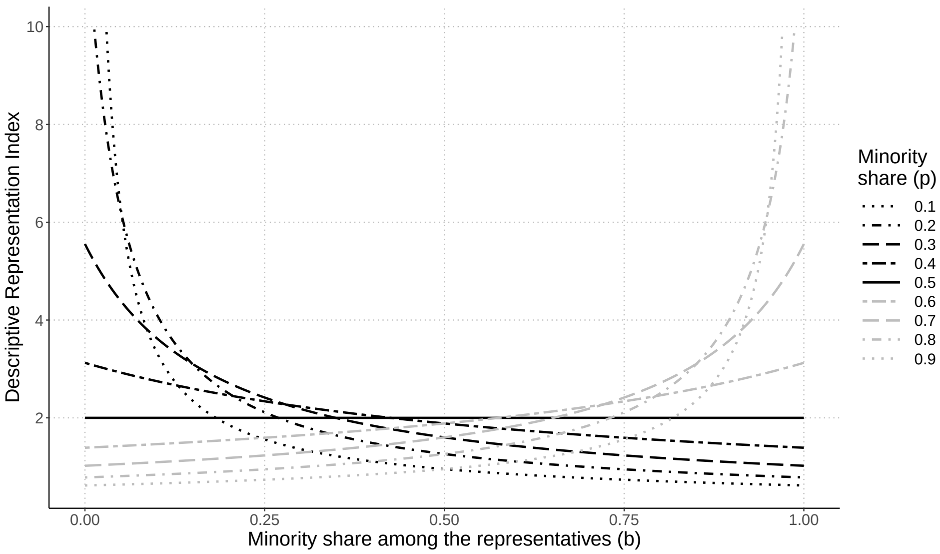
2.2 DRI properties

Figure 1 depicts some of DRI's properties. Panel (a) shows how the DRI changes depending on the values of the minority share among the population (p) across the whole domain of the minority share among the represented (b). When $p = 0.5$, DRI is a linear function of b. In all other cases, the index is strictly increasing in b, although the slope changes depending on b. For instance, when the minority share is small both among the population and among the representatives (i.e., when b and p are small), DRI's slope is steeper. For $p < 0.5$, as b increases, the rate of change decreases. For $p > 0.5$ the opposite happens: the derivative is increasing in b. This pattern can be observed in Panel (b), where we plot the values of DRI's derivative with respect to b for different values of p.

Figure 1: Descriptive Representation Index



(a) DRI as function of b and p



(b) DRI derivative with respect to b, as function of b and p

Notes: Plots use illustrative data to analyze the behavior of DRI depending on p (i.e., the minority share in the population) and b (i.e., the minority share in the representatives). Plot (a) shows the actual DRI, while Plot (b) presents DRI's derivative with respect to b. Black lines concerns values of the DRI when $p \leq 0.5$ and gray lines when $p > 0.5$.
Source: by the authors.

How can these differences in DRI's behavior across the minority share among the represented (p) be interpreted? In this sense, the inverse of the derivative can be viewed as the political cost of rebalancing representation: when the minority share in the whole population is small, moving towards equality is relatively easy, since a small increase in minority representation moves DRI faster towards zero. However, when the (political) minority comprises a large fraction of the represented, the efforts required to achieve equality are substantial.

Consider, for instance, the case of the United States and South Africa. In the first, Blacks comprise nearly 12% of the population, while in the second Blacks comprise around 3/4 of the population.[4] Therefore, moving from 5% Black representation to 10% would represent a large advancement for Blacks in the US, but would barely affect their representation in South Africa. DRI captures this difference, while a simple difference in percentage points would regard both cases as symmetric.

Given the widespread use of disproportionality indexes to measure descriptive representation, we assess DRI using a set of properties employed by Koppel and Diskin (2009) to evaluate the performance of these different measures. These measures are designed to assess how imbalanced are parties' vote shares and their share of seats in legislatures. They can be applied to contexts with any number of parties or social groups, contrary to DRI which is restricted to only two social groups. Among these measures are the Duncan & Duncan index (Duncan and Duncan, 1955) and the Gallagher measure (Gallagher, 1991), which in the case of only two parties both simplify to $|b - p|$ in our notation.[5] They list a total of eight properties, which translated to our context can be summarized in the following way:

1. *Continuity*: The index is continuous in both inputs (b and p);
2. *Irrelevance of virtual parties*: Adding an extra party (or minority group) with no presence among represented and representatives should not alter the value of the index;
3. *Symmetry*: Switching the minorities share in the population (p) and the minorities share among the representatives (b) yields the same outcome;
4. *Insensitivity to order*: Replacing minority shares, p and b, with the majority shares, $(1 - p)$ and $(1 - b)$, yields the same outcome;
5. *Scale invariance*: Changing the number of representatives and holding the minority participation constant should not change the index;
6. *Specified limits*: The index ranges between 0 and 1;
7. *Dalton's principle of transfers*: Changes in representation that bring b closer to p should reduce the absolute value of the index;

[4]For South Africa we exclude Coloured and other Non-white minorities.
[5]See Section 4.3.

8. *Optimality of equality*: The minimal absolute value for the index is obtained when the minority share in the electorate (p) is equal to its share among the representatives (b).

Of these properties, *irrelevance of virtual parties* cannot be applied in our context, as DRI is only designed for two groups. Regarding the other properties, DRI satisfies all of them but *symmetry* and *specified limits*. We do not provide formal proofs for the properties that DRI satisfies, as these can be easily verified using numeric examples or are implied by the fact that there are only two groups. In the context of DRI, *symmetry* would require that replacing the minority share among the represented with the minority share among the representatives should result in the same value for the index. Although this property is relevant in contexts discussed by Koppel and Diskin (2009), particularly for the case of volatility in party composition across elections, it is less clear whether this property is essential when looking at descriptive representation.[6] For instance, $DRI(b = 0.2, p = 0.3) = -0.24$, while $DRI(b = 0.3, p = 0.2) = 0.21$. It is not straightforward whether these (absolute) values should be the same.

For the case of *specified limits*, two considerations must be made. First, to be restricted to the $[0, 1]$ range, we must use the absolute value or the square of DRI.[7] This poses no complication. Second, DRI does not satisfy the principle of *orthogonality*, a subproperty of *specified limits* by which the index should attain only maximal value when all representatives are outsiders. However, we do not believe this property is relevant in the context of descriptive representation, particularly when we are interested in evaluating representation in contemporary democracies. Instead, our index achieves the maximal value when all representatives belong to only one group, which we believe is a more interesting comparison case.

2.3 When does DRI actually mean lack of representation?

Although DRI is capable of ranking any representation profile as being more or less imbalanced than some other, a related question is what values should we use to consider a situation imbalanced. In other words, how large should deviations from zero be so we can discard the possibility that it had just occurred by chance rather than by lack of representation? This is a relevant question because studies show that the size of political bodies and electorate composition are important determinants of proportionality in representation (Gerring et al., 2022).

[6] In the case of volatility across years, *symmetry* imposes that switching years would not change the result.

[7] DRI's absolute value or its square are also required to satisfy *insensitivity to order.*

To address this concern, we compute DRI's probability distribution, both considering a fixed number of individuals in a representative group and asymptotically, when the size of the representative body grows indefinitely. Details on the results can be found in Appendix B. The main assumption in deriving the statistical properties of DRI is that of no bias in representation. Therefore, we suppose that N representatives are randomly sampled from the population based on a binomial distribution, with the probability of drawing a minority candidate being equal to their share in the population, p.

For each pair (N, p), we derive DRI's probability distribution under the assumption of no bias in the sampling process. In Appendix B, we show that this distribution is centered around zero and that it becomes symmetric as N increases (Figure B.4). For small representative bodies and $p \neq 0.5$, the probability distribution will be skewed, with larger odds of imbalance towards negative values if $p < 0.5$ and the opposite if $p > 0.5$.

From DRI's probability distribution, we construct confidence intervals based on the following algorithm: i) We construct a window of length $h = 0.001$ around zero and compute the probability mass within it; ii) We continue increasing this window's length until the probability mass becomes larger than $1 - \alpha$, with α being some confidence level such as 10% or 5%. Figure B.5 shows how the confidence interval grows as the size of the representative body increases, for different minority shares among the represented. Because this distribution is based on finite values, it has many jumps. Figure B.5 also includes interpolated values of this confidence interval. It shows that for a small representative body, the confidence interval is large. For instance, when N is around 25, the confidence interval is close to ± 0.4 for every considered value of p. However, these confidence intervals become smaller as N increases: as N reaches 150, the confidence intervals lie in a 0.1 window.

Because the formula for the finite probability distribution includes factorials of N (see Equation B.16), it can only be computed for a limited representative body size, as the factorial of N tends quickly towards infinity. Therefore, we also rely on asymptotic theory to assess the behavior of DRI as $N \to \infty$. In Appendix B, we show that the distribution of $\sqrt{N}DRI$ converges to a Normal distribution with mean zero and standard deviation 1 under the assumption of random sampling of representatives from the population. This result can be used to derive asymptotic confidence intervals around zero using Equation B.22. This equation only takes as input the confidence level $1 - \alpha$ and the size of the representative body, N, because as N increases, the probability distribution becomes symmetrically centered around zero for any $0 < p < 1$.[8] For $p = 0.5$, both confidence intervals are similar even for a small N. For $p \neq 0.5$, the finite confidence interval is larger the further p is from 0.5. However, as N grows, the finite confidence interval always converges to the asymptotic confidence interval regardless the value of p.[9]

Finally, it is also possible to derive pseudo p-values for DRI using the asymptotic distribution. This is shown in Equation B.23. It gives us an indicator of how large the confidence level should be, given the observed value of DRI and the size of the representative body, so the hypothesis of no bias in the sampling process could not be rejected. In the following empirical application, we report these p-values together with DRI's values.

3 Application: context, data and methodology

3.1 The Brazilian context

Legislative elections in Brazil At the federal level, the Brazilian Legislative Power is represented by the National Congress, which is a bicameral entity composed by the Federal Senate (the upper house) and the Chamber of Deputies (the lower house). At the State and Municipal levels, the Legislative Power is represented only by one legislative chamber. We restrict our application to the Chamber of Deputies and to the State Chambers. We exclude the Federal Senate because it is elected through a majority vote system, with only three representatives per State and the Federal District. Furthermore, we also exclude municipal councils for two reasons: i) they tend to be small for most of Brazilian municipalities; and ii) we only have estimates of minority share among the represented per municipality based on the Brazilian Census, which was last carried out in 2010.

Representatives for the Chamber of Deputies and State Chambers are elected every four years for four-year terms, with the whole chambers renewed after each term. State and Federal Deputies can be reelected indefinitely and elections for these positions are held using a proportional system within each State and the Federal District. Each State and Federal District have a given number of seats that are proportional to their population size. These seats are then distributed between parties based on their

[8] See Figure B.4 in the Appendix.

