Department of Economics- FEA/USP

Heterogeneity, electoral rules and the number of candidates: an empirical investigation sing a quasi-natural experiment

CARLOS EDUARDO S. GONÇALVES

MAURO RODRIGUES JR.

RICARDO A. MADEIRA

WORKING PAPER SERIES Nº 2012-25

DEPARTMENT OF ECONOMICS, FEA-USP
WORKING PAPER Nº 2012-25

Heterogeneity, electoral rules and the number of candidates: an empirical investigation using a quasi-natural experiment

Carlos Eduardo S. Gonçalves (cesg@usp.br)

Mauro Rodrigues jr (mrodrigues@usp.br

Ricardo A. Madeira (rmadeira@usp.br)

Abstract:

There is an open debate in social sciences concerning the impact of different electoral rules and societal heterogeneity on the number of candidates vying for a seat during election times. Using data from Brazil's municipal mayoral elections, this paper assesses the empirical validity of the so-called nuanced view, which claims the interaction between societal heterogeneity and institutional permissiveness of electoral rules (presence of a runoff, in our case) is key to explain the number of candidates. Our study differs from others in the literature in two major aspects: (i) we have a truly exogenous source of variation in electoral rules due to a change in legislation introduced by the constitutional reform of 1988 and, (ii) we use panel-data techniques that allow for a more reliable identification of the parameters. Our results provide support for the nuanced-view: the coefficient of the interaction between heterogeneity and the presence of a runoff is always positive and statistically significant.

Keywords: runoff, heterogeneity, number of candidates

JEL Codes: D72; D63 e D02

Heterogeneity, electoral rules and the number of candidates: an empirical investigation using a quasi-natural experiment

Carlos Eduardo S. Gonçalves* Mauro Rodrigues Jr.† Ricardo A. Madeira‡

August 2012

Abstract

There is an open debate in social sciences concerning the impact of different electoral rules and societal heterogeneity on the number of candidates vying for a seat during election times. Using data from Brazil's municipal mayoral elections, this paper assesses the empirical validity of the so-called nuanced view, which claims the interaction between societal heterogeneity and institutional permissiveness of electoral rules (presence of a runoff, in our case) is key to explain the number of candidates. Our study differs from others in the literature in two major aspects: (i) we have a truly exogenous source of variation in electoral rules due to a change in legislation introduced by the constitutional reform of 1988 and, (ii) we use panel-data techniques that allow for a more reliable identification of the parameters. Our results provide support for the nuanced-view: the coefficient of the interaction between heterogeneity and the presence of a runoff is always positive and statistically significant.

Key words: runoff, heterogeneity, number of candidates

*University of São Paulo, Department of Economics

†University of São Paulo, Department of Economics

‡This is the corresponding author for this submission. University of São Paulo, Department of Economics, Av. Prof. Luciano Gualberto, 908, São Paulo, Brazil, 05508-900, rmadeira@usp.br, tel: + 55 11 3091 6070.

1 Introduction

There is an open debate in social sciences concerning the impact of different electoral rules and societal heterogeneity on the number of candidates vying for a seat in an election. The theme is an old one: more than fifty years ago, Duverger (1954) put forth the idea that second-ballot voting rule (runoff) and proportional representation favors multipartyism. Since then, much of the empirical literature has focused on assessing the validity of Duverger's claim about the direct impact of electoral rules on the number of candidates and parties. Recently, though, researchers have begun to espouse a more "nuanced view" that electoral rules do matter, but not in isolation. This view's advocates argue the number of candidates competing in elections must also be a function of some intrinsic measure of divergence in societies' most preferred policies[1]. To be sure, in the extreme case where everyone has the same preferences regarding the menu of available public policies, the number of candidates should be greatly reduced independently of the institutional framework governing the election.

This paper uses data from five municipal mayoral elections in Brazil and provides evidence lending credence to the "nuanced view". Our results are in line with a handful of empirical articles supporting the idea that both heterogeneity within the population and the permissiveness of the electoral system (degree of proportionality; runoffs) matter. In particular, the interaction between heterogeneity and the permissiviness of electoral rules is found to be positive in many studies (Ordeshook and Shvetsova (1994); Cox (1997); Cox and Amorim Neto (1997); Jones (1999); Golder (2006))[2]. That is, a permissive electoral rule magnifies the influence exerted by the degree of heterogeneity on the number of candidates.

