

Cash-transfers and voting behavior: An empirical assessment of the political impacts of the Bolsa Família program.*

Cesar Zucco

Woodrow Wilson School & Department of Politics

Princeton University

zucco@princeton.edu

This version: August 21, 2009

Introduction

In the 2006 Brazilian election, incumbent president Lula obtained a sweeping majority of the votes in the less developed areas of the country, reversing a two decade personal and party history of performing much better in more developed regions. However one looks at it, there is not doubt that these elections represented a significant change in Lula's electorate (Zucco Jr. 2008, Nicolau & Peixoto 2007, Hunter & Power 2007). Higher socioeconomic level had been shown to be a strong predictor of party identification with the PT (Samuels 2006), and until very recently the PT was all but absent from the most backward regions of the country. How, then, could such a radical shift happen in such a short time?

In this paper, we explore the argument that a considerable part of this change in voting patterns was due to a large scale cash transfer program — Bolsa Família — implemented by the Lula government in his first term. Though this hypothesis finds support in the above mentioned literature, it has been questioned by Carraro, Araújo Jr, Damé, Monasterio & Shikida (2007) who have attributed Lula's new constituency to the effects of pro-poor economic growth. It has also been noted that incumbent party candidates historically do better in poorer municipalities, and this incumbency effect is yet another confounding factor at work that helps complicate the analysis (Zucco Jr. 2008). For this reason, this paper takes the empirical analysis of the effects

*This version was prepared for delivery at the Annual Meeting of the American Political Studies Association, Toronto, September 2009. Work in progress. The author thanks Joseph Wright, Irineu Carvalho Filho, Nelson Souza Sobrinho, and Maurício Canêdo Pinheiro for insights and suggestions.

of the program one step further by combining inference in aggregate data, ecological analysis techniques, and individual level survey data.

We start with a quick presentation of previous evidence in support for the Bolsa Familia hypothesis that relies basically on aggregate data at the municipal level. We then apply standard ecological inference techniques to the data and present revised estimates of the program’s electoral effects. Subsequently, we attempt to separate the effects of the Bolsa Familia into its direct effect over beneficiaries and the indirect effects it has as a economic stimulus. Collectively, these analyses suggest that Lula received more votes wherever the program had greater scope; that this effect remains after controlling for economic growth; that the program has indirect effects channeled through economic growth; that BF beneficiaries voted for Lula at a higher rate than non-beneficiaries; and that this difference is larger in more well off localities.

While these results are intuitive, it is important to note they rely on indirect comparisons between beneficiaries and non-beneficiaries of the program. Given that these two groups differ significantly in many other respects besides membership in the program, this is probably not the best conceptualization of a “Bolsa Familia electoral effect.” For this reason, we spend some time defining an alternative concept of this effect, and discuss research designs that could be used to evaluate it. This discussion is followed by an attempt to implement one such design using a not-too-reliable survey data that provide a preliminary estimate of the effect of the electoral effect of receiving the Bolsa Familia measured as a comparison between otherwise similar individuals. We then proceed to a cursory exercise that attempts to reconcile both sets of results and which finds that they are, in fact, of comparable magnitude. We conclude the paper with a brief digression into how these results can inform a theoretically grounded research agenda on the subject.

1 Lula’s Shifting Vote-Base

Lula’s departure from his previous voting patterns — the basic phenomenon to be explained — is relatively uncontroversial and can be easily spotted at the individual level using survey data from 2002 to 2006. Figure 1(a) reflects what was considered to be Lula’s traditional constituency, as the probability of voting for him was highest in the middle income brackets. The 2006 figure, on the other hand, displays Lula’s “new” constituency, where he fares significantly better among those with lowest family income.

One important factor between the two elections was the *Bolsa Familia*, a massive cash transfer program implemented and maintained by the federal government. It is the main component