[9] In Figure B.6, we plot the evolution of the asymptotic confidence interval as well as the finite confidence interval for a few values of p.

share of votes within that particular Federative Entity, each of which comprises a separate electoral district. Within parties, the order of candidates is determined based on the number of votes each one has received, in an open-list format.[10] For State Chambers, the seats are allocated based on the same proportionality rule, with the difference that there is only one district for each chamber. The Chamber of Deputies has a total of 513 seats, while the number of seats in the State Chambers varies according to population size, ranging from 24 to 94 seats.

Racial representation in Brazil The racial classification system used by Brazil's Statistics Bureau (*Instituto Brasileiro de Geografia e Estatística* - IBGE) consists of five categories that are generally self-identified by interviewees: White, Black, Brown, Asian, and Indigenous.[11] Black and Brown (mixed-race) are commonly aggregated into a Non-white group. From the Brazilian Census 2010, 47.7% of Brazilians identify as White, while Black and Browns represent 7.6% and 43.1%, thus being the majority of the population (50,7%). Asian and Indigenous individuals comprise less than 2% of the population and because of it, we exclude them from the analysis. These numbers, however, hide large heterogeneity across Brazilian regions. In Southern Brazil, the share of Non-white is much smaller than in the Northeast, for instance.

Despite comprising the demographic majority, Non-whites in Brazil are a clear political minority. Even though Brazil has been regarded as a "racial democracy" until recently and many still deny the existence of racial discrimination in this country, academic research shows unambiguously that Non-whites are excluded from access to economic resources and political power in Brazil (Telles, 2004). For instance, in the Constitutional Assembly held in 1988, 100 years after the abolition of slavery in Brazil, only 11 representatives were classified as Non-white out of 559 assembly members (Pires, 2013). In contrast, 47.4% of the population declared themselves as Non-white in the 1991 Demographic Census.

Studies show that Non-white candidates are less likely to be elected in Brazilian elections. These differences seem to stem from the odds of being elected, or the ability of Non-white candidates to receive a large share of votes, as the evidence suggests no underrepresentation among candidates (Campos and Machado, 2015, 2017; Bueno and Dunning, 2017). The literature shows that Non-white candidates have less access to campaign resources that include money, time on TV propaganda and ballot num-

[10]In practice, parties can form coalitions, in which case the votes of each party in the coalition are aggregated to determine the number of seats they have, and the list is ordered across all candidates in these parties.

[11]Respectively, in Portuguese, *branco*, *preto*, *pardo*, *amarelo*, and *indígena*.

bers easier to memorize (Bueno and Dunning, 2017; Wylie, 2020; Janusz, 2021; Janusz and Sells, 2021). Whether black candidates are discriminated by the electorate remains an open question (Aguilar et al., 2015; Bueno and Dunning, 2017; Janusz, 2018, 2021). Finally, studies also reveal that racial descriptive representation is possibly connected with substantive representation. Janusz and Campos (2021) shows that, in Rio de Janeiro municipal elections, black candidates are ten times more likely than white candidates to discuss topics about racial inequality in their political advertisements on TV.

In this paper, we focus only on measuring how imbalanced is racial descriptive representation in Brazil. We do not delve into its causes or consequences. However, since we investigate racial inequalities in representation both at the candidate and elected representatives levels, we also contribute to the literature that tries to assess whether a shortage of Non-white candidates might be an important factor explaining the small presence of Non-whites in the legislative bodies in Brazil.

We also investigate racial imbalances across ideological groups, by comparing the descriptive representation at the national level between parties in the left, center and right of the political spectrum, as classified in Brazil by Power and Rodrigues-Silveira (2018). This application not only sheds light on the relevant question of whether some ideological group is more inclined to represent ethnic minorities than others,[12] but also illustrates how DRI can be applied when aggregating results across geographical areas with different racial profile.

3.2 Data and Methodology

The data used in this study comes from three different sources. First, the Brazilian Electoral Court (*Tribunal Superior Eleitoral* – TSE) provides detailed information on elections. More specifically, microdata on Brazilian elections is made available at TSE Electoral Data Repository, and contains socioeconomic information on candidates (e.g., gender, age, occupation, party affiliation, and declared assets) and their electoral outcomes (e.g., total votes and electoral outcomes). Since 2014, the TSE has also released the self-declared race of candidates. Hence, due to data constraints, this analysis is carried out only for 2014 and 2018.[13]

[12] See, for instance, Sobolewska (2013); Zapata-Barrero (2017); Farrer and Zingher (2018).

[13] Following IBGE's classification system, candidates declare themselves as White (*branco*), Brown (*pardo*), Black (*preto*), Asian (*yellow*), or Indigenous (*indígena*). We merge Browns and Blacks into a Non-white group and exclude Asian and Indigenous from the analysis, as they comprise a small share of the population and politicians. This procedure is customary in the context of Brazilian race studies (Telles, 2004)

Second, because TSE does not provide information on the electorate's race, we use data from the Continuous National Household Sample Survey (*Pesquisa Nacional Por Amostra de Domicílios Contínua – Continuous PNAD*) to compute the Non-white participation in the electorate.[14] We estimate this value using the proportion of Non-white individuals aged 18 years old or more in each state. We use this age threshold as this is when voting becomes mandatory in Brazil. Although voting ceases to be mandatory at age 70, we do not cap our estimate of the minority voting at this age.

Third, the ideological positioning of parties follows Power and Rodrigues-Silveira (2018) classification. In this study, the authors created an index ranging from -1 to $+1$, with more negative values representing parties to the left and more positive values representing parties to the right. Using this index, we separated Brazilian parties into three groups of the same size according to the scale of the index, forming the Left, Center, and Right party groups.

In this application, we are interested in comparing the share of Non-white individuals running as candidates (or elected) for some position j — federal or state representative — in state s. Therefore, $b_{j,s}$ will denote the share of Non-white candidates or elected representatives in the position j in state s. Similarly, p_s represents the proportion of Non-white voters in state s. We compute DRI for each state and position separately.

When computing racial descriptive representation across ideological groups at the national level, we aggregate results across several states for each ideological group. However, this poses a challenge to our approach, as different parties are not equally represented across all Brazilian regions. To deal with this problem, the proportion of Non-white voters used as reference, p, is the average of each state, p_j, weighted by the share of the candidates in each ideological group within each of these states.

For instance, we want to know what is the descriptive representation index among candidates in left-wing parties across the whole country. As an example, suppose left-wing parties have a total of 50 candidates in three states, 10 in the first, 15 in the second, and 25 in the third. In addition, suppose that these states have a proportion of Non-white voters of respectively 80%, 50%, e 30%. Therefore, the proportion of Non-white voters used as a reference for measuring DRI is:

$$\left(\frac{10}{50} \times \frac{80}{100}\right) + \left(\frac{15}{50} \times \frac{50}{100}\right) + \left(\frac{25}{50} \times \frac{30}{100}\right) = \frac{46}{100}$$

[14] These surveys use the same classification system as TSE and we proceed the same way as before with respect to individuals' racial classification.

4 Application: evaluating descriptive representation in Brazil

In this section, we present the results on the racial descriptive representation in Brazilian Legislative elections. First, we present results regarding the balance in representation among candidates running for seats both at the National and State Chambers, in 2014 and 2018. We include both levels of representation to investigate whether racial imbalances are higher for more important offices, with a lower number of available seats. Moreover, measuring descriptive representation in two distinct years allow us to assess the stability of racial (im)balances in these positions.

Next, we evaluate racial balances in representation among elected representatives, again in both representation levels and both elections. This analysis allows us to assess to what extent imbalances in representation are explained by a shortage of candidates or whether it arises due to other reasons, such as lower resources for campaigns, lack of support from parties, or electoral preferences (Aguilar et al., 2015; Campos and Machado, 2015; Bueno and Dunning, 2017; Wylie, 2020; Janusz and Sells, 2021; Janusz, 2018). [15] In addition, contrasting candidates and elected officials also help us illustrate the statistical properties of our index. In particular, we show that, because the number of elected officials is much smaller than candidates, it is much harder to be positive when asserting if a given imbalance is just the result of chance rather than an actual racial bias of representation.

This empirical application is also useful for contrasting the results obtained by using the DRI or the other disproportionality measures. We highlight the cases where DRI and other disproportionality measures diverge and explain this divergence based on the properties discussed in Section 2.2.

Finally, we use our index to assess how racial representation differs across ideological groups. Whether left or right-wing parties have advantages in terms of representation remains an open question in the literature (Sobolewska, 2013; Zapata-Barrero, 2017; Farrer and Zingher, 2018). Our exercises show that the DRI can be used to study such issues in contexts where both the minority shares and the presence of political parties differ across regions.

[15] We emphasize that assessing the relative importance of these factors is beyond the scope of this study.

4.1 Racial representation among candidates

Figure 2 presents the results on descriptive representation among candidates across all Brazilian states.[16] The top two maps show the results concerning candidates running for office in the National Chamber of Deputies. The left and right-sided panels comprise results for the 2014 and 2018 elections, respectively. Considering all Brazilian candidates, the overall DRI was -0.26 in 2014 and -0.28 in 2018, meaning the Non-white population is underrepresented among candidates. There is some heterogeneity in the level of imbalance across Brazilian states. In 2014, the State of Tocantins (TO) presented the most significant imbalance, with a DRI of -0.62. At the same time, the place where DRI value was closest to zero DRI was the State of Santa Catarina (SC), with a DRI of -0.03, reflecting a statistically balanced representation. In 2018, even though DRI changes across individual states, the overall situation is similar to 2014. The State of Roraima (RR) has DRI closest to equality, slightly favoring Non-white candidates (0.01), while the highest difference in representation occurs in Alagoas (AL), with a DRI of -0.55. For the absolute majority of states, we reject the hypothesis that the observed imbalances in racial representation were simply due to chance.

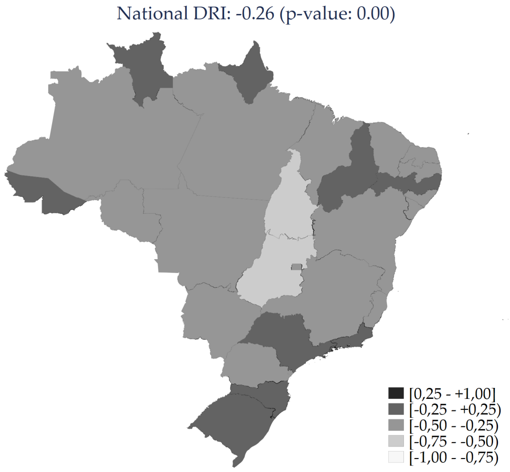
Conclusions are similar when we consider State Assemblies rather than the National Chamber of Deputies. These results are presented in the bottom panels of Figure 2. The overall DRI, aggregating all Brazilian candidates, indicates that the lack of representation is marginally smaller among those running for seats in the State Chambers than those running for seats in the Chamber of Deputies, respectively -0.23 and -0.19 in 2014 and 2018. The range in which DRI varies is similar, going from small positive values (indicating a slight over-representation of Non-whites) to values reaching almost -0.6. The conclusions remain true for both elections. Again, for the absolute majority of Brazilian states, we reject the hypothesis that the observed imbalances arose by chance. One exception is the state of Acre (AC), in which we cannot reject the hypothesis that racial representation is balanced for both years at usual confidence levels.