We believe to have a more clear-cut identification strategy than other papers in the literature, which do not address the issue of endogeneity – reverse causality and omitted variables bias. Endogeneity bias is an important drawback in this context because more fragmented societies should naturally be inclined to opt for electoral rules that allow for broader representation (e.g, runoffs). By exploring a major change in Brazil's electoral law that accompanied the constitutional reform of 1988, we believe to able to truly identify causal relationships.

In particular, before the 1988 nationwide reform, mayoral elections in all municipalities

[1] See Cox (1997) for a comprehensive discussion of this point.

[2] Jones (2004) brings evidence against the "nuanced view" for presidential democracies. He finds that the pure institutional model has a higher explanatory power when compared to an interactive model.

were governed by simple plurality rule. However, from the 1992 election onwards municipalities with more than 200,000 eligible voters were forced to adopt a runoff stage – whereas below that threshold simple plurality remained mandatory. Since these political units could not choose what voting rule to follow, the legislation change endowed us with a quasi-natural experiment that allows for a clear test of the above mentioned "nuanced view" in a panel-set framework. To the advantage of our identification strategy, not only our data set comprises municipalities with exogenously different electoral rules at the same point in time, but also features cross-sectional units that change status through time. Moreover, here the impact of all other features of the electoral system (aside from runoff) are "controlled for" since they apply uniformly to all municipalities, reducing the likelihood of omitted variables bias. Finally, whilst other papers in the literature use ethnico-linguistic fragmentation as a measure of societal cleavage, we employ an income inequality index instead. This is because we believe income inequality to be the most telling indicator of social divide in Brazilian society.

We find a robust positive effect of the interaction between heterogeneity and the runoff dummy on the number of candidates, thus providing support for the nuanced view. That is, the presence of a runoff magnifies the effect of heterogeneity on the number of candidates. Heterogeneity alone is an important determinant of the number of candidates only for subsamples of municipalities with more than 50,000 voters.

The rest of this paper is organized as follows. In section 2, we review the literature. In section 3 we introduce our dataset and comment upon the identification strategy employed. In section 4, we present our results, and section 5 finally concludes.

2 Literature

Duverger's propositions spawned two strands of research. On the theoretical front, a set of studies using game-theoretic approaches and endogenizing politicians' entry decision lent credence to the idea that two candidate elections are indeed more likely under runoff elections. Osborne and Slivinski (1996), for instance, model citizens' strategic decisions as to whether or not compete for a seat. In their model, the combination of parameters yielding two-candidate elections is more stringent under runoff than plurality[3]. Other papers, such as Feddersen

[3]Nevertheless, in two other theoretical contributions, Callander (2005) and Bouton (2010) find that runoffs are frequently associated with two candidate elections.

(1992), point to *voters'* strategic behavior, stressing their unwillingness to squander their votes on "hopeless" candidates. This is also emphasized in Cox (1997), who forcefully argues that votes get concentrated on a small number of candidates in one-seat and simple plurality elections (as opposed to proportional elections and runoffs) because of strategic considerations of voters and political elites. The argument is similar to Feddersen's: people avoid wasting their vote on hopeless candidates if the rule is first-past-the-pole and so does the political elite having to allocate scarce resources to finance and endorse candidacies.

There are two other theoretical reasons as to why runoff elections end up with more candidates when compared to simple plurality contests. First, in a runoff election, a politician does not have to be the first to pass the pole to have a meaningful chance of winning – a fact that increases his willingness to participate. Second, some candidates have incentives to enter the race even if they do not perceive a great chance of arriving among the first. This happens because by garnering a high voters' support in the first round, they may be able to sell their endorsement to the front runners at a higher price in the second round of the election. Clearly, this motivation for entry is absent in simple plurality elections.

On the empirical front, many studies have investigated the existence of a relationship between electoral rules – such as district size and the presence of a runoff stage – and the number of candidates, mostly using cross-country data. In an important contribution, Lijphart (1994) presents systematic evidence supporting the idea that the level of proportionality embedded in the various electoral rules is significantly associated with the degree of multi-candidate competition. Regarding the effect of runoffs, Wright and Riker (1989) use data from Democratic primaries in U.S. States between 1950 and 1982 to investigate whether those held under simple plurality entailed less candidates than the ones employing the two-ballot system. They find that the average number of candidates in plurality primaries was less than 3, whereas the figure for runoff primaries was above 5. Controlling for other factors, such as the participation of the incumbent in the election, the conditional difference in the number of candidates falls to 2 but is still significant.