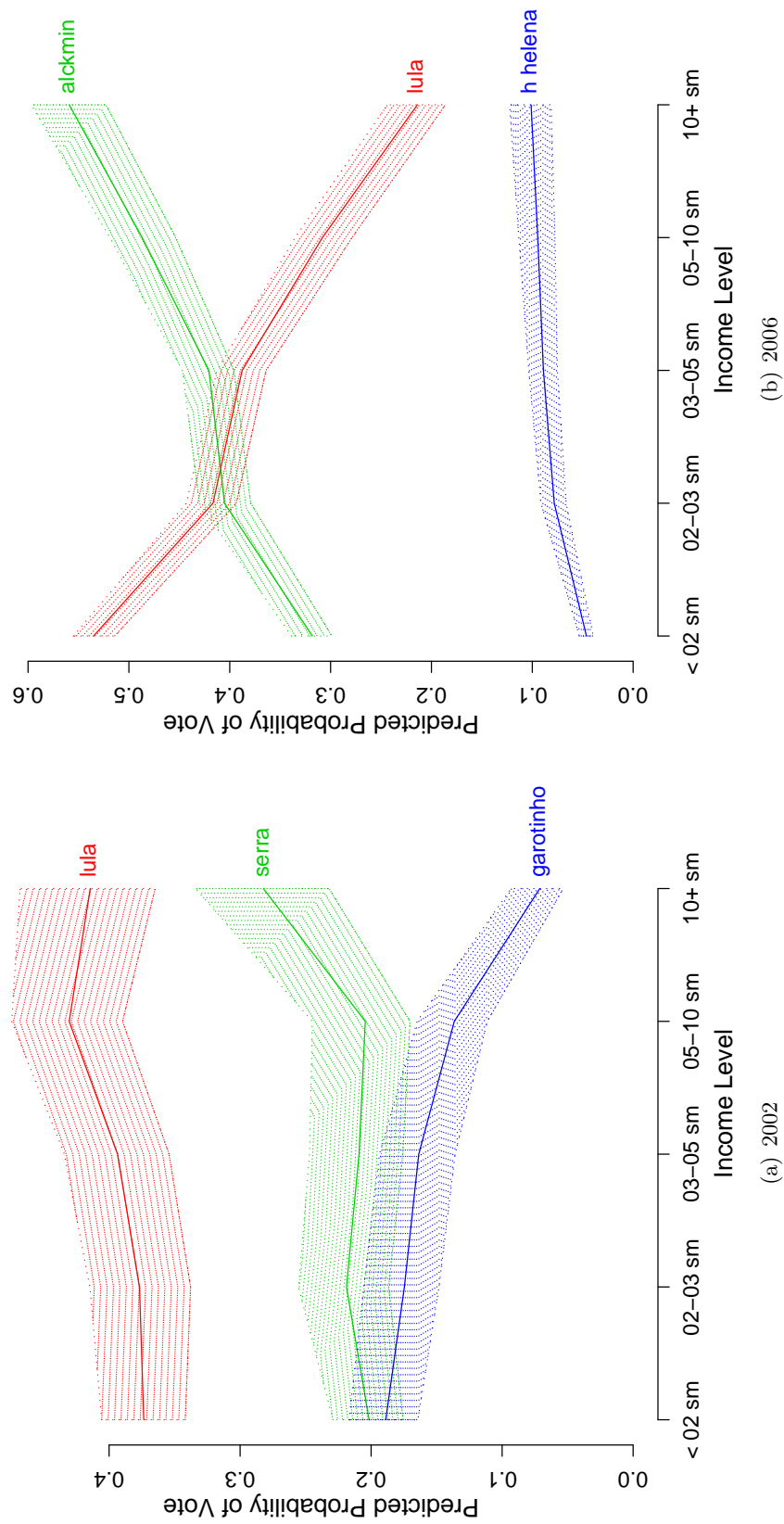


Figure 1: Vote for Lula and Income

Notes: Figures show the predicted probability of voting for each of the main presidential candidates by family income bracket, expressed in minimum wages (sm). Data is from the last pre-electoral Datafolha survey prior to the first round of each election. Figures were estimated by a multinomial logit regression of vote intention on income and a series of other socioeconomic variables — which are held to their median categories for presentation.

of a larger umbrella program called *Fome Zero* (Zero Hunger) and reaches families with monthly income of up to R\$ 120 (just over US\$ 60). Most of its benefits depend on the number of children in the household, and are conditional on them attending school as well on keeping immunization records and maintaining a schedule of visits to the doctor. Extremely poor families also receive a flat benefit on top of the per child one, in which case total benefits can add up to just under R\$ 100 per family.

Table 1 reproduces results from Zucco Jr. (2008).¹ The dependent variable here, which will be used in later portions of the text as well, is the (log of the) scope of the Bolsa Familia program in each municipality.² To compute this measure we obtained the number of families covered by the program as of October 2006, the month of the election,³ and compared it to the total number of households in each municipality computed using IBGE’s data on population and on the average size of household in the state which the municipality belonged to. Alongside the obvious dependent variables such as Lula’s vote share in the first round of the 2002 presidential election, and the HDI-M — Human Development Index at the Municipal Level — for the year 2000, we included others that capture both political and socio-demographic factors at work. Political variables included dummies indicating whether the mayor of the municipality elected in 2004, or the governor of the state elected in 2002, were from the PT, and the vote share obtained by the PT backed candidate in the 2004’s mayoral election. Socio-demographic variables included the log of the municipalities’ population, as well as variables that capture potential racial and religious cleavages, namely the proportion of non whites and pentecostal Christians in each municipality.⁴

Regardless of the controls used, the scope of the *Bolsa Familia* has a positive, significant and substantially relevant effect on Lula’s vote share. The negative association between HDI-M and Lula’s vote share and the effect of Lula’s previous vote-share are also quite stable.⁵ Though a full interpretation of these results transcends the scope of this paper, it is worth

¹Differently than in the original source of these results, variables are here expressed in logs in order to facilitate interpretation.

²There are just over 5500 municipalities in the country. The median population is just over 10 thousand people.

³These data are available from the Ministry of Social Development’s website, which also publishes the actual roll of recipients.

⁴All data was obtained from publicly available online sources, namely Brazil’s Superior Electoral Court (TSE), the Applied Economics Research Institute (IPEA), and from the National Geographical and Statistical Institute (IBGE). Detailed description of the variables and sources, as well as the complete data set is available from the author upon request.