4.2 Racial representation among elected representatives

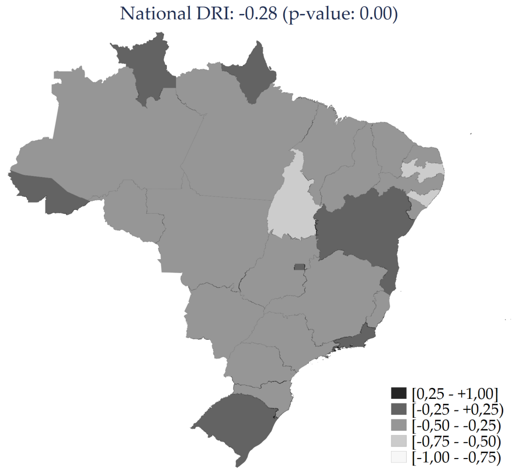
How does the results change when we restrict the analysis to those candidates that were in fact elected? Figure 3 shows that the lack of representation of Non-white

[16]The exact values, together with pseudo p-values, can be found in the Appendix C, Tables C.1 and C.2.

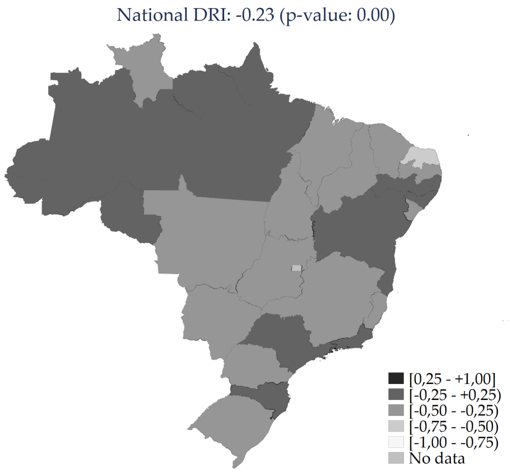
Figure 2: Candidates' DRI by State



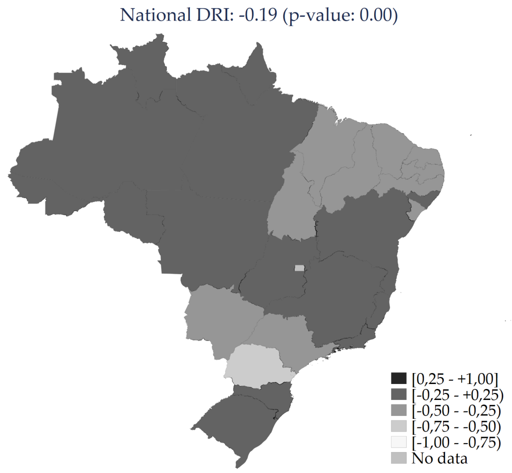
(a) Federal Congress Representatives - 2014



(b) Federal Congress Representatives - 2018



(c) State Assemblies - 2014



(d) State Assemblies - 2018

Notes: Maps present DRI calculated for the pool of candidates for each of the Brazilian states. Maps (a) and (b) show DRI for candidates to Federal Congress in the 2014 and 2018 elections, respectively. Maps (c) and (d) present DRI for candidates to State Assemblies.
Source: DRI is calculated using data from the Brazilian Electoral Court (*Tribunal Superior Eleitoral* – TSE) and from the Continuous National Household Sample Survey (*Pesquisa Nacional Por Amostra de Domicílios Contínua*– Continuous PNAD).

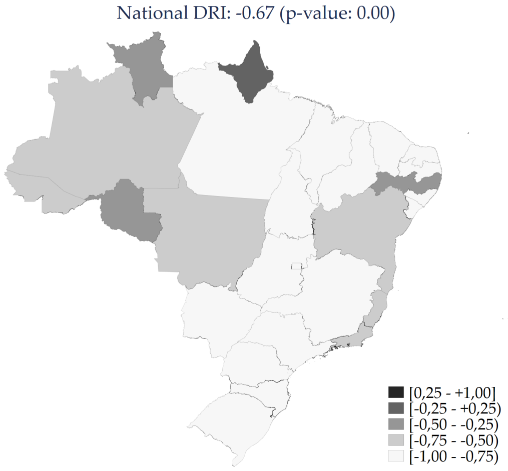
Brazilians is much larger among the elected representatives than among candidates.[17] At the national level, DRI among elected representatives in the Chamber of Deputies was -0.67 and -0.62 in 2014 and 2018, respectively. In contrast, the same values among candidates were -0.26 and -0.28.

Again, representation is very heterogeneous across Brazilian states. In 2014, the state of Amapá (AP) had the least imbalanced racial descriptive representation, with a DRI of -0.05. At the same time, the state of Roraima (Roraima, RO) had a much larger

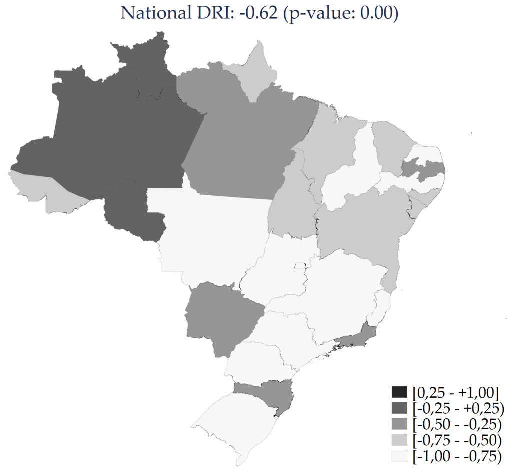
[17] Complete results, including pseudo p-values, are shown in Tables C.3 and C.4 in Appendix C.

DRI (-0.38) and ranked as the second least imbalanced in descriptive representation. Four states did not elect any Black representative in the 2014 elections, meaning their DRI was -1. In 2018, Roraima (RO) and Amazonas (AM) states had the smallest DRI: -0.12 and -0.13, respectively. Meanwhile, the states of Rio Grande do Sul (RS) and Rio Grande do Norte (RN) did not elect any Non-white representative.

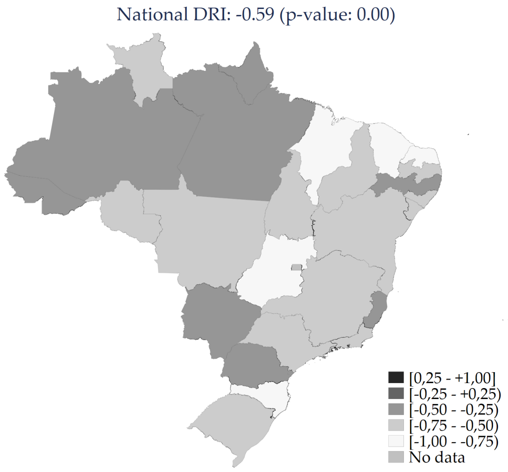
Figure 3: Elected Representatives' DRI by State



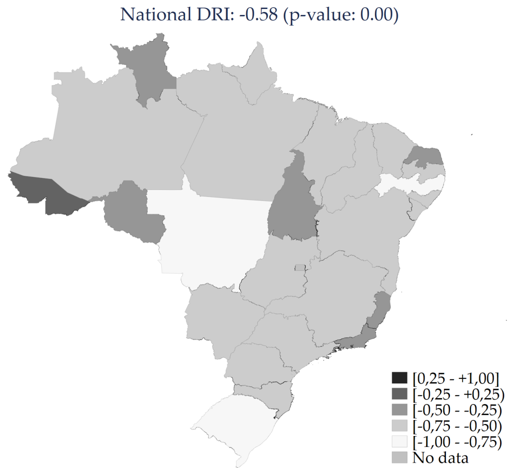
(a) Federal Congress Representatives- 2014



(b) Federal Congress Representatives - 2018



(c) State Assemblies - 2014



(d) State Assemblies - 2018

Notes: Maps present DRI calculated for the pool of elected representatives for each of the Brazilian states. Maps (a) and (b) show DRI for Federal Congress Representatives in the 2014 and 2018, respectively. Maps (c) and (d) present DRI for State Assemblies Representatives.
Source: DRI is calculated using data from the Brazilian Electoral Court (*Tribunal Superior Eleitoral* – TSE) and from the Continuous National Household Sample Survey (*Pesquisa Nacional Por Amostra de Domicílios Contínua* – Continuous PNAD).

Results for elected Federal representatives in 2018 (Table C.3) are particularly interesting to discuss some of DRI's properties. First, take the example of Tocantins (TO) and Mato Grosso do Sul (MS). Both states elected 38% of Non-white representatives. However, in absolute terms, TO's DRI is considerably larger than MS's (–0.64 vs. -0.32) since the Non-white group represents 77% of voters in the first state and only 54% in the latter. In contrast, the state of Rio de Janeiro (RJ), while having the same share of Non-white voters as MS, had a more negative DRI (-0.38) because of electing a slightly smaller share of Non-white representatives (35%).

The second property of DRI worth discussing is that it is not always the case that a more significant distance between the minority's share among representatives and voters ($b - p$) will imply a more imbalanced situation. For instance, for Santa Catarina (SC), this distance is 14 p.p. ($b = 6\%$ and $p = 20\%$), while for Rio de Janeiro (RJ), this distance is 19 p.p. ($b = 35\%$ and $p = 54\%$). Nonetheless, DRI is more imbalanced in SC (-0.48) than in RJ (-0.38). This occurs because the standardization based on the binomial distribution deems as more likely the second rather than the first scenario.

What is more interesting is that, considering the size of SC and RJ representative bodies ($N = 16$ and $N = 46$), the likelihood that these states' racial descriptive imbalances happened by chance (i.e., their pseudo p-values) are 0.05 and 0.01, respectively. Hence we cannot reject the hypothesis that racial imbalance in SC is random at a 95% confidence level. This points to the importance of considering the legislative bodies' size when investigating disparities in descriptive representation: in small legislatures, such as SC, the likelihood of observing random descriptive representation imbalance is higher than in larger legislatures, such as RJ. In fact, replacing one White representative with a Non-white one in SC would increase the share of Non-white representatives to 12% implying a considerable reduction of DRI from -0.48 to approximately -0.21.