Wright and Riker (1989) is a highly cited paper supporting Duverger's hypothesis, but it has an important weakness: it implicitly assumes the decision of adopting plurality or runoff to be exogenous. Unfortunately, it is possible that some counties self-selected into these two categories of ballot structure. If unobservable characteristics affect both the decision to opt for

runoff primaries and politicians motivations to enter the race, then the estimated coefficient is not reflecting the pure effect of the voting system on the number of candidates[4].

Jones (1999) also tests the influence of runoffs on the number of candidates using cross-country data from presidential elections. Resorting to different econometric methods and samples, he finds that runoff elections have on average approximately one more candidate than those with simple majority. Nonetheless, this paper also features no exogenous source of variation in the ballot structure variable. Furthermore, the use of cross-country dataset probably aggravates the selection bias problem (since we expect institutional characteristics to vary more strongly among countries than within sub-national units).

Advocates of the "sociological view" criticize Lijphart's "institutional view" arguing that the number of parties/candidates should be related to societal cleavages and not to artificial (or endogenous) institutional design characteristics[5]. As mentioned in the Introduction, currently some consensus is building around a more nuanced view combining both the institutional and the sociological approaches. This hybrid view suggests heterogeneity only matters in (or is magnified by) the presence of permissive electoral rules, such as a runoff stage. Arguably, heterogeneity should not be key in explaining the number of candidates if the voting structure leaves no room for its effects to play out (an idea not fully corroborated by our results). In this vein, Ordeshook and Shvetsova (1994) provide cross-country evidence that it is the interaction of institutional design (district size, in their paper) and social cleavages that matters, not any of each separately. Similarly, Cox and Amorim Neto (1997) find that the number of presidential candidates is positively related to the interaction between a runoff dummy and an index of ethnic fragmentation (also using cross-country data). Importantly, in his paper neither the runoff variable nor ethnic fragmentation are by themselves statistically significant. Using a larger dataset for presidential elections, Golder (2006) finds similar results.

3 Data and Identification Strategy

In Brazil, mayors are elected every four years, with elections been held in the same day in all municipalities. We use data from five of these elections, which took place in 1988, 1992, 1996, 2000 and 2004. In the 1988 election, the prevailing legislation mandated simple plurality for

[4] As discussed in the next section, in our case self-selection is not a concern.

[5] See Cox (1997).

all municipalities across the country, independently of size. But, crucially for our identification strategy, the constitutional reform – approved at the end of that same year – imposed a new rule for later elections: municipalities with more than 200,000 registered voters were required to have a second ballot[6]. Below this threshold, simple plurality rule remained in place. Hence, in all subsequent mayoral elections there have been instances of both simple majority and runoff contests. Table 1 presents the number of municipalities in our sample for all five elections considered.

Table 1: Ballot structure in municipal elections

	1988	**1992**	**1996**	**2000**	**2004**
Plurality	2,536	3,637	5,356	5,502	5,490
Runoff	0	30	47	57	68
Total	2,536	3,667	5,403	5,559	5,558

The panel data structure and the law shift at the end of 1988 allow us to explore both cross-section and time-series variations in order to identify the effect of runoffs and heterogeneity on the number of competing candidates. Identification thus comes from two sources: differences in electoral rules across municipalities in a same election, and differences arising from municipalities changing status between elections.

The electoral data – number of candidates and eligible voters – come from the Superior Electoral Court (TSE) and Regional Electoral Courts (TREs) datasets. Electoral data on all Brazilian municipalities is readily available from TSE only for the 1996, 2000 and 2004 elections. For the remaining two elections, we obtained information from TREs, either through their websites or direct contact. This effort yielded a sample of around 60% of all Brazilian municipalities for the 1988 and 1992 elections. Details are described in the Data Appendix.

Table 2 below presents some summary statistics.

[6]In this case, the election follows the usual format of a runoff: in the first stage, a candidate is elected if he/she gathers more votes than the sum of all other candidates. Otherwise, a second round takes place with the two candidates with most votes in the first round. The winner of the second ballot is finally elected.