⁵The use of a bounded dependent variable such as vote-share of a candidate can pose problems to a standard OLS regression. However, in the present case, Lula’s vote share across municipalities is quite symmetric and presents very few extreme values: only 3% of the observations fall outside of the 0.15—0.85 range, and less than 0.5% fall outside the 0.05—0.95 range. This fact, coupled with the ease of interpretation, make OLS a feasible and attractive option for the problem at hand.

mentioning that political variables indicate that the presence of a PT governor and or mayor actually decrease Lula's vote share and that Lula performed better in municipalities with larger shares of non-white population, and worse were Pentecostals following is greater.

Part of the association between the HDI-M and Lula's vote share in 2006 seems to have been channeled through the *Bolsa Familia* program: municipalities with lower HDI-M received greater program coverage, which boosted support for Lula. However, the regressions shown indicate that there is also a *direct* effect of HDI-M on the vote for Lula in 2002, that persists even after accounting for the *Bolsa Familia*. Given Lula's electoral record of traditionally performing better among the urban middle class, how can one explain this positive *direct* effect of HDI-M on Lula's 2006 vote share?

The inclusion of variables that attempt to capture the lack of economic alternatives that characterizes much of the less developed areas in the country provides a hint of an answer. These variables are the (log of) municipal GDP per capita, a measure of the relative size of the public sector in the municipality's economy, and the proportion of the municipality's operational revenue that is raised locally through tax, as opposed to received through transfers from the federal or state governments. Results suggest that Lula tended do better where the public sector represents a larger share of the economy, and where the local government is *less* reliant on locally raised taxes (and more reliant on transfers).

As the reader will have already realized, reliance on the Federal Government is a characteristic that cannot change as fast as the shift in Lula's vote base did. Therefore, if this argument is true, it implies that the less developed regions of the country should *always* tend to vote for the government candidate. This, in fact, has been shown to be the case (Zucco Jr. 2008), with the additional caveat that while it is always true for the least developed "places", it does not seem to be necessarily true for the poorest voters. The conclusion is that while it looks like BF played a role in this shifting vote base, there is also evidence that an "incumbency" factor is at work.⁶ With that said, the rest of this paper focuses on better defining the former, leaving the latter for future work.

⁶For additional evidence supporting this point, see Canêdo-Pinheiro (2009).

Table 1: OLS Regressions for Lula’s 2006 Vote Share by Municipality

	(1)	(2)	(3)	(4)	(5)	(6)
log(HDI)	-1.695	-1.190	-1.144	-1.001		
SE	0.043	0.046	0.050	0.064		
p-value	0.000	0.000	0.000	0.000		
log(Lula 2002)	0.297	0.281	0.294	0.311	0.266	0.227
	0.011	0.010	0.011	0.013	0.013	0.012
	0.000	0.000	0.000	0.000	0.000	0.000
log(Scope)	0.170	0.171	0.166	0.147	0.205	0.321
	0.007	0.008	0.009	0.011	0.010	0.007
	0.000	0.000	0.000	0.000	0.000	0.000
log(Nonwhite)		0.106	0.115	0.111	0.145	
		0.006	0.006	0.007	0.007	
		0.000	0.000	0.000	0.000	
log(Pentecostal)		-0.027	-0.030	-0.023	-0.037	
		0.004	0.005	0.005	0.005	
		0.000	0.000	0.000	0.000	
log(Population)		0.040	0.037	0.047	0.049	
		0.003	0.003	0.004	0.004	
		0.000	0.000	0.000	0.000	
log(PT 2004)			0.034	0.035	0.036	
			0.016	0.019	0.020	
			0.039	0.068	0.070	
PT Governor			-0.098	-0.099	-0.066	
			0.014	0.016	0.016	
			0.000	0.000	0.000	
PT Mayor			-0.037	-0.032	-0.039	
			0.012	0.013	0.014	
			0.002	0.017	0.005	
log(GDPpc)				-0.010	-0.067	
				0.017	0.017	
				0.531	0.000	
log(PublicSector)				0.033	0.026	0.135
				0.018	0.018	0.007
				0.068	0.159	0.000
Local Taxes				-0.034	-0.051	
				0.007	0.007	
				0.000	0.000	
Const.	-1.621	-1.796	-1.730	-1.732	-1.618	-1.352
	0.017	0.043	0.048	0.074	0.076	0.031
	0.000	0.000	0.000	0.000	0.000	0.000
R ²	0.597	0.649	0.639	0.651	0.627	0.514
N	5501	5464	4576	3404	3404	5554

Notes: The dependent variable is (log) of Lula’s proportion of votes in each municipality in the first round of the 2006 presidential election. P-values are shown below the estimates. The main variable of interest is the (log of) scope of *Bolsa Familia*, which is the proportion of families in each municipality included in the program. Data set is available from author.

2 Beyond Aggregate Data

The evidence shown above is clear to point out that Lula’s constituency changed between 2002 and 2006, but the argument that this change was related to the Bolsa Familia relies mostly on aggregate data. Even though these data cover more than 5000 generally small municipalities (the size of the median municipality is just over ten thousand), it is fundamental to keep in mind that this analysis does not allow us to infer individual level behavior. We cannot, for instance, say whether the beneficiaries of the program are the ones in fact voting for Lula. Nor can we say that those voting for Lula are doing so because of the program. In this sense, there might not be much difference between the voting patterns of recipients and non-recipients of transfers, and it is even possible — though not very likely — that non recipients support Lula more than recipients.