Therefore, it is not surprising that we find more racially balanced states from a statistical perspective when analyzing the pool of elected representatives than when examining the pool of all candidates.[18] This occurs despite the levels of racial imbalance being much larger among elected representatives than among candidates.

Results for candidates that were elected to the State Chambers are shown in the bottom panels of Figure 3. The size of the disproportionality in racial descriptive representation is much larger among elected candidates than among candidates. Considering all elected State representatives in Brazil, the overall DRI was -0.59 in 2014 and -0.58 in 2018, much larger than -0.23 and -0.19 among candidates in those same years. There is a large variation in racial descriptive representation among states. The DRI that was closest to zero was found in the state of Acre (AC), both in 2014 and 2018,

[18]At a 95% confidence level, this happened in eleven states in 2018 for elected Federal representatives, but only 3 times among candidates in these elections.

reaching respectively -0.30 and -0.15. In 2014, the State of Santa Catarina (SC) did not elect any Black representative for its State Legislative Chamber, while the largest disproportionality is found in Pernambuco (PE) in 2018, with a DRI of -0.84.

Contrary to what happens with representatives at the Federal level, for the vast majority of State Chambers we can reject the hypothesis that racial imbalances occurred by chance. Even though the levels of imbalance in descriptive representation are similar between State Chambers and States' representation at the Federal level, differences in the size of these bodies produce this divergence in statistical significance. For instance, in 2018 only three State Chambers could not be rejected as balanced at the 5% confidence level, while for the Federal Chamber there were 11 states in this situation. This is the result of State Chambers being much larger than the number of state representatives elected to the Federal Deputies Chamber. The smallest number of Federal representatives per State is 8, while the smallest State Chambers have 24 members. This makes an important difference when evaluating whether imbalances are the result of simple chance rather than actual distortions in descriptive representation.

4.3 Comparison with other disproportionality measures

As we have mentioned before, some studies have employed disproportionality indexes to assess the level of descriptive representation in cross-country analyses (Ruedin, 2009; Gerring et al., 2022). Here, we compare DRI with Duncan & Duncan's (D) and the Gallagher (G) indexes (Koppel and Diskin, 2009). Let the vectors $\mathbf{x} = (x_0, x_1)$ and $\mathbf{y} = (y_0, y_1)$ contain the minority (subscript 0) and majority (subscript 1) shares among representatives and represented, respectively. In the special case with only two racial groups, $x_1 = 1 - x_0$ and $y_1 = 1 - y_0$. When this happens, both indexes simplify to the absolute difference between the minority share among the representatives and the minority share among the represented. Duncan & Duncan's index is given by:

$$D(\mathbf{x}, \mathbf{y}) = \Sigma \left| x_i - y_i \right| / 2 = |x_0 - y_0| \tag{3}$$

and Gallagher's index by:

$$G(\mathbf{x}, \mathbf{y}) = \left(1/2 \times \Sigma \left(x_i - y_i \right)^2 \right)^{1/2} = |x_0 - y_0| \tag{4}$$

Figure D.7, in Appendix D, plots the absolute value of DRI against the absolute difference between b and p, corresponding to both Duncan & Duncan's and Gallaher's indexes in the case of only two groups. The plot includes results for candidates and elected representatives, in all 27 Brazilian states, for the National Congress and State Assemblies. We observe a strong correlation between these measures of descriptive representation. This is no surprise, as all of them rely heavily on the difference between the proportion of elected representatives of each race and their share among voters. DRI, however, almost always takes values above the other two indexes.

In general, the difference between the indexes follows a somewhat concave pattern across the whole range of these indexes. But there are some cases in which the differences between them are large. The largest discrepancies are found when DRI is close to one but $|b - p|$ is low. The reason why this arises is twofold. First, in the extreme case of no minority representative ($b = x_0 = 0$), DRI takes an absolute value of 1 by definition, while the Duncan & Duncan and Gallagher's indexes take a value equal to the minority share among voters, $p = y_0$.

Second, small changes in b when it is already too low are weighted heavily by DRI when p is farther away from 0.5, while it behaves more linearly the closer it gets to this value, as can be seen in Figure 1. Therefore, cases in which b is very small and p is very large or very small will lead to larger discrepancies between DRI and the other two measures.

For most of situations, DRI will provide results similar to Duncan & Duncan's and Gallagher's indexes. How should we proceed when they do not? That is, which is best when they diverge? The answer to this question will depend on how the researcher weighs the different properties of these indexes. The situation when there is no minority representative is particularly relevant in this choice. DRI will always consider this case as an extreme lack of representation and assign to it the highest possible value, -1. Disproportionately indexes, however, will treat this situation as less problematic the smaller the minority share among the represented is. As we have discussed previously, this distinction between the indexes arises precisely from the principle of *orthogonality*, which we believe is not the best to characterize an index concerned with descriptive representation in contemporary societies.

At the same time, DRI will reduce fast as the minority share increases, especially for small p. This is due to the non-linear nature of DRI, which is exacerbated when the minority share is too small or too large. This lack of linearity differs from the other two indexes when there are only two groups, as they will change linearly with reductions in b irrespective of p. As we have argued, the lack of linearity in DRI can be interpreted as a "cost" to achieve equality of representation. When the minority share is small, modest concessions from the elites can drive representation towards balance. But when the minority is highly underrepresented and it comprises a large share of the population, the elite will have to renounce a lot of power to allow for some balance.

Finally, another difference between these indexes includes the property of *symmetry*, by which y and x should be interchangeable in the indexes. As we have argued, we do not believe this property is relevant in the context of descriptive representation, although it is certainly necessary in the case of the volatility of representation across time, for instance.

4.4 Differences in racial representation across ideological groups

As our last exercise in this empirical application, we investigate whether there are differences in racial descriptive representation across ideological groups. For instance, it has been argued that center-left parties are more likely to nominate ethnic minorities because these parties' platforms tend to attract more minority support (Zingher, 2014; Sobolewska, 2013) and that this effect should be stronger in districts with higher minority vote-shares (Farrer and Zingher, 2018). At the same time, it is possible that center-right political parties nominate minorities in districts where they are more likely to win (Sobolewska, 2013), or even that right-wing parties might use descriptive representation to attract votes from poorer electorate (Desai and Frey, 2021).

Table 1 summarizes the results for DRI contrasting parties in the Left, Center and Right of the political spectrum in Brazil, as classified by Power and Rodrigues-Silveira (2018), in the 2014 and 2018 elections, considering again both candidates and elected representatives. We observe large disparities in racial representation among candidates and elected representatives for all ideological groups. Again, in all instances inequalities are larger among those elected than candidates. We can reject the hypothesis that these imbalances occurred by chance in all cases evaluated, for all ideological groups.

Looking more closely at our results, we observe that left-wing and centrist parties have a better performance in terms of racial descriptive representation than right-wing parties among candidates. This happens in the 2014 and 2018 elections, in State Assemblies and the Federal Deputy Chamber alike. For instance, in 2018, the DRI for right-wing parties at the Federal level was -0.41 while it was -0.20 and -0.23 for the center and left-wing parties, respectively.

Table 1: DRI by ideology: Candidates and elected, Federal and State elections, 2014 and 2018

	2014						2018					
Ideology	DRI	P-value	Black	Total	b	p	DRI	P-value	Black	Total	b	p
A. Federal Representatives - Candidates												
Left	–0.21	0.00	897	2,144	0.42	0.52	–0.23	0.00	1,088	2,431	0.45	0.57
Center	–0.25	0.00	1,245	3,055	0.41	0.53	–0.20	0.00	1,282	2,851	0.45	0.55
Right	–0.49	0.00	191	624	0.31	0.56	–0.41	0.00	747	2,260	0.33	0.54
B. State Assemblies - Candidates												
Left	–0.18	0.00	2,396	4,987	0.48	0.57	–0.15	0.00	2,571	4,906	0.52	0.60
Center	–0.24	0.00	3,420	7,540	0.45	0.57	–0.15	0.00	3,204	6,223	0.51	0.59
Right	–0.36	0.00	654	1,665	0.39	0.58	–0.30	0.00	2,069	4,840	0.43	0.58
C. Federal Representatives - Elected												
Left	–0.50	0.00	47	161	0.29	0.56	–0.59	0.00	40	152	0.26	0.59
Center	–0.75	0.00	35	223	0.16	0.57	–0.70	0.00	15	65	0.23	0.64
Right	–0.73	0.00	20	129	0.16	0.55	–0.61	0.00	69	293	0.24	0.57
D. State Assemblies - Elected												
Left	–0.48	0.00	101	318	0.32	0.57	–0.59	0.00	87	306	0.28	0.62
Center	–0.62	0.00	128	501	0.26	0.60	–0.54	0.00	62	187	0.33	0.64
Right	–0.67	0.00	45	213	0.21	0.58	–0.59	0.00	144	541	0.27	0.59

Notes: table shows the DRI calculated for candidates and elected representatives coming from parties of different ideologies according to Power and Rodrigues-Silveira (2018). The left-hand side of the table displays the DRI for 2014 elections, whereas the right-hand side presents the DRI for 2018 elections. Blocks (A) and (B) calculate the DRI by ideology for the pool of candidates. Blocs (C) and (D) calculate the DRI by ideology for the pool elected representatives. *Source:* DRI is calculated using data from the Brazilian Electoral Court (*Tribunal Superior Eleitoral – TSE*), the Continuous National Household Sample Survey (*Pesquisa Nacional Por Amostra de Domicílios Contínua – Continuous PNAD*), and Power and Rodrigues-Silveira (2018).

Moving from candidates to elected representatives, we observe not only a change in the levels of racial imbalance in representation but also a qualitative change in terms of differences among ideological groups, particularly for the 2018 election. In that year, left-wing parties had an aggregate DRI of -0.59, while right-wing parties' DRI was -0.61. Among state Assemblies, both party groups had the same DRI of -0.59. This means that although left-wing parties are more likely to nominate Non-white candidates in Brazilian Legislative elections, this does not translate into they electing more Non-white representatives. These results corroborate other findings in the literature regarding the nomination of minorities (Zingher, 2014; Sobolewska, 2013; Farrer and Zingher, 2018), at the same time that it highlights that different electoral tactics might result in similar minority election rates. In particular, we notice that the approval rate of right-wing parties is larger than left-wing and centrist ones. Investigating the factors accounting for these results is an important next step to the analysis of imbalances in racial descriptive representation.

5 Conclusion

This paper introduces a new measure of descriptive representation that allows for the comparison of imbalances across groups that can vary both in size and proportion in the electorate. The proposed measure improves upon alternatives by providing an accurate range of values — especially in extreme situations of misrepresentation — and by including confidence intervals, assisting researchers in deciding whether observed imbalances in representation are potentially due to randomness or due to actual lack of representation.