Table 2: Summary statistics (averages across municipalities)

	1988	1992	1996	2000	2004	Total
Candidates/municipality	3.07	2.91	2.84	2.71	2.83	2.84
	(1.35)	(1.21)	(1.22)	(1.06)	(1.11)	(1.18)
Electorate/municipality	19,224	17,105	18,352	19,517	21,556	19,317
	(136,577)	(110,852)	(119,652)	(124,387)	(134,808)	(125,281)
Incumbent vying	–	–	–	65.7	42.7	26.5
for reelection (%)				(47.5)	(49.5)	(44.1)

	1991	2000
Gini index	0.53	0.56
	(0.06)	(0.06)

Standard errors in parentheses

Differently from other studies, which resort to ethnico-linguistic measures of social fragmentation, we use the local Gini inequality index as a proxy for social cleavage – data comes from the 1991 and 2000 population Censuses. This choice is based on the fact that inequality in income is arguably the most important indicator of different public policy preferences in a country with extreme disparities in income. Brazil has persistently ranked amongst the worst countries in the world when it comes to income inequality. Furthermore, nearly all the population speaks Portuguese, the country's official language, and there is no history of persistent racial conflicts (such as in Africa or in the U.S.)[7].

In our analysis we also control for electorate size and the presence of an incumbent[8]. The former is important in our context since the rule establishing the second ballot is a discontinuous function of this variable. If for some reason the number of candidate is dependent on electorate size, not adding this variable could lead to omitted variable bias.

Finally, it is worth to emphasize the advantage of working with a set of political units for which the electoral legislation decided on a higher sphere applies uniformly. This renders

[7] A long tradition in Brazilian sociology argues that differences in race were never crucial to explain social tensions within the country (see, for instance, Freyre (1936)).

[8] Reelection was prohibited in Brazil until 2000. A change in legislation then allowed sitting mayors to run for an additional consecutive term.

improbable that other confounding effects are driving the results and, of course, eliminates the possibility of self-selection (this is not the case in the famous Wright and Riker study, for instance)

What does the raw data tell us about the number of candidates in elections with and without a second-ballot? Table 3 below displays unconditional averages of this variable. As can be seen, simple averages seem to support Duverger's Hypothesis: the number of candidates competing in elections where there is a runoff stage is considerably greater than the number of contenders in simple plurality elections. However, as we will show in the next section, econometric analysis suggest the runoff effect operates mainly through its interaction with the heterogeneity measure.

Table 3: Average number of candidates

	1988	**1992**	**1996**	**2000**	**2004**	**Total**
Simple Plurarity	3.07	2.89	2.80	2.67	2.78	2.81
	(1.35)	(1.18)	(1.14)	(0.98)	(1.01)	(1.11)
Runoff	–	5.73	7.11	6.11	6.43	6.39
		(1.84)	(2.19)	(2.49)	(2.28)	(2.29)
Total	3.07	2.91	2.84	2.71	2.83	2.84
	(1.35)	(1.22)	(1.21)	(1.06)	(1.11)	(1.18)

Standard errors in parentheses

4 Results

4.1 Pooled OLS Estimations

We first run a set of pooled OLS regressions controlling for State[9] and time dummy variables. The model estimated thus has the following specification[10]:

$$y_{it} = \beta_0 + \beta_1.D_{it} + \beta_2.H_{it} + \beta_3.D_{it}.H_{it} + \beta_4.E_{it} + S_i + T_t + \epsilon_{it} \quad (1)$$

[9] Each of Brazil's municipalities is a member of one of the 27 States that comprise the federation.

[10] The fact that the existence of runoff depends on a fixed population threshold (population larger than 200,000) calls for the employment of regression discontinuity techniques (as discussed by Lee and Lemieux (2010)). However, since there is not enough mass of municipalities around such threshold, we do not have statistical power to employ such techniques.

Where: y_{it} is the number of candidates in municipality i in election t; D_{it} assumes value 1 if there is a runoff stage and 0 otherwise; H_{it} is the local Gini coefficient; E_{it} is electorate size (in tens of thousands of voters); S_i are State dummies aimed at isolating State-specific characteristics that do not vary over time but may idiosyncratically influence the number of candidates (such as local culture and institutions), and T_t are election-year dummies that capture changes common to all municipalities, such as shifts in party structure and in other dimensions of the national electoral system[11]. For the 1988 and 1992 elections the Gini coefficient employed is the one from the 1991 Census, whereas for the remaining elections the Gini comes from the 2000 Census[12].