Short of obtaining individual level data, there is no absolute fix for this problem. There exist, however, techniques that can be applied to aggregate data to reduce the ecological inference problem. In the world of ecological inference, the known aggregate quantities are the marginals of a contingency table, and the inner cells are the unknowns. In the present case, each municipality is represented by its own 2x2 contingency table, dividing voters between those that voted for Lula and those that did not, and those that are beneficiaries of the BA and those that are not. The marginals of these tables are known, and while the inner cells are strictly not observable, the marginals can provide informative bounds on their values. Current ecological inference techniques combine this deterministic information contained in the data, using different statistical models.⁷

Table 2 presents aggregate (national level) estimates of how Lula fared among beneficiaries and non beneficiaries of the *Bolsa Familia*. These partial tables were constructed from the estimates obtained at the municipal level. In the first table, we ran the ecological inference procedures on all 5500 municipalities in the data set at once. In the second table the interior cells were estimated using subsets of the data broken in quintiles according to their levels of development, and the estimates were combined back so that the results reported include estimates for every municipality.

To understand why the results vary according to how the data are aggregated, it is important to understand that careful analysis of the data using tomography plots (omitted for shown for brevity) show that municipalities seem to draw from different distributions according to their

⁷Most of the statistic component of these techniques are parametric (King 1997), but in this paper we make use of a recently developed non-parametric method (Imai, Lu & Strauss 2008).

levels of development. If one estimates the inner cells of each municipal contingency table using the whole set of data, the many poorer municipalities for which there is greater certainty about Lula’s vote share among beneficiaries, pull the estimate of that quantity up.

If, as we discuss below, one expects the quantities of interest to vary according to level of development, it makes sense to estimate the inner cells of the contingency tables using a population of roughly similar characteristics. While we opted to break the municipalities into five different groups, the aggregate estimates are basically unaltered once municipalities are broken into at least four groups. Regardless of how the data are aggregated, one can say that Lula fared better among beneficiaries of the program. This result holds nationally, in all states and at the municipal level, and we can be reasonably certain that this result holds in a majority of the municipalities.⁸

Table 2: Aggregate Voting Patterns

	(a) Full Sample		(b) by HDI Quintiles		
	BA	Not BA		BA	Not BA
Lower Bound	0.89	0.38	l.b.	0.80	0.40
Lula’s vote share	0.91	0.38	Lula	0.82	0.40
Upper Bound	0.92	0.39	u.b.	0.86	0.41
Total	0.19	0.81	Total	0.19	0.81

Notes: Tables show the point estimates, as well as the bounds of the 95% conf. intervals on Lula’s performance among recipients and non recipients of benefits from the Bolsa Familia program.

More interesting than this national effect, however, is the variation across municipalities that it hides. While Lula always does better among beneficiaries, the size of this effect is larger in the more developed municipalities. More to the point, the analysis that follows shows that this result is driven mostly the variation in Lula’s support among those voters that are *not* beneficiaries of the program.

In Table 3, we show the results from a pair of seemingly unrelated regressions fit to both the partitioned data sets. The two dependent variables are the quantities of interest — namely Lula’s vote share among beneficiaries, and among non beneficiaries of the program — which are logit transformed and then regressed on the *same* set of explanatory variables.⁹ This set consists

⁸While the number of municipalities in which this positive electoral effect of the Bolsa Familia program is significant varies depending on the aggregation method (5515 and 2642 for the full sample and the development-partitioned set respectively) there is no municipality in which Lula performed significantly *worse* among beneficiaries than among non-beneficiaries.

⁹The logit transformation was required because the distribution of Lula’s vote share among beneficiaries was highly skewed and contained many observations above 0.9. Because of this transformation, caution should be taken in interpreting the coefficients.

of a series of controls (for political and social characteristics) and two variables of substantive interest: level of socio-economic development and the level of inequality (Gini coefficient).

Table 3: Voting for Lula Among Recipients and Non Recipients, By Municipality

State	Development	
	$f\left(\frac{Lula}{BF}\right)$	$f\left(\frac{Lula}{NotBF}\right)$
log(HDI)	0.22	-3.17
SE	0.14	0.12
p-value	0.11	0.00
log(GINI)	0.21	-0.45
	0.10	0.09
	0.04	0.00
PT Governor	-0.19	-0.14
	0.04	0.04
	0.00	0.00
PT Mayor	-0.04	-0.05
	0.04	0.03
	0.24	0.10
log(Nonwhite)	0.18	0.22
	0.02	0.02
	0.00	0.00
log(Pentecosals)	-0.07	-0.08
	0.01	0.01
	0.00	0.00
log(Population)	0.05	0.13
	0.01	0.01
	0.00	0.00
log(Scope of BA)	0.03	-0.07
	0.03	0.02
	0.32	0.00
Constant	1.13	-3.90
	0.13	0.12
	0.00	0.00
N	5460	5460
R2	0.06	0.41

Notes: The model was estimated using ecological inference development-partitioned estimates. The dependent variables are the logit transformed Lula’s vote share among recipients and non recipients of the *Bolsa Familia*. These values are themselves estimates, obtained by applying ecological inference techniques to municipal level data.

