6.Identification Strategy

In order to understand the empirical strategy, it is important to remember that the sample is restricted to municipalities where there was at least one dynastic candidate. y_i is defined as a general political outcome in the 2005 -2008 mandate[10] in the municipality where the dynastic candidate i ran for election, D_i as a dummy variable, which is one if the dynastic candidate is elected and 0 otherwise and $\mathcal{E}_i$ is the error term. The simplest way to estimate these effects would be to run the following regression, using the Ordinary Least Squares Method (OLS):

$$y_i = \alpha + \rho D_i + \beta X_i + \mathcal{E}_i \qquad (2)$$

Where X_i is a matrix with observables that potentially affect the outcomes used (y_i). Nonetheless, the OLS estimator of the dynastic effect ($\hat{\rho}_{OLS}$) will be biased for the causal effect of dynasties (ρ) due to the simultaneous relationship between the victorious dynastic candidates ($D_i = 1$) and outcomes used (y_i) and because of omitted variables correlated with D_i.

It is plausible, for instance, that dynasties' political influence is correlated with D_i at the same time that it is responsible for some of the variation in the capacity of the mayor to obtain Discretionary Transfers' with the federal government (which serve as a measure of mayor's effort). Moreover, it is equally likely that dynastic candidates' victories not only affect outcomes, but are also affected by these political and socioeconomic factors. The apparatus size, besides being a potential result of political dynasties, may help these families to perpetuate themselves in power, for example.

To identify political dynasties' effects, an exogenous variation of dynastic mayors is necessary. The ideal experiment would be to randomize among the Brazilian municipalities dynastic and non-dynastic mayors. Since this is not possible, we exploit a

[10] The exceptions are the reelection rate and the campaign spending that refers to the 2008 elections.

quasi-experiment that randomly assigns dynastic mayors among these municipalities, following Lee, Moretti and Butler (2004). In particular, the focus is on electoral races in which dynastic candidates barely won or lost. Maintaining the same structure described in the above discussion, consider the Brazilian electoral rule which establishes that:

$$D_i = \begin{cases} 1, & if\ x_i > 0 \\ 0 & if\ x_i < 0 \end{cases} \qquad (3)$$

Where x_i is the vote margin defined as the difference between the number of votes received by the dynastic candidate i in the 2004 elections and the main opponent's number of votes over the total number of valid votes. Since Brazilian municipalities with more than 200,000 inhabitants have their mayors elected in two rounds, the sample is restricted to the last round of each municipality, so that in every election the treatment is defined as described in equation (3). Assuming that the relation between x_i and y_i can be described by the function $f_1(.)$ before the discontinuity and by $f_2(.)$ after it, the effect of political dynasties can be consistently estimated by the OLS regression of the following equation:

$$y_i = \alpha + \rho D_i + f_1(x_i) + D_{it} * f_2(x_i) + \beta X_i + \mathcal{E}_i \qquad (4)$$

It is important that the functions of equation (4) are properly estimated. For this reason, this exercise estimates three different specifications for each regression using linear, quadratic and cubic polynomials for both functions. For each of these specifications, the sample was restricted to candidates that won or lost by less than 60% the valid votes[11]. Furthermore, non-parametric estimations are also presented using the optimal bandwidth of Imbens and Kalyanaraman (2009), and bandwidths twice larger and half smaller than that one.

The identification assumption is that the distribution of municipal pre-determined characteristics is smooth in the discontinuity (Hahn, Todd and Van der Klaauw, 2001). As a consequence, municipalities where a dynastic candidate barely won should be similar in all characteristics to those in which the dynastic contender barely lost. We test

[11] We also estimate the same specifications restricting the sample to vote margins narrower than 30% and 10%. They are available upon request.

this assumption estimating regressions as the one described in equation (4), but using pre-determined variables in the place of the outcomes (y_{it}). If this assumption is correct, municipalities in both sides of the zero vote margin should present similar characteristics. However, when this test is performed with the entire sample, we find that municipalities where the dynastic candidate won with a small margin seem less educated, with lower population density and poorer than municipalities where these candidates lost by a small margin.

This discontinuity appears only in the 1996 and 2000 elections, as it can be seen in Table 4 and Table 5[12]. The absence of a discontinuity in all the observable characteristics, as presented in Table 6, supports the identification assumption for the 2004 data. This table presents the result of the non-parametric regressions using the optimal bandwidth (column 1), a bandwidth 50% smaller (column 2), and another one twice as large (column 3). Out of the 13 variables analyzed, only two are marginally significant in the first specification. The results using alternative parametric specifications lead to a similar conclusion.

It is important to note that the treatment effect identified is local in two different ways. First, due to the experiment design, the sample is restricted to the municipalities where at least one dynastic candidate ran for office. Second, since the exploited discontinuity is in the zero margin, the effect is identified in municipalities with high political competition.

[12] The discontinuities are less robust in the 1996 sample in the non-parametric estimations. However in the parametric specifications, the results are similar to the ones of 2000.