The results appear in Table 4 below. In all regressions reported in this paper, standard errors are robust to heteroskedasticity and clustered by municipality, so as to correct for autocorrelation problems. Column (1) shows the result of a simple regression including only the runoff dummy (besides time and State dummies). As one can see the presence of a runoff is associated with an increase of approximately 3.5 candidates per municipal election. Column (2) includes also our heterogeneity indicator in the spirit of the "sociological view". Here, the runoff coefficient changes only slightly, while the positive coefficient on the Gini index seems to support the idea of a positive relationship between heterogeneity and the number of candidates. However, this effect is quantitatively very small: an increase of 10 percentage points in the Gini index is related to only 0.12 more candidates[13].

In column (3) we report the results adding the interaction term ($D_{it}.H_{it}$) that portrays the nuanced view. In this case the effect of heterogeneity on the number of candidates is given by $\beta_2 + \beta_3.D_{it}$. Therefore, if $\beta_3 > 0$, the presence of a runoff raises the effect of heterogeneity on the number of candidates. The coefficient on the interaction term is positive and highly significant. Moreover, this specification reveals that heterogeneity's marginal impact is now meaningful in municipalities featuring a runoff. More precisely, it goes from 1.11 for $D_{it} = 0$ to roughly 18 (=1.11 + 16.9) for $D_{it} = 1$. To put it differently, these initial results suggest heterogeneity is relevant only when a more permissive electoral rule is in place.

[11] Again, national electoral legislation applies commonly to all municipalities.

[12] Our main conclusions are not sensitive to this choice. For instance, results are essencially the same if we assign the 1991 Gini index to the 1988, 1992 and 1996 elections, and the 2000 Gini to the remaining elections.

[13] The Gini index assumes values between zero and one. Since the coefficient on this variable is approximately 1.2 in column (2), an increase in 10 percentage points in the Gini leads to 0.12 more candidates.

Finally, in column (4) we add two additional controls: electorate size and a incumbency dummy. Our main conclusions are robust to the inclusion of these variables. The coefficients on Gini and its interaction with runoff have the same sign as those on column (3) and remain highly significant, although the impact of heterogeneity is somewhat dampened. Additionally, electorate size has positive but quantitatively small effect: an increase of 10,000 voters leads to approximately 0.02 candidates, all else equal. The coefficient of the incumbency dummy, only marginally significant, has an unexpected negative sign[14].

Table 4: Pooled OLS[15]

Dependent Variable: Number of Candidates				
	(1)	(2)	(3)	(4)
Runoff	3.449	3.409	-5.917	-2.175
	(0.237)***	(0.232)***	(1.537)***	(1.290)*
Gini		1.234	1.120	0.986
		(0.187)***	(0.184)***	(0.180)***
Runoff × Gini			16.112	7.495
			(2.829)***	(2.515)***
Incumbent				0.034
				(0.018)*
Electorate/10,000				0.021
				(0.007)***
Observations	22796	22689	22689	22663
R-squared	0.15	0.15	0.16	0.19

Robust standard errors (clustered by municipality) in parentheses

* significant at 10%; ** significant at 5%; *** significant at 1%

In Table 5 we subject our main result to some robustness checks. In column (1) we simply reproduce the findings from column (4) in Table 4 for the sake of comparison. In column (2) only candidates with more than 1% of the votes are computed in the total number of

[14] Nonetheless, when we run our fixed-effects estimations, the coefficient the incumbency dummy becomes negative (see Section 4.2).

[15] For the 1988 and 1992 regressions the Gini coefficient employed is the one from the 1991 Census, whereas for the 1996, 2000 and 2004 regressions the Gini comes from the 2000 Census.

candidates. The objective here is to assess whether the results are being driven by non-representative candidates. As can be seen, even if the point estimate of the interaction term is reduced, it remains statistically significant. In column (3) we follow other papers and use the effective number of candidates (ENC) instead of total candidates as the dependent variable. Algebraically, $ENC = 1/\left[\sum_i s_i^2\right]$, where s_i is candidate i's voting share. This measure takes into account the distribution of votes across candidates. For instance, in the case of two elections with the same number of candidates, the one displaying more unequal vote shares will end up with a lower ENC. Though it is more difficult to interpret the size of the point estimate when the dependent variable is ENC, both the heterogeneity and interaction terms remain positive and significant in this specification.

Columns (4)-(6) are analogous to (1)-(3) except that in the former we restrict our sample to municipalities for which we have data for all five elections. We do this to check if our results are being influenced by composition effects (changes in the sample of municipalities across time). This does not seem to be the case since the new estimations yield similar results.