To illustrate the properties of our index, we use Brazilian legislative elections as empirical setting and investigate to what extent racial imbalances are a pervasive feature of Brazilian politics. Because racial composition in Brazil is highly heterogeneous, it provides the perfect case-study to assess the proposed measure. We show that descriptive representation is higher among candidates compared to elected representatives. However, there is considerable variation across Brazilian States. We also show that left-wing parties are more balanced with respect to candidates, but this does not translate into better representation among elected representatives.

Nonetheless, our index has limitations. The most important one is that it was designed to assess only the situation of one group in contrast to another, in our context, Whites versus Non-whites. One way around this limitation is to contrast separately several groups. For instance, to account for gender and race imbalances in Brazil, researchers can individually compare White men with Non-white men, Non-white women, and White women. We believe, however, that future research can generalize our index to n groups by replacing the binomial distribution with a multinomial distribution as the starting point of the derivation of the index.

References

Aguilar, R., Cunow, S., Desposato, S., and Barone, L. S. (2015). Ballot structure, candidate race, and vote choice in brazil. *Latin American Research Review*, pages 175–202.

Boas, T. C. and Smith, A. E. (2019). Looks like me, thinks like me: descriptive representation and opinion congruence in brazil. *Latin American Research Review*, 54(2):310–328.

Bueno, N. S. and Dunning, T. (2017). Race, resources, and representation: evidence from brazilian politicians. *World Politics*, 69(2):327–365.

Campos, L. A. and Machado, C. (2015). A cor dos eleitos: determinantes da sub-representação política dos não brancos no brasil. *Revista Brasileira de Ciência Política*, pages 121–151.

Campos, L. A. and Machado, C. (2017). O que afasta pretos e pardos da representação política? uma análise a partir das eleições legislativas de 2014. *Revista de Sociologia e Política*, 25:125–142.

Desai, Z. and Frey, A. (2021). Can descriptive representation help the right win votes from the poor? evidence from brazil. *American Journal of Political Science*.

Dovi, S. (2002). Preferable descriptive representatives: Will just any woman, black, or latino do? *American Political Science Review*, 96(4):729–743.

Duncan, O. D. and Duncan, B. (1955). A methodological analysis of segregation indexes. *American sociological review*, 20(2):210–217.

Farrer, B. D. and Zingher, J. N. (2018). Explaining the nomination of ethnic minority candidates: how party-level factors and district-level factors interact. *Journal of Elections, Public Opinion and Parties*, 28(4):467–487.

Firpo, S., França, M., and Rodrigues, L. (2020). Índice de equilíbrio racial: Uma proposta de mensuração da desigualdade racial entre e dentro das categorias ocupacionais.

Fraga, B. L. (2016a). Candidates or districts? reevaluating the role of race in voter turnout. *American Journal of Political Science*, 60(1):97–122.

Fraga, B. L. (2016b). Redistricting and the causal impact of race on voter turnout. *The Journal of Politics*, 78(1):19–34.

Gallagher, M. (1991). Proportionality, disproportionality and electoral systems. *Electoral studies*, 10(1):33–51.

Geese, L. (2022). Does descriptive representation narrow the immigrant gap in turnout? a comparative study across 11 western european democracies. *Political Studies*, pages 1–21.

Gerring, J., Jerzak, C., Erzen, and Oncel (2022). The composition of descriptive representation.

Griffin, J. D. and Newman, B. (2007). The unequal representation of latinos and whites. *The Journal of Politics*, 69(4):1032–1046.

Grose, C. R. (2011). *Congress in black and white: Race and representation in Washington and at home*. Cambridge University Press.

Hänni, M. and Saalfeld, T. (2020). Ethnic minorities and representation. In *Research Handbook on Political Representation*, pages 222–235. Edward Elgar Publishing.

Janusz, A. (2018). Candidate race and electoral outcomes: evidence from brazil. *Politics, Groups, and Identities*, 6(4):702–724.

Janusz, A. (2021). Race and resources in brazilian mayoral elections. *Political Research Quarterly*, page 10659129211032279.

Janusz, A. and Campos, L. A. (2021). Candidate advertisements and afro-brazilian political marginalization. *Latin American Research Review*, 56(4):761–778.

Janusz, A. and Sells, C. (2021). Race and campaign resources: Candidate identification numbers in brazil. *Journal of Politics in Latin America*, page 1866802X211052625.

Koppel, M. and Diskin, A. (2009). Measuring disproportionality, volatility and malapportionment: axiomatization and solutions. *Social Choice and Welfare*, 33(2):281–286.

Mansbridge, J. (1999). Should blacks represent blacks and women represent women? a contingent" yes". *The Journal of politics*, 61(3):628–657.

Monroe, B. L. (1994). Disproportionality and malapportionment: Measuring electoral inequity. *Electoral Studies*, 13(2):132–149.

Pantoja, A. D. and Segura, G. M. (2003). Does ethnicity matter? descriptive representation in legislatures and political alienation among latinos. *Social Science Quarterly*, 84(2):441–460.

Phillips, A. (1995). *The Politics of Presence*. OUP Oxford.

Pires, T. (2013). *Criminalização do racismo: entre política de reconhecimento e meio de legitimiação do controle social dos não reconhecidos*. PhD thesis, PUC-Rio.

Pitkin, H. (1967). *The Concept of Representation*. Campus (Berkeley, Calif.). University of California Press.

Power, T. J. and Rodrigues-Silveira, R. (2018). *Replication Data for: Mapping Ideological Preferences in Brazilian Elections, 1994-2018: A Municipal-Level Study*. Harvard Dataverse.

Ransom, R. L. and Sutch, R. (2001). *One kind of freedom: The economic consequences of emancipation*. Cambridge University Press.

Rocha, R. R., Tolbert, C. J., Bowen, D. C., and Clark, C. J. (2010). Race and turnout: Does descriptive representation in state legislatures increase minority voting? *Political Research Quarterly*, 63(4):890–907.

Rodrigues, L. C. (2021). *Sobre as cinzas do cativeiro: educação, trabalho e terra entre ex-escravos e descendentes após a abolição no oeste paulista*. PhD thesis, Universidade de São Paulo.

Ruedin, D. (2009). Ethnic group representation in a cross-national comparison. *The Journal of Legislative Studies*, 15(4):335–354.

Schwindt-Bayer, L. A. and Mishler, W. (2005). An integrated model of women's representation. *The Journal of politics*, 67(2):407–428.

Sobolewska, M. (2013). Party strategies and the descriptive representation of ethnic minorities: The 2010 british general election. *West European Politics*, 36(3):615–633.

Stockemer, D. (2015). Women's descriptive representation in developed and developing countries. *International Political Science Review*, 36(4):393–408.

Taagepera, R. and Grofman, B. (2003). Mapping the indices of seats–votes disproportionality and inter-election volatility. *Party Politics*, 9(6):659–677.

Telles, E. E. (2004). *Race in another America*. Princeton University Press.

Wallace, S. J. (2014). Representing latinos: Examining descriptive and substantive representation in congress. *Political Research Quarterly*, 67(4):917–929.

Wängnerud, L. (2009). Women in parliaments: Descriptive and substantive representation. *Annual Review of Political Science*, 12(1):51–69.

Whitby, K. J. (2007). The effect of black descriptive representation on black electoral turnout in the 2004 elections. *Social Science Quarterly*, 88(4):1010–1023.

Wylie, K. (2020). Taking bread off the table: race, gender, resources and political ambition in brazil. *European Journal of Politics and Gender*, 3(1):121–142.

Zapata-Barrero, R. (2017). How do political parties deal with the "diversity gap"? democratic deficits and party strategies. *Ethnic and Racial Studies*, 40(5):766–786.

Zingher, J. N. (2014). An analysis of the changing social bases of america's political parties: 1952–2008. *Electoral Studies*, 35:272–282.

A Derivation of the Descriptive Representation Index

We start deriving the Descriptive Representation Index (DRI) in the same way Ransom and Sutch (2001) did, by assuming that in a given occupation the number of Black workers, or in our case, the number of Non-white candidates or elected representatives, follows a binomial process:

$$P_j(X = x_j) = \binom{n_j}{x_j} p_j^{x_j} (1 - p_j)^{n_j - x_j} \tag{A.1}$$

where x_j is the number of black candidates or elected representatives running for office in state j, which has a total number of candidates or elected representatives n_j and p_j is the share of black voters in that state.

Based on the properties of the binomial distribution, it is possible to define a level of racial balance in an election by comparing the observed number of Non-white candidates or elected representatives, x_j, with its expected value, $n_j p_j$. We first standardize the distance $x_j - n_j p_j$ by dividing it by the standard deviation of the binomial distribution:

$$v_j = \frac{x_j - n_j p_j}{\sqrt{n_j p_j (1 - p_j)}} \tag{A.2}$$

The DRI is simply a standardized version of v_j in such a way that it lies between -1 and 1, while zero represents perfect racial balance. To do so, define $b_j = x_j / n_j$, the share of Non-white individuals in the population of interest, such as candidates, elected representatives or workers in a given occupation. It can be shown that:

$$v(b_j) = \sqrt{\frac{n_j \left(b_j - p_j\right)^2}{p_j \left(1 - p_j\right)}} \tag{A.3}$$

Since $b_j \in [0, 1]$, $v(b_j)$ has two limits given by:

$$v(0) = \sqrt{\frac{np}{(1 - p)}}\text{, if } b_j = 0 \tag{A.4}$$

and

$$v(1) = \sqrt{\frac{n\,(1 - p)}{p}}\text{, if } b_j = 1 \tag{A.5}$$

Up to this point, our derivation of the DRI follows exactly that of Ransom and Sutch (2001). We depart from theirs by replacing the denominator used to ensure that the index falls within -1 and 1 bound. Ransom and Sutch divide v_j in Equation A.3 by:

$$\phi(b_j) = \sqrt{n_j \left(\frac{p_j}{1-p_j}\right)^{1-2b_j}} \tag{A.6}$$

When b is 0 and 1, the ratio between A.3 and Equation A.6 is the same as Equations A.4 and A.5, respectively. Instead, we replace Equation A.6 by the linear combination of Equations A.4 and A.5:

$$\omega(b_j) = \sqrt{n_j \left[b_j \left(\frac{n_j(1-p_j)}{p_j} \right) + (1-b_j) \left(\frac{n_j p_j}{(1-p_j)} \right) \right]} \tag{A.7}$$

It is easy to see that when b takes values 0 and 1, ω is the same as in expressions A.4 and A.5. Dividing $v(b_j)$ in Equation A.3 by $\omega(b_j)$, after some algebra we arrive at the Descriptive Representation Index (DRI):

$$\text{DRI}(b_j, p_j) = \frac{b_j - p_j}{\sqrt{(b_j - p_j)^2 + b_j(1-b_j)}} \tag{A.8}$$

Like Ransom and Sutch's measure, the DRI also assume -1 and 1 when b_j is respectively 0 and 1, while the value of zero only happens when $b_j = p_j$. However, contrary to Ransom and Sutch, whose measure can take values outside the -1 and 1 interval in some cases, ours always lies within this interval.