The most important result is that level of development has *no impact* on Lula’s vote share among beneficiaries while being substantially negatively associated with Lula’s vote share among non beneficiaries.¹⁰ Our initial interpretation of these results is that they capture positive externalities generated by the program, which are much more apparent in the less developed

¹⁰The use of the SUR setup allows to test linear hypotheses about estimates in different equations. In the present case, we can confidently reject the null that the coefficients on HDI are the same in both equations in each data sets. A complete table with the results of such tests for all variables is reported in the Appendix.

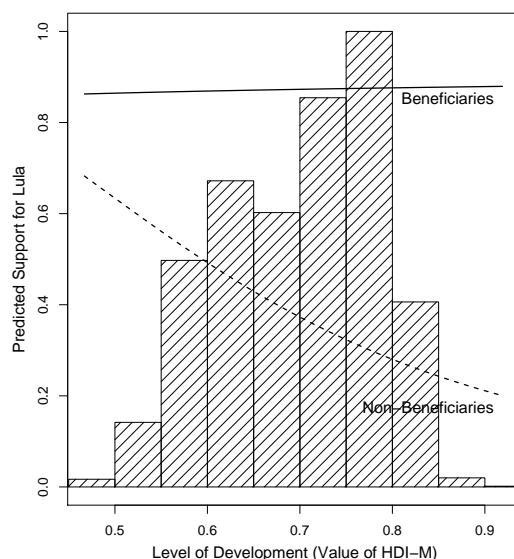


Figure 2: Predicted Support for Lula Given Level of Development

Notes: Predicted values were computed from results using the development-partitioned data set. Histogram in the background depicts the distribution of the HDI-M variable.

regions. Note that this result holds regardless of whether the scope of the program itself (which is the marginal on the original 2x2 table) is included as a regressor, as is the case with the regressions that are actually reported. This suggests that the positive externalities are not due only to having the large proportion of people on the program’s payroll.

The fact that Lula’s vote share among recipients is not (conditionally) associated with the level of development suggests there is a fixed “direct effect” of the program. Support for Lula is generally high among recipients,¹¹ regardless of whatever other processes is going on. On the regressions for Lula’s vote share among the other voters, the negative coefficient on HDI suggests the existence of an “indirect” effect, which allowed Lula to win over even voters that are not on the payroll. This result has profound implications for the logic of cash distribution. Not only the poorer voters are the ones which yield the best marginal political return to the dollar invested, but these returns seem to be further boosted by investing dollars in poor voters in generally poor places.

These results can be more conspicuously seen graphically (Figure 2). Holding all other variables at their means, we can see how support for Lula among non beneficiaries of the cash transfers drops dramatically over the range of HDI-M — the density of which is also shown — while support among beneficiaries hardly budges.

¹¹Note the substantially different intercept terms within each pair of equations.

A similar, but weaker result holds for the levels of inequality. In more unequal places, Lula tends to perform worse among non recipients and better¹² among recipients. The interpretation here is less clear, but the coefficients on inequality might be capturing the degree of polarization, or the perceived differences between the groups. If an “us vs. them” dynamic exists in the more unequal places, it might very well magnify the direct effects of the cash hand outs.

3 The “It’s the Economy” Hypothesis

While the previous result reinforces the link between the program and pro-Lula voting, not all analyses see the Bolsa Familia program as the cause for voting behavior. In particular, Carraro et al. (2007) have argued that it was the economic improvements for the poorest segments of the population that prompted most of the country’s poor to support Lula’s reelection.

One approach to isolate the effects of the economy from those of cash transfer policies could include the observation of other elections, which would allow one to observe the same municipalities under different economic conditions. Obviously, other issues of comparability would arise and require further control, but the selection of the 1998 election, when as in 2006 there was an incumbent president running, should help mitigate these problems. The problem is that the Bolsa Familia program did not exist then. Several antecessor programs were put in place by the Cardoso government, but mostly in the second half of his second term (after 2000). While we have recently obtained such data and are working to include it in the analysis, these data will only allow for a comparison with the 2002 election which, which was significantly different than the 1998 and 2006 election because the incumbent candidate was not running.

Within the framework of a single election, the most natural approach would be simply to include municipal level figures for economic growth in a regression of Lula’s vote share on the scope of the Bolsa familia program, plus a large battery of controls, such as the regressions presented in the previous section. This would still ignore the important fact that the program itself probably had non trivial effects on municipal level growth rates, at least in the poorest municipalities. As these municipalities are also the ones where the program has greater scope and where Lula received a larger share of the vote, it is likely that such a design would underestimate the political effects of the program.