Table 5: Pooled OLS

	Full Sample			Restricted Sample		
	number of candidates	candidates with >1% of votes	effective number of candidates	number of candidates	candidates with >1% of votes	effective number of candidates
	(1)	(2)	(3)	(4)	(5)	(6)
Runoff	-2.175	-1.290	-0.725	-1.422	-1.887	-1.132
	(1.290)*	(1.054)	(0.634)	(1.348)	(1.174)	(0.851)
Gini	0.986	0.814	0.273	1.151	0.949	0.118
	(0.180)***	(0.162)***	(0.090)***	(0.293)***	(0.267)***	(0.149)
Runoff × Gini	7.495	4.862	2.130	6.393	6.083	2.939
	(2.515)***	(1.998)**	(1.139)*	(2.406)***	(2.066)***	(1.488)**
Incumbent	0.034	0.026	0.001	0.055	0.026	-0.014
	(0.018)*	(0.017)	(0.011)	(0.030)*	(0.028)	(0.017)
Electorate	0.021	0.008	0.003	0.017	0.006	0.002
	(0.007)***	(0.005)	(0.002)*	(0.005)***	(0.004)	(0.001)
Observations	22663	22619	22619	11360	11360	11360
R-squared	0.19	0.12	0.08	0.17	0.10	0.07

Robust standard errors (clustered by municipality) in parentheses

* significant at 10%; ** significant at 5%; *** significant at 1%

4.2 Fixed-Effects Estimations

In this section we present the results from panel regressions with the following specification:

$$y_{it} = \beta_0 + \beta_1.D_{it} + \beta_2.H_{it} + \beta_3.D_{it}.H_{it} + \beta_4.E_{it} + F_i + T_t + \epsilon_{it} \quad (2)$$

Where F_i is the unobservable time-invariant characteristic of municipality i.

The gain from using fixed-effects relative to pooled regressions is twofold. First, the inclusion of the fixed effect eliminates the possibility of omitted variable bias due to time invariant characteristics, like geography, culture, institutions at the municipality level[16]. Secondly,

[16] See Wooldridge (2002).

within a panel framework we are able to fully explore variations in status that happen through time in the same municipality (whereas OLS estimations treat different observations of the same municipality as independent). This leads to a sharper identification of the parameters of interest[17].

As can be seen in the Table 6, the previous main result continues to hold in this new setting: the interactive variable is always strongly statistically significant and has a positive sign. Moreover, its impact is magnified with the inclusion of fixed effects. When a runoff is in place, an increase of 10 percentage points in the Gini index is associated with nearly two more candidates – for columns (1) and (4) where the dependent variable is the number of candidates. Notice that the coefficient on Gini now becomes insignificant and often times negative. Nonetheless, its magnitude remains quite small, that is, heterogeneity does not seem to affect the number of candidates in the absence of runoffs. Finally, incumbency now has the "expected" sign: an incumbent running for re-election is associated with a smaller number of candidates. This effect is nonetheless immaterial.

[17] As far as we know, no empirical work testing the influence of heterogeneity and runoffs on the number of candidates controls for fixed effects.

Table 6: Fixed-effects

	Full Sample			Restricted Sample		
	number of candidates	candidates with >1% of votes	effective number of candidates	number of candidates	candidates with >1% of votes	effective number of candidates
	(1)	(2)	(3)	(4)	(5)	(6)
Runoff	-11.329	-5.971	-3.601	-12.031	-6.907	-3.687
	(2.466)***	(2.350)**	(1.093)***	(3.097)***	(2.610)***	(1.305)***
Gini	0.118	-0.129	-0.246	-0.175	-0.180	-0.353
	(0.259)	(0.246)	(0.156)	(0.356)	(0.342)	(0.211)*
Runoff × Gini	19.295	10.409	6.461	20.757	12.372	6.670
	(4.401)***	(4.264)**	(1.993)***	(5.307)***	(4.548)***	(2.344)***
Incumbent	-0.057	-0.060	-0.058	-0.028	-0.053	-0.063
	(0.020)***	(0.018)***	(0.012)***	(0.031)	(0.029)*	(0.018)***
Electorate	-0.008	-0.014	-0.012	-0.004	-0.010	-0.013
	(0.014)	(0.015)	(0.007)*	(0.013)	(0.014)	(0.008)
Observations	22663	22619	22619	11360	11360	11360
R-squared	0.60	0.54	0.46	0.57	0.52	0.43