B DRI's statistical properties

Our evaluation of the DRI's statistical properties rests on the binomial distribution. Consider a population from which the minority share, $0 < p < 1$, is known. If N individuals are randomly selected from this population to become their representatives, the total number of cases in which a minority is picked is $X = \sum_{n=1}^{N} x_n$, where x is a binary variable that takes value 1 when the person belongs to the minority group. The distribution of the number of selected representatives belonging to the minority group, X, is given by the binomial distribution:

$$Pr\left[X = x | N, p\right] = p^X (1-p)^{(N-X)} \frac{N!}{X!\,(N-X)!} \tag{B.9}$$

From Equation A.8 and defining $b = X/N$, the distribution of the DRI can be written as

$$Pr\left[DRI = d|N,p\right] = Pr\left[\frac{\frac{X}{N}-p}{\sqrt{\left(\frac{X}{N}-p\right)^2+\frac{X}{N}\left(1-\frac{X}{N}\right)}}\right] \tag{B.10}$$

Let $U = X/N - p$, then after some algebra, Equation B.10 can be written, as

$$\begin{aligned} Pr\left[DRI = d|N,p\right] &= Pr\left[\frac{U^2}{\sqrt{U+p-2pU-p^2}} = d\right] \\ &= Pr\left[\frac{U^2}{U+p-2pU-p^2} = d^2\right] \\ &= Pr\left[U^2 - d^2\left(1-2p\right)U - d^2p\left(1-p\right) = 0\right] \end{aligned} \tag{B.11}$$

Using the properties of second degree polynomials, we know that the equality in the last probability will hold when

$$U = u_1 = \frac{d^2\left(1-2p\right)+\sqrt{d^2\left(d^2\left(1-2p\right)^2\right)+4p\left(1-p\right)}}{2} \tag{B.12}$$

or

$$U = u_2 = \frac{d^2\left(1-2p\right)-\sqrt{d^2\left(d^2\left(1-2p\right)^2\right)+4p\left(1-p\right)}}{2} \tag{B.13}$$

Hence, the probability of the DRI taking some value d is given by:

$$Pr\left[DRI = d|N,p\right] = \mathbb{1}\left[DRI \geq 0\right]Pr\left[U = u_1\right] + \mathbb{1}\left[DRI < 0\right]Pr\left[U = u_2\right] \tag{B.14}$$

Replacing U with $X/N - p$, as before, we arrive at

$$\begin{aligned} Pr\left[DRI = d|N,p\right] =& \mathbb{1}\left[DRI \geq 0\right]Pr\left[\frac{X}{N}-p = u_1\right] \\ &+ \mathbb{1}\left[DRI < 0\right]Pr\left[\frac{X}{N}-p = u_2\right] \\ =& \mathbb{1}\left[DRI \geq 0\right]Pr\left[X = \left(u_1+p\right)N\right] \\ &+ \mathbb{1}\left[DRI < 0\right]Pr\left[X = \left(u_2+p\right)N\right] \end{aligned} \tag{B.15}$$

Because X follows a binomial distribution, the distribution of the DRI takes the following form:

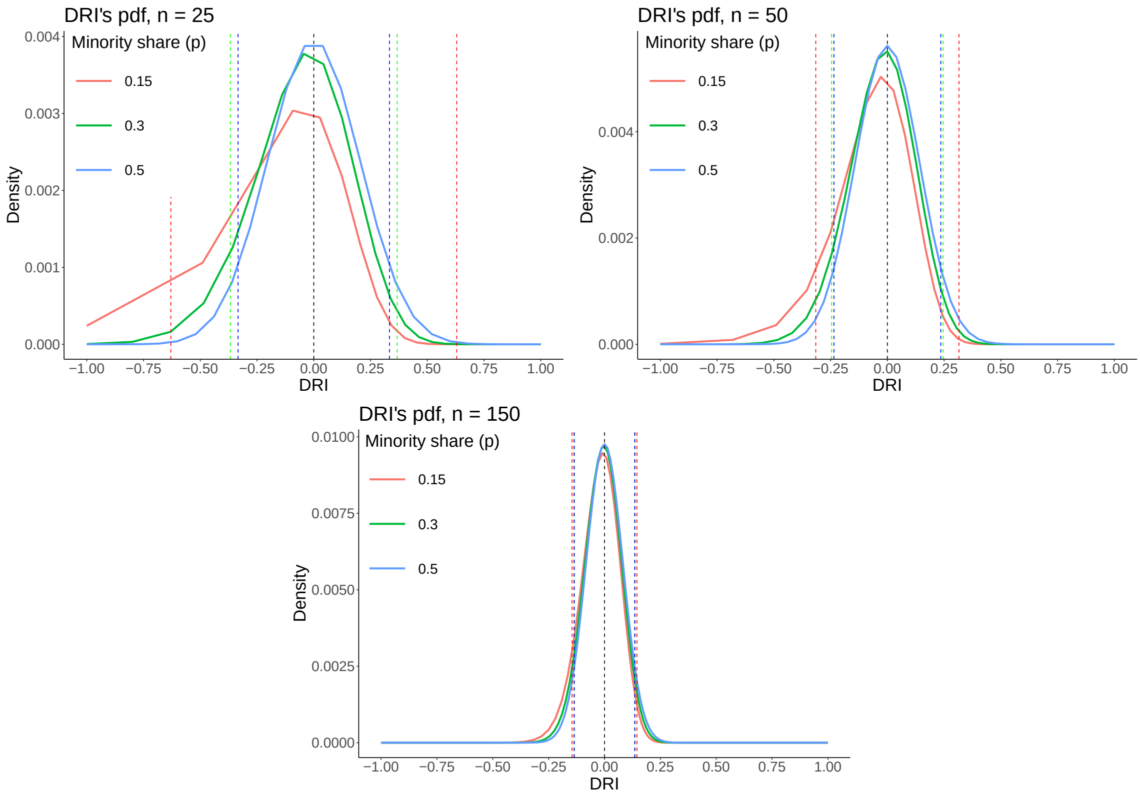
$$Pr\left[DRI = d|N, p\right] = \mathbb{1}[DRI \geq 0] \frac{p^{(u_1+p)N} \left(1-p\right)^{(N-(u_1+p)N)} N!}{(u_1+p)N! \left(N-(u_1+p)N\right)!} + \mathbb{1}[DRI < 0] \frac{p^{(u_2+p)N} \left(1-p\right)^{(N-(u_2+p)N)} N!}{(u_2+p)N! \left(N-(u_2+p)N\right)!} \tag{B.16}$$

We can use the distribution in Equation B.16 to investigate how likely a given value of the DRI is to have arisen by chance, given the number of members in a group of representatives and the minority share in the population. In particular, the smaller the number of representatives, the higher the odds of extreme values of the DRI occurring even if the distribution of representatives is purely random.

To show that, we plot two graphs. Figure B.4 shows the probability distribution of the DRI under different scenarios, considering values of $p = \{0.15, 0.3, 0.5\}$ and $N = \{25, 50, 150\}$.[19] We can see that for small minority shares, the distribution of the DRI is skewed to the left, meaning a higher probability of large imbalances. However, as sample sizes increase, this skewedness reduces significantly, and the DRI is centered around zero for any p.

[19]The plotted probabilities are based on interpolation of discrete values to one thousand points.

Figure B.4: DRI's probability distributions for given pairs (N, p)

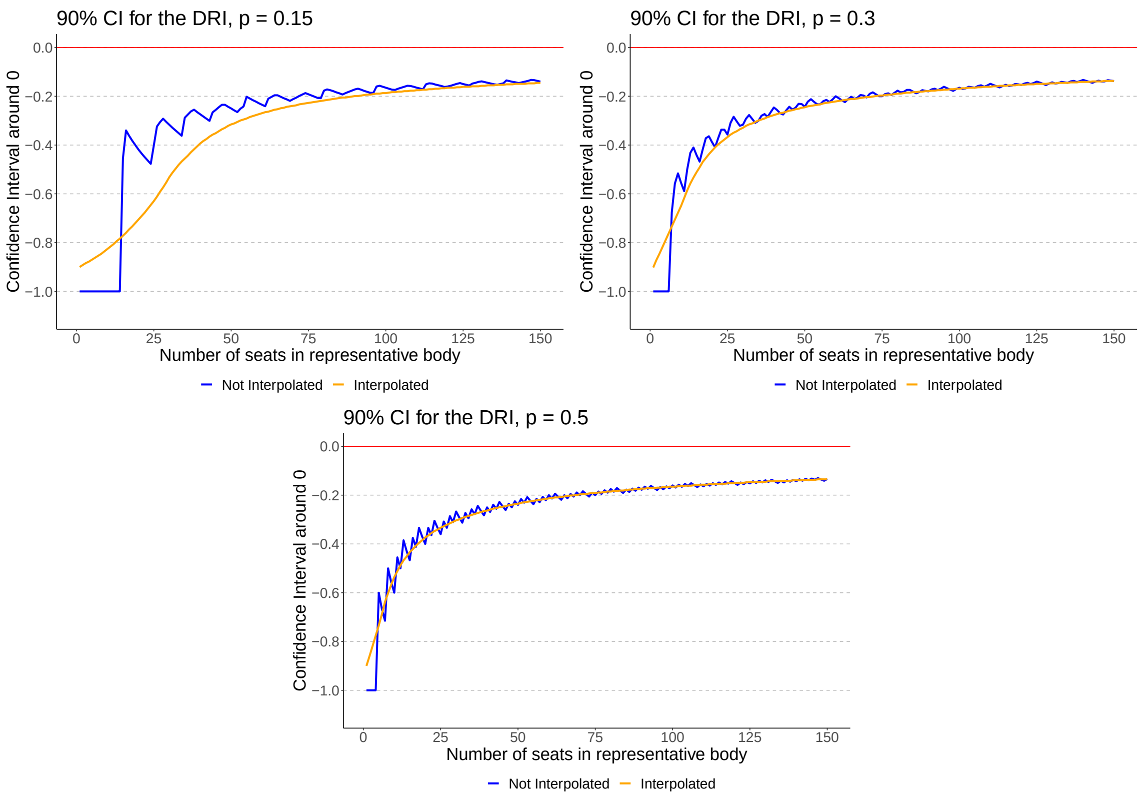


Notes: These figures plot in solid lines the probability distribution of the DRI, based on Equation B.16. These probabilities are obtained via interpolation of one thousand values of the actual function based on integer values. Dashed lines represent symmetric bounds that amount to 90% probability mass around zero.