Given this interdependency between the scope of the BF program and economic growth, we can parcel the program’s effects into direct and indirect effects. Besides any direct effect the Bolsa Familia might have on recipients, it should also have an indirect electoral effect that passes

¹²though the effect is only significant on the quintile setup.

through its contribution to economic growth. As such, the issue can be stated as a system of two equations that with the following structure:

$$Lula2006 = \beta_0 + \beta_1 BF\text{Scope} + \beta_2 Growth + \beta OtherVariables \quad (1)$$

$$Growth = \gamma_0 + \gamma_1 BF\text{Scope} + \gamma_1 Exports + \gamma OtherVariables \quad (2)$$

The most important feature here is that the municipal level growth rate is endogenous to the system. There is also the issue that these equations probably have correlated errors, because of the presence of the same variables in both. Such a system could be estimated by three-stage least-squares (3SLS), which is a generalization of the two-stage-least-squares method to take into account the association between equations.¹³ However, as 3SLS is highly sensitive to alternate specifications, we also estimate this system using 2SLS, and ignoring the correlation between the errors of the regressions.

Note that the goal of the design is not to study the determinants of economic growth at the municipal level, but rather to capture the indirect effect that the Bolsa Familia program had on Lula's 2006 vote share. For this reason, we do not aspire for a complete specification of in Equation 2. We do include an interaction term between the scope of the program and the level of income per head in the municipality, as we expect the former to have a greater impact on growth in the more developed municipalities. The endogeneity of economic growth requires it be instrumented, which merits further elaboration in the subject.

While economic growth has a great number of immediate and long term causes, one particular aspect of the problem might be of particular relevance to the separation between the effects of the Bolsa Familia program and that of economic growth generally considered. We need an instrument for growth that only affects Lula's vote share through its effect on economic growth¹⁴, and that is not linked in any way to the scope of the program. In essence, we are looking for instruments that capture determinants of growth that are exogenous (to the system).

It is a well known fact that exchange rate policies are one of the instruments that more clearly generate political and electoral effects, and for that reason have been widely used and abused by Latin American governments (Dornbusch & Edwards 1990). Overvalued exchange rate periods tend to be accompanied by consumption expansion and are for this reason generally

¹³As such, three stage least can be thought of as a combination of seemingly-unrelated regression (SUR) and two-stage least squares (SLS).

¹⁴Which is equivalent to saying that the instrument must be uncorrelated with the errors in Eq. 1.



Figure 3: Exchange Rate

Notes: Figure shows the bilateral real exchange rate between the Brazilian Real and the US Dollar. Horizontal line represents the historical average, roughly understood as the long term equilibrium rate.

highly popular with the electorate (Bonomo & Terra 1999). However, overvalued exchange rates do not affect all sectors — and consequently all regions — equally. While consumers may rejoice with an overvalued currency, the exporting sector is generally hurt. This differential impact has been exploited by (Carvalho Filho & Chamon 2008), who have recently argued that the mechanism by which exchange rate appreciations are transmitted to voters is mainly through changes in factor incomes rather than in consumption prices.

As Figure 3 shows, between 2002 and 2006 the Real moved from an extremely undervalued position against the US Dollar to an extremely overvalued one in historical terms. Hence, it is plausible that, all things equal, municipalities that rely to larger extent on exports would exhibit lower economic growth than municipalities that are essentially “consumers”. Granted, not all people in export relying municipalities earn wages from exports, and even those that do can benefit as consumers from overvalued exchange rates. The assumption here is simply that greater aggregate consumption capacity makes everybody on average better off, and that in non exporting municipalities this improvement is not offset by any loss in welfare resulting from forgone exports. More importantly to the task at hand, this greater capacity to consume is exogenous to the BA program.

The results shown in Table 4 confirm that export relying municipalities had lower growth, that the Bolsa Familia Program had a significant effect on growth, that its effect is smaller the richer the municipality, and that the program affected Lula’s vote share directly and indirectly. The first two specifications impose no constraints on the predicted vote shares for Lula, which

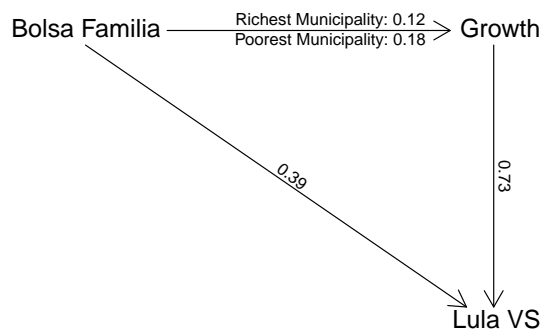


Figure 4: Direct and Indirect Effects

Notes: Figure shows the marginal effects of the change in one percentage point in the scope of the Bolsa Família, as estimated by the two-stage least squares model with logit transformation of the dependent variable, reported in Table 4. Result is shown for average municipality which also happens to be the range in which the direct effect is higher. It falls to around 0.28 for municipalities where the scope of the BA is higher.

in seven cases falls above the natural bound of 100%. While these add up to a very reduced fraction of the cases, general inspection of the residuals suggests that some transformation of the dependent variable would be warranted. Once the same system is estimated using a logit transformed dependent variable, with the exception of a few outlying extremely high growth municipalities in which Lula did not perform substantially well, residuals exhibit the desired (non) patterns. Effects are stable in the three specifications reported, and results from the second 2SLS estimates are reported graphically in Figure 4.