Robust standard errors (clustered by municipality) in parentheses

* significant at 10%; ** significant at 5%; *** significant at 1%

4.3 Sub-Samples According to Electorate

Our analysis so far has been based on the comparison of two fairly distinct groups. Those municipalities for which $D_{it} = 1$ are much bigger (more than 200,000 voters) than those with $D_{it} = 0$, but account for approximately only 1% of the sample. Even if we control for electorate size, there remains the possibility of omitted variables influencing the results when we lump together very dissimilar municipalities. To allay these concerns, in this subsection we replicate our panel regressions using subsamples comprising only larger municipalities. Table 7 displays

the results for the number of candidates as dependent variable[18]. To facilitate comparison, column (1) exhibits results for the full sample, whereas columns (2)-(4) report results for sub-samples of municipalities with at least 50,000, 100,000 and 150,000 voters, respectively.

Table 7: Fixed-effects (sub-samples according to electorade)

Dependent Variable: Number of Candidates				
	(1)	(2)	(3)	(4)
Runoff	-11.329	-10.707	-10.581	-10.274
	(2.446)***	(2.552)***	(2.578)***	(2.697)***
Gini	0.118	9.469	23.147	22.467
	(0.259)	(4.451)**	(10.311)**	(13.514)*
Runoff × Gini	19.295	18.609	18.499	18.008
	(4.401)***	(4.520)***	(4.618)***	(4.841)***
Incumbent	-0.057	0.033	-0.134	-0.233
	(0.020)***	(0.143)	(0.237)	(0.337)
Electorate/10,000	-0.008	0.005	0.007	0.009
	(0.014)	(0.007)	(0.006)	(0.006)
Electorate larger than	0	50,000	100,000	150,000
Observations	22663	1174	526	328
R-squared	0.60	0.63	0.65	0.68

Robust standard errors (clustered by municipality) in parentheses

* significant at 10%; ** significant at 5%; *** significant at 1%

Our main conclusions are robust to these modifications in sample size. Despite the drastic reduction in the number of municipalities, the coefficient on the interaction term remains highly statistically significant, with little change in magnitude. Interestingly, the coefficient on the Gini index alone turns statistically significant and rises in magnitude as the sample is curtailed to include only larger municipalities. In particular, in columns (3) and (4), an increase in 10 percentage point in the Gini index is associated with approximately 2.2 more candidates even

[18] Results are similar for the other two dependent variables considered before (candidates with at least 1% of total votes and effective number of candidates), except for the fact that the coefficient on the Gini index becomes insignificant in column (4).

in the absence of a runoff[19]. So when we consider only municipalities less disparate in size, the data reveals some support also for the pure sociological view, that is, heterogeneity seems to matter whether a runoff is in place or not.

The evidence favoring the nuanced view remains unchallenged: as can be seen in all estimations – columns (1)-(4) – the presence of a runoff importantly magnifies the impact of heterogeneity.

5 Conclusion

Theory suggests the number of candidates should vary with the degree of societal heterogeneity and the nature of electoral rules. In particular, the so-called nuanced view argues that electoral permissiveness should intensify the role played by heterogeneity. We present evidence supporting this line of reasoning. More specifically, we find that in Brazilian municipalities where a runoff is mandatory heterogeneity influences more heavily the number of candidates vying for mayor. This result is robust to different econometric methods (OLS, fixed effects) and to changes in the sample of municipalities. Since we resort to panel data estimations and have a truly exogenous source of variation in electoral rules coming from the constitutional change of 1988, we believe to be better able to deal with the sort of endogeneity problems that affect most of articles in this literature using cross-section estimations.

References

[1] Bouton, L. 2010. "A Theory of Strategic Voting in Runoff Elections", Mimeo.

[2] Callander, S. "Duverger's Hypothesis, the Run-Off Rule, and Electoral Competition", *Political Analysis* 13: 209-232.

[3] Cox, G. 1987. "Electoral Equilibria Under Alternative Voting Institutions", *American Journal of Political Science* 31: 82-108.

[4] Cox, G. 1990. "Centripetal and Centrifugal Incentive in Electoral Systems", *American Journal of Political Science* 34: 903-935.

[19] The reduction in sample size renders the coefficients on the incumbent dummy and the electorate size insignificant, though.