Figure B.4 also shows (in dashed lines) the intervals centered around zero that comprise 90% of the probability mass for each distribution given (N, p). These intervals are constructed in the following way: we start at $DRI = 0$ and compute the odds of this happening. Then, we create a window around -0.01 and 0.01 and compute the new mass. We continue increasing the length of this window by 0.01 until it contains 90% of the mass. We see that for small p and small N, these windows are large. However, as N grows, the windows become more similar.

The behavior of these windows as the number of representatives (N) grows can be analyzed in Figure B.5. Each panel considers a population with a given minority share $p = \{0.15, 0.30, 0.50\}$. The blue lines are confidence intervals around zero created using the process described above, by marginally increasing the size of the window until each reaches a probability mass of 90%. However, because the binomial distribution is discrete, it has many jumps along the line. To deal with this limitation, we also interpolate it, resulting in the smooth orange line. In both cases, it is clear that the intervals reduce considerably in size as N grows.

Figure B.5: Confidence intervals for the DRI's as number of representatives grow



Notes: These figures plot 90% probability intervals of the DRI around zero, based on Equation B.16. This probability mass is obtained by summing all probabilities within increasingly large symmetric windows around zero. The process of enlarging these windows ceases when the mass probability becomes as large as 0.9. The blue line shows the interval obtained by using Equation B.16 with integer values, while the orange line is obtained via interpolation of this function to one thousand values.

Notice, however, that these confidence intervals are only valid for relatively small N, as the factorial of N in Equation B.16 rapidly becomes too large and cannot be computed. For large samples, confidence intervals must be constructed using asymptotic theory. Under the fact that the number of representatives is randomly generated from the binomial distribution, we know from the Law of Large Numbers that $X/N \xrightarrow{p} p$, which implies that the DRI converges to:

$$plim(DRI) = \frac{plim\left(\frac{X}{N} - p\right)}{plim\left(\sqrt{\left(\frac{X}{N} - p\right)^2 + \frac{X}{N}\left(1 - \frac{X}{N}\right)}\right)} = \frac{0}{\sqrt{p\left(1-p\right)}} = 0 \tag{B.17}$$

Now, from the Central Limit Theorem, we know that:

$$\sqrt{N}\left(\frac{X}{N} - p\right) \xrightarrow{d} N\left(0, p\left(1-p\right)\right) \tag{B.18}$$

Then, combining B.17 and B.18 using the Slutsky Theorem, we find the limiting distribution for the DRI:

$$\sqrt{N}DRI = \frac{\sqrt{N}\left(\frac{X}{N} - p\right)}{\sqrt{\left(\frac{X}{N} - p\right)^2 + \frac{X}{N}\left(1 - \frac{X}{N}\right)}} \xrightarrow{d} N\left(0, \frac{p(1-p)}{\left(\sqrt{p(1-p)}\right)^2}\right) \sim N(0,1) \quad \text{(B.19)}$$

Therefore, based on the limiting distribution B.19, for some confidence level $0 < \alpha < 1$, we can establish bounds around zero such that:

$$\begin{aligned} 1-\alpha &= Pr\left[-c \leq DRI \leq c\right] \\ &= Pr\left[-\sqrt{N}c \leq \sqrt{N}DRI \leq \sqrt{N}c\right] \\ &= \Phi(\sqrt{N}c) - \Phi(-\sqrt{N}c) \end{aligned} \quad \text{(B.20)}$$

Where $\Phi(a)$ is the cumulative density distribution of the standard normal. Because this distribution is symmetric, $\Phi(a) = 1 - \Phi(-a)$, and we have:

$$\begin{aligned} 1-\alpha &= Pr\left[-c \leq DRI \leq c\right] \\ &= \Phi(\sqrt{N}c) - \left(1 - \Phi(\sqrt{N}c)\right) \\ &= 2\Phi(\sqrt{N}c) - 1 \end{aligned} \quad \text{(B.21)}$$

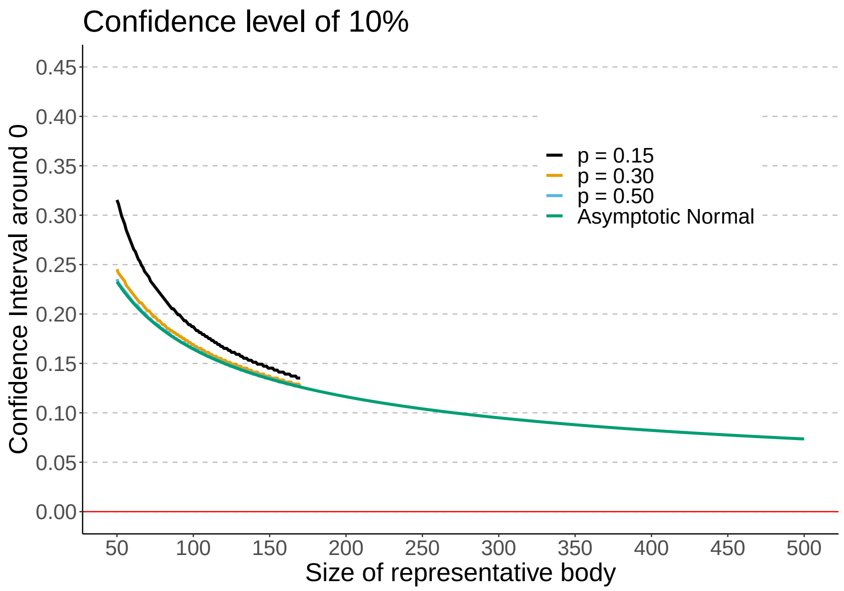
Which implies that

$$c = \Phi^{-1}\left(\frac{1+\alpha/2}{\sqrt{N}}\right) \quad \text{(B.22)}$$

where $\phi^{-1}(.)$ is the inverse function of the standardize normal probability density function.

Figure B.6 shows how these confidence intervals evolve as N grows. It can be seen that as the size of the representative body grows, the size of the confidence interval becomes smaller. For comparison, we also plot the confidence intervals obtained by interpolation using the binomial distribution. For minorities shares closer to the 0.5, these intervals are almost similar even for small number of representatives. However, for all minority shares, as the size of the representative bodies grows, the confidence intervals approximate the one obtained asymptotically.

Figure B.6: Evolution of asymptotic confidence intervals for the DRI under no imbalance assumption



Notes: These figures plot 90% probability intervals of the DRI around zero. The green solid line is obtained from the asymptotic distribution of the DRI, Equation B.22. The other lines are based on the discrete distribution, in Equation B.16. See note in Figure B.5 on how these discrete bounds are computed with interpolation.

The same logic can also be used to derive pseudo p-values for the DRI. That is, assuming a random distribution of representatives, what is the confidence level that includes a given empirical value $\hat{d}$ of the DRI in the region where we cannot discard imbalances that arise by chance. This value if obtained in the following way:

$$\begin{aligned}\text{p-value} &= 1 - Pr\left[-\hat{d} \leq DRI \leq \hat{d}\right] \\ &= 1 - \left(2\Phi\left(\sqrt{N}\hat{d}\right) - 1\right) \\ &= 2 - 2\Phi\left(\sqrt{N}\hat{d}\right) \\ &= 2\left(1 - \Phi\left(\sqrt{N}\hat{d}\right)\right) \end{aligned} \tag{B.23}$$

C Additional Tables

Table C.1: Federal Representatives - DRI by State - Candidates

State	DRI	P-value	Black	Total	b	p
AC	–0.24	0.06	42	60	0.70	0.81
AL	–0.40	0.00	52	98	0.53	0.75
AM	–0.32	0.01	48	75	0.64	0.80
AP	–0.20	0.04	68	101	0.67	0.77
BA	–0.29	0.00	182	282	0.65	0.79
CE	–0.38	0.00	96	190	0.51	0.71
DF	–0.30	0.00	51	124	0.41	0.57
ES	–0.43	0.00	54	148	0.36	0.60
GO	–0.55	0.00	28	92	0.30	0.61
MA	–0.44	0.00	128	222	0.58	0.82
MG	–0.31	0.00	238	589	0.40	0.57
MS	–0.34	0.00	40	110	0.36	0.54
MT	–0.37	0.00	39	85	0.46	0.66
PA	–0.33	0.00	111	170	0.65	0.82
PB	–0.33	0.00	41	90	0.46	0.63
PE	–0.16	0.05	91	151	0.60	0.68
PI	–0.25	0.02	56	85	0.66	0.78
PR	–0.41	0.00	40	286	0.14	0.29
RJ	–0.15	0.00	392	861	0.46	0.53
RN	–0.44	0.00	30	80	0.38	0.61
RO	–0.31	0.01	44	81	0.54	0.70
RR	–0.20	0.08	51	77	0.66	0.76
RS	–0.24	0.00	29	304	0.10	0.17
SC	–0.03	0.75	14	124	0.11	0.12
SE	–0.39	0.00	39	72	0.54	0.75
SP	–0.19	0.00	311	1,220	0.25	0.34
TO	–0.62	0.00	18	46	0.39	0.78

(a) 2014

State	DRI	P-value	Black	Total	b	p
AC	–0.12	0.27	57	79	0.72	0.78
AL	–0.55	0.00	29	71	0.41	0.73
AM	–0.30	0.00	80	122	0.66	0.80
AP	–0.10	0.32	73	95	0.77	0.81
BA	–0.21	0.00	317	444	0.71	0.81
CE	–0.29	0.00	126	226	0.56	0.71
DF	–0.21	0.01	79	158	0.50	0.61
ES	–0.41	0.00	62	155	0.40	0.62
GO	–0.46	0.00	75	196	0.38	0.63
MA	–0.40	0.00	110	186	0.59	0.80
MG	–0.32	0.00	350	808	0.43	0.60
MS	–0.36	0.00	41	117	0.35	0.54
MT	–0.31	0.00	55	107	0.51	0.68
PA	–0.33	0.00	82	128	0.64	0.81
PB	–0.53	0.00	50	136	0.37	0.67
PE	–0.39	0.00	152	325	0.47	0.68
PI	–0.47	0.00	65	122	0.53	0.80
PR	–0.44	0.00	64	412	0.16	0.33
RJ	–0.16	0.00	481	1,043	0.46	0.54
RN	–0.40	0.00	48	113	0.42	0.64
RO	–0.27	0.01	50	91	0.55	0.69
RR	0.01	0.94	91	132	0.69	0.69
RS	–0.24	0.00	51	400	0.13	0.21
SC	–0.30	0.00	23	229	0.10	0.20
SE	–0.29	0.00	70	107	0.65	0.80
SP	–0.26	0.00	398	1,460	0.27	0.39
TO	–0.51	0.00	38	80	0.47	0.77