What this figure shows is that in an average municipality, an increase in the program's coverage in one percentage point is directly associated with an increase in Lula's vote share of 0.39 percentage points. Given the non-linear transformation of the dependent variable, the estimated direct effect falls to around 0.28 for municipalities where the scope of the BA is higher. Indirectly, however, this same increase in coverage is associated with higher growth, which in turn leads to more votes for Lula. The indirect effect ranges from an increase in 0.08 percentage points for Lula in the richest municipalities to 0.13 percentage points in the poorest ones. When all is said and done, the program seems to have had significant effects on the electoral results, independent of the general economic outlook.

Table 4: Simultaneous Estimation of Direct and Indirect Electoral Effects

(a) Equation 1: DV = Lula 2006 VS				(b) Equation 2: DV = Growth GDP-h			
	2SLS	3SLS	2SLS logit(DV)		2SLS	3SLS	2SLS logit(DV)
BA Scope	0.350	0.366	0.016	BA Scope	0.182	0.182	0.182
BA Scope	0.050	0.048	0.002	BA Scope	0.060	0.060	0.060
BA Scope	0.000	0.000	0.000	BA Scope	0.003	0.003	0.003
Growth GDP-h	0.618	0.618	0.029	BA Scope×log(GDP-h)	-0.012	-0.012	-0.012
Growth GDP-h	0.225	0.216	0.010	BA Scope×log(GDP-h)	0.062	0.062	0.062
Growth GDP-h	0.006	0.004	0.005	BA Scope×log(GDP-h)	0.849	0.849	0.849
PT Mayor VS	1.724	2.326	0.081	Exports-h	-0.196	-0.196	-0.196
PT Mayor VS	0.976	0.791	0.045	Exports-h	0.049	0.049	0.049
PT Mayor VS	0.077	0.003	0.077	Exports-h	0.000	0.000	0.000
PT Governor	-4.136	-4.641	-0.171	log(GDP-h)	4.760	4.760	4.760
PT Governor	0.822	0.723	0.038	log(GDP-h)	1.599	1.599	1.599
PT Governor	0.000	0.000	0.000	log(GDP-h)	0.003	0.003	0.003
log(Pop)	1.669	1.766	0.076	Const.	-7.296	-7.296	-7.296
log(Pop)	0.178	0.155	0.008	Const.	2.146	2.146	2.146
log(Pop)	0.000	0.000	0.000	Const.	0.001	0.001	0.001
log(GDP-h)	-6.004	-6.125	-0.274	N	4576	4576	4576
log(GDP-h)	0.903	0.871	0.042	R2	0.03	0.03	0.03
log(GDP-h)	0.000	0.000	0.000				
Lula 2002 VS	0.382	0.357	0.017				
Lula 2002 VS	0.020	0.017	0.001				
Lula 2002 VS	0.000	0.000	0.000				
Nonwhite	0.273	0.255	0.012				
Nonwhite	0.013	0.011	0.001				
Nonwhite	0.000	0.000	0.000				
Pentecostals	-0.125	-0.083	-0.005				
Pentecostals	0.041	0.036	0.002				
Pentecostals	0.002	0.022	0.007				
PT Mayor	-2.849	-2.809	-0.132				
PT Mayor	0.673	0.537	0.031				
PT Mayor	0.000	0.000	0.000				
Const.	2.133	2.263	-2.126				
Const.	2.104	1.889	0.098				
Const.	0.311	0.231	0.000				
N	4576	4576	4576				
R2	0.50	0.50	0.46				

4 Alternative Research Design

Thus far, the actual mechanism by which this effect is generated remains somewhat of a mystery. The aggregate association of greater vote share with greater BF coverage at the municipality level could be caused by several different processes. Even the ecological inference results that allow as a glimpse of the differences of voting behavior between recipients and non-recipients ignore the fact that both groups differ systematically with many other respects.

The ultimate test of the political effects of the BF program would rely on a comparison of individuals that receive the benefit with those that do not, and which are otherwise similar. Participation in the program would be the “treatment,” and the determination of its effect would require a control group of similar individuals that do not participate in the program.

The natural route for this comparison would be to rely on individual level survey data on participation in the program and political behavior. Short of designing a specific survey for this purpose, however, these data are not easily found. Not only very few electoral surveys have asked respondents about their participation in the program, but those that have, such as the LAPOP Barometer of the Americas, were conducted after the election and produced data does not match the actual election results very well, as we will show momentarily.

Furthermore, even if such data were available, the program’s effective targeting and wide coverage should make such comparisons difficult. Official data suggest that in Brazil, the treatment — as defined as participation in the program — is very well dispensed. Most qualifying poor individuals are effectively covered by the program, so it is no easy to find a control group, especially with probabilistic samples of individuals.