[5] Cox, G. 1997. *Making Votes Count*. Cambridge University Press.

[6] Cox, G.; Amorim Neto, O. 1997. "Electoral Institutions, Cleavage Structures and the Number of Parties", *American Journal of Political Science* 41: 149-174.

[7] Duverger, M. 1954. *Political Parties*. New York, Wiley.

[8] Fedderesen, T. 1992."A Positive Model Implying Duverger's Law and Positive Turnout", *American Journal of Political Science* 36: 938-962.

[9] Freyre, G. 1936. *Casa-Grande e Senzala*. 2006 Edition, Global Editora.

[10] Golder; M. 2006. "Presidential Coattails and Legislative Fragmentation", *American Journal of Political Science* 50: 34-48.

[11] Jones, M. 2004. "Electoral institutions, social cleavages, and candidate competition in presidential elections", *Electoral Studies* Vol. 23: 73-106

[12] Jones, M. 1999. "Electoral Laws and the Effective Number of Candidates in Presidential Elections", *The Journal of Politics*, Vol.61, 171-184.

[13] Lee, D.; Lemieux, T. 2010. "Regression Discontinuity Designs in Economics", *Journal of Economic Literature* Vol. 48: 281–355

[14] Lijphart, A. 1994. *Electoral Systems and Party Systems: A Study of 27 Democracies*. Oxford, Oxford University Press.

[15] Ordeshook, P.; Shvetsova, O. 1994. "Ethnic Heterogeneity, District Magnitude and the Number of Parties", *American Journal of Political Science* 38: 100-123.

[16] Osborne and Slivinski. 1996. "A Model of Political Competition with Citizen Candidates", *Quarterly Journal of Economics* 111: 65-92.

[17] Wooldridge, J. 2002. *Introductory Econometrics: A Modern Approach*. South-Western College.

[18] Wright, S.; Riker, W. 1989. "Plurality and Runoff Elections and Numbers of Candidates", *Public Choice* 60: 155-176.

Data Appendix

This appendix provides further details on the data used in the paper.

Electoral data: for 1996, 2000 and 2004, election results for all Brazilian municipalities are available at the webpage of the national electoral authority[20] – *Tribunal Superior Eleitoral* (TSE). For the 1988 and 1992 elections, we obtained data from State electoral authorities – *Tribunais Regionais Eleitorais* (TREs) – either through their websites or direct contact. For these two elections, we have information on a subset of about 60% of all Brazilian municipalities.

In what follows, we present the list of States for which we have information on the number of candidates regarding the 1988 and 1992 elections:

1988 election: Alagoas, Amapa, Ceara, Mato Grosso do Sul, Minas Gerais, Paraiba, Pernambuco, Piaui, Rio de Janeiro, Rio Grande do Norte, Rondonia, Santa Catarina, Sao Paulo, and the municipality of Aracaju (capital of the State of Sergipe).

1992 election: Alagoas, Amapa, Amazonas, Bahia, Minas Gerais, Mato Grosso, Mato Grosso do Sul, Para, Paraiba, Pernambuco, Piaui, Rondonia, Rio Grande do Sul, Santa Catarina, Sergipe and Sao Paulo.

In most of the cases above, we have information on electoral results, enabling us to calculate the three dependent variables used in this paper – number of candidates, number of candidates with at least 1% of total votes, and effective number of candidates. For a small number of municipalities, we were unable to calculate the last two measures, since we have only information on the set of candidates and the winner.

Finally, electorate data was obtained from TSE's webpage, for all elections considered here.

Inequality data: the municipal Gini index is based on microdata from the population Censuses of 1991 and 2000, collected by *Instituto Brasileiro de Geografia e Estatistica* (Brazilian Institute for Geography and Statistics). Data were gathered from the webpage of Instituto Brasileiro de Pesquisa Economica Aplicada (Institute for Applied Economic Research)[21]. In our regressions, we assign the 1991 Gini index to the 1988, 1992 and 2000 elections, and the 2000 Gini index to the remaining elections.

[20] www.tse.gov.br.

[21] www.ipeadata.gov.br.

Presence of incumbent: reelection was prohibited in Brazil until 1997. A change in legislation then allowed mayors, state governors and the president to run for an additional term. Nevertheless, neither TSE nor TREs have information on whether the incumbent is participating in an election. To assess this information, we used the names of candidates in each election and checked if the winner of a certain election was participating in the subsequent one.