(b) 2018

Table C.2: State Assemblies - DRI by State - Candidates

State	DRI	P-value	Black	Total	b	p
AC	0.08	0.09	409	486	0.84	0.81
AL	–0.19	0.00	165	252	0.65	0.75
AM	–0.19	0.00	383	539	0.71	0.80
AP	–0.21	0.00	220	329	0.67	0.77
BA	–0.19	0.00	387	548	0.71	0.79
CE	–0.38	0.00	269	535	0.50	0.71
ES	–0.25	0.00	214	459	0.47	0.60
GO	–0.35	0.00	296	691	0.43	0.61
MA	–0.40	0.00	286	472	0.61	0.82
MG	–0.30	0.00	422	1,025	0.41	0.57
MS	–0.27	0.00	149	372	0.40	0.54
MT	–0.25	0.00	144	273	0.53	0.66
PA	–0.22	0.00	440	611	0.72	0.82
PB	–0.32	0.00	142	308	0.46	0.63
PE	–0.20	0.00	269	465	0.58	0.68
PI	–0.34	0.00	129	214	0.60	0.78
PR	–0.26	0.00	136	717	0.19	0.29
RJ	–0.21	0.00	726	1,710	0.42	0.53
RN	–0.58	0.00	67	234	0.29	0.61
RO	–0.22	0.00	213	360	0.59	0.70
RR	–0.29	0.00	231	380	0.61	0.76
RS	–0.29	0.00	56	665	0.08	0.17
SC	–0.20	0.00	28	398	0.07	0.12
SE	–0.32	0.00	92	157	0.59	0.75
SP	–0.19	0.00	447	1,762	0.25	0.34
TO	–0.26	0.00	150	230	0.65	0.78

(a) 2014

State	DRI	P-value	Black	Total	b	p
AC	0.04	0.39	343	432	0.79	0.78
AL	–0.17	0.01	179	274	0.65	0.73
AM	–0.18	0.00	408	565	0.72	0.80
AP	–0.17	0.00	314	428	0.73	0.81
BA	–0.13	0.00	429	570	0.75	0.81
CE	–0.25	0.00	296	508	0.58	0.71
ES	–0.14	0.00	319	581	0.55	0.62
GO	–0.24	0.00	406	795	0.51	0.63
MA	–0.32	0.00	293	456	0.64	0.80
MG	–0.24	0.00	580	1,213	0.48	0.60
MS	–0.29	0.00	123	316	0.39	0.54
MT	–0.25	0.00	179	325	0.55	0.68
PA	–0.15	0.00	466	627	0.74	0.81
PB	–0.30	0.00	186	364	0.51	0.67
PE	–0.30	0.00	319	616	0.52	0.68
PI	–0.32	0.00	129	203	0.64	0.80
PR	–0.50	0.00	95	704	0.13	0.33
RJ	–0.10	0.00	1,103	2,238	0.49	0.54
RN	–0.36	0.00	134	298	0.45	0.64
RO	–0.11	0.04	238	372	0.64	0.69
RR	0.12	0.02	315	427	0.74	0.69
RS	–0.25	0.00	100	797	0.13	0.21
SC	–0.21	0.00	54	430	0.13	0.20
SE	–0.29	0.00	188	287	0.66	0.80
SP	–0.26	0.00	526	1,937	0.27	0.39
TO	–0.35	0.00	122	206	0.59	0.77

(b) 2018

Table C.3: Federal Representatives - DRI by State - Elected Candidates

State	DRI	P-value	Black	Total	b	p
AC	–0.67	0.06	3	8	0.38	0.81
AL	–0.78	0.02	2	9	0.22	0.75
AM	–0.51	0.15	4	8	0.50	0.80
AP	–0.05	0.90	6	8	0.75	0.77
BA	–0.52	0.00	19	39	0.49	0.79
CE	–0.81	0.00	4	22	0.18	0.71
DF	–0.80	0.02	1	8	0.13	0.57
ES	–0.54	0.09	3	10	0.30	0.60
GO	–0.75	0.00	3	17	0.18	0.61
MA	–0.87	0.00	3	18	0.17	0.82
MG	–0.79	0.00	7	53	0.13	0.57
MS	–1.00	0.00	0	8	0.00	0.54
MT	–0.51	0.15	3	8	0.38	0.66
PA	–0.81	0.00	4	17	0.24	0.82
PB	–0.78	0.01	2	12	0.17	0.63
PE	–0.44	0.03	11	25	0.44	0.68
PI	–0.82	0.01	2	10	0.20	0.78
PR	–0.82	0.00	1	30	0.03	0.29
RJ	–0.61	0.00	10	46	0.22	0.53
RN	–0.83	0.02	1	8	0.13	0.61
RO	–0.38	0.29	4	8	0.50	0.70
RR	–0.46	0.19	4	8	0.50	0.76
RS	–1.00	0.00	0	31	0.00	0.17
SC	–1.00	0.00	0	16	0.00	0.12
SE	–0.88	0.01	1	8	0.13	0.75
SP	–0.77	0.00	4	70	0.06	0.34
TO	–1.00	0.00	0	8	0.00	0.78

(a) 2014

State	DRI	P-value	Black	Total	b	p
AC	–0.64	0.07	3	8	0.38	0.78
AL	–0.50	0.13	4	9	0.44	0.73
AM	–0.12	0.73	6	8	0.75	0.80
AP	–0.53	0.13	4	8	0.50	0.81
BA	–0.57	0.00	18	39	0.46	0.81
CE	–0.52	0.01	9	22	0.41	0.71
DF	–0.83	0.02	1	8	0.13	0.61
ES	–0.87	0.01	1	10	0.10	0.62
GO	–0.85	0.00	2	17	0.12	0.63
MA	–0.59	0.01	8	18	0.44	0.80
MG	–0.81	0.00	7	53	0.13	0.60
MS	–0.32	0.37	3	8	0.38	0.54
MT	–0.86	0.02	1	8	0.13	0.68
PA	–0.49	0.04	9	17	0.53	0.81
PB	–0.45	0.12	5	12	0.42	0.67
PE	–0.82	0.00	4	25	0.16	0.68
PI	–0.83	0.01	2	10	0.20	0.80
PR	–0.85	0.00	1	29	0.03	0.33
RJ	–0.38	0.01	16	46	0.35	0.54
RN	–1.00	0.00	0	8	0.00	0.64
RO	–0.13	0.70	5	8	0.63	0.69
RR	–0.23	0.55	4	7	0.57	0.69
RS	–1.00	0.00	0	31	0.00	0.21
SC	–0.48	0.05	1	16	0.06	0.20
SE	–0.66	0.06	3	8	0.38	0.80
SP	–0.82	0.00	4	69	0.06	0.39
TO	–0.64	0.07	3	8	0.38	0.77

(b) 2018

Table C.4: State Assemblies - DRI by State - Elected Candidates

State	DRI	P-value	Black	Total	b	p
AC	–0.30	0.15	16	24	0.67	0.81
AL	–0.66	0.00	9	27	0.33	0.75
AM	–0.40	0.05	14	24	0.58	0.80
AP	–0.42	0.04	13	24	0.54	0.77
BA	–0.63	0.00	25	63	0.40	0.79
CE	–0.77	0.00	10	46	0.22	0.71
ES	–0.47	0.01	10	29	0.34	0.60
GO	–0.76	0.00	7	41	0.17	0.61
MA	–0.76	0.00	12	42	0.29	0.82
MG	–0.68	0.00	15	77	0.19	0.57
MS	–0.40	0.05	8	24	0.33	0.54
MT	–0.74	0.00	5	24	0.21	0.66
PA	–0.43	0.01	24	41	0.59	0.82
PB	–0.58	0.00	11	36	0.31	0.63
PE	–0.45	0.00	21	49	0.43	0.68
PI	–0.69	0.00	10	30	0.33	0.78
PR	–0.44	0.00	7	54	0.13	0.29
RJ	–0.72	0.00	11	70	0.16	0.53
RN	–0.89	0.00	2	24	0.08	0.61
RO	–0.54	0.01	9	23	0.39	0.70
RR	–0.67	0.00	8	24	0.33	0.76
RS	–0.75	0.00	1	55	0.02	0.17
SC	–1.00	0.00	0	40	0.00	0.12
SE	–0.66	0.00	8	24	0.33	0.75
SP	–0.71	0.00	7	93	0.08	0.34
TO	–0.54	0.01	11	24	0.46	0.78

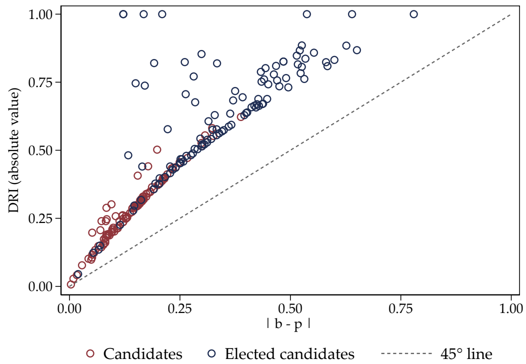
(a) 2014

State	DRI	P-value	Black	Total	b	p
AC	–0.15	0.46	17	24	0.71	0.78
AL	–0.74	0.00	7	27	0.26	0.73
AM	–0.66	0.00	9	24	0.38	0.80
AP	–0.67	0.00	9	24	0.38	0.81
BA	–0.59	0.00	28	63	0.44	0.81
CE	–0.69	0.00	13	46	0.28	0.71
ES	–0.47	0.01	11	30	0.37	0.62
GO	–0.63	0.00	11	41	0.27	0.63
MA	–0.73	0.00	13	42	0.31	0.80
MG	–0.69	0.00	16	77	0.21	0.60
MS	–0.63	0.00	5	24	0.21	0.54
MT	–0.76	0.00	5	24	0.21	0.68
PA	–0.57	0.00	19	41	0.46	0.81
PB	–0.58	0.00	12	36	0.33	0.67
PE	–0.84	0.00	7	49	0.14	0.68
PI	–0.56	0.00	14	30	0.47	0.80
PR	–0.58	0.00	6	54	0.11	0.33
RJ	–0.44	0.00	22	70	0.31	0.54
RN	–0.48	0.02	9	24	0.38	0.64
RO	–0.36	0.08	12	24	0.50	0.69
RR	–0.28	0.17	13	24	0.54	0.69
RS	–0.82	0.00	1	55	0.02	0.21
SC	–0.74	0.00	1	40	0.03	0.20
SE	–0.56	0.01	11	24	0.46	0.80
SP	–0.68	0.00	10	93	0.11	0.39
TO	–0.48	0.02	12	24	0.50	0.77

(b) 2018

D DRI and other descriptive representation indicators

Figure D.7: Differences between DRI with respect to $|b - p|$



Note: This Figure plots, in the vertical axis, the empirical values of the DRI obtained for each Brazilian state in the 2014 and 2018 election for State Chambers and The Federal Deputy Chamber against the absolute difference between the same of Non-white among representatives and among the electorate $|b - p|$