Had the program been implemented in randomly selected municipalities, as it was initially the case with Mexico’s Progresá (De La O 2008), this would serve as a natural experiment that could help identify the political consequences of the program through the use of aggregate municipal data. However, coverage is very close to the estimated number of poor families in most municipalities, thereby preventing attempts to identify the effects by comparing municipalities.¹⁵

The 2007 LAPOP Barometer of the Americas survey asked a couple of questions regarding participation in the Bolsa Familia program, and as such is the best suited existing survey to identify individual level political effects of the program. Though it was conducted less than two

¹⁵Some small variations in the coverage of the program in each municipality relative to the targets estimated from census data. In principle, municipalities that are below their coverage targets could be compared with similar ones that are above their targets, allowing for the estimation of coverage effects. This would amount to a more implementation of the notion of the treatment, but would again not be estimating effects at the individual level. We will pursue this route in subsequent studies, but it not clear, at this point, whether the differences in coverage across municipalities are sufficiently large to be captured by such a design.

months after the election — in the end of December 2006 — the survey disappointingly misses the election results by a wide margin. Lula, who obtained 46% of the valid vote in the first round appears in the survey with more than 69% of respondent recall. Alckmin, the runner up who obtained 41% of the valid votes in the actual election would have received just over 20% of the vote according to the survey. This problem is compounded by the fact that 22% of respondents did not respond to this question, so any serious use of these data leaves us to make heroic assumptions about the preferences of voters.¹⁶

Additionally, the phrasing of the two questions is not particularly conducive to the research we would like to do. The first question asked whether the interviewee participates in the Bolsa Familia Program and the other whether “a family member or somebody he/she knows” participates in. The first question is too restrictive, as the program is directed at families, and not at individuals. The second is too broad, as it defines neither family nor acquaintances. Perhaps because of this wording, 10% of respondents responded positively to the first question and 45% of respondents responded positively to the second one. Given that the program serves 11 million families we would expect this figure to be close to 20% in a national representative sample.

Another reason to doubt these data is that coverage of the Bolsa Familia does not seem to match well respondents stated income levels, despite the fact that the program meets its coverage targets in most municipalities in the country. Though we doubt that this survey has adequately captured our variables of interest, in practical terms, this result allows for the existence of control individuals within various income brackets.

These problems point to very limited reliability of conclusions drawn from these data, but with all the possible caveats involved, the predicted effect of being a recipient of the Bolsa familia on the probability of voting for Lula is stated in Table 5. To the very limited extent to which these data can be trusted, participating in the program increases in 15% the probability that somebody will vote for the government, an effect which is quite stable across different income brackets.

¹⁶Licio, Castro & Rennó (2009) use the LAPOP data and claim that it does not distort the election results. They cite figures where Lula receives 56% of the votes as evidence that only those who voted for other candidates “forgot” who they voted for. However, this percentage only obtains if one assumes missing respondents all voted for candidates other than Lula, which seems like a strong assumption.

Table 5: Bolsa Familia Political Effect Estimated From Individual Level Survey Data

Family Income Per Head	Frequencies		Prob(Vote-Lula)		Risk Ratios
	BF	Not BF	BF	Not BF	
Less than R\$ 60.00	30	49	0.74	0.86	1.15
R\$60.01 – R\$100.00	39	100	0.74	0.85	1.16
R\$100.01 – R\$140.00	15	112	0.73	0.85	1.16
R\$140.01 – R\$180.00	14	61	0.73	0.85	1.16
R\$180.01 – R\$220.00	14	120	0.72	0.84	1.17
R\$220.01 – R\$260.00	5	65	0.71	0.84	1.17
More than R\$260.01	7	498	0.71	0.83	1.18

Notes: Data was obtained from 2007 edition of the LAPOP Barometer of the Americas. Income per head is an approximation, computed from questions that asked family income and size of household. Probability of voting for Lula was computed from a logit regression, including income and whether respondent declared participate in the Bolsa Familia program. *All risk ratios are significant at the 0.95 level.

5 Reconciling the Results

What, after all these exercises, can we say about the electoral effects of the Bolsa Familia program in the Brazilian 2006 election? We are left with a few different estimates of different quantities. The simultaneous equations model suggests that one percentage point more of coverage is associated to about 0.4 percentage point higher vote-share. The ecological inference analysis suggests that recipients of the program were more likely to vote for Lula, about 18% more likely in poorer municipalities, and considerably higher than that in richer ones. Our survey data analysis says that beneficiaries are about 15% more likely to vote for Lula than otherwise similar non-beneficiaries. How do these estimates compare?

First, consider a comparison between the ecological inference results and the survey data analysis. In particular, reconsider the results displayed in Figure 2. Recall that this result does not take into consideration all the other differences between beneficiaries and non-beneficiaries. However, it is probably the case that in poorer municipalities these two groups are more similar to each other than in other places, as much of the poorest municipalities are, indeed, almost uniformly poor. If this is the case, the small difference between the voting behavior of recipients and non-recipients in the left of the Figure would be our best estimate of the electoral effects of the Bolsa Familia on similar individuals. This effect is slightly higher than what we found with a “better” control group in the survey data analysis, a fact that we take as a comforting assurance that we are not too far off.

The comparison of this individual level effect with aggregate results requires some gymnastics. To this end, we make use of a simulation where we consider, hypothetically, a poor municipality in which recipients voted for Lula with probability 0.85 and non recipients voted

for him with probability 0.73 (the predicted probabilities from Table 5). With a coverage of 90% of the population, Lula’s predicted vote share is 83.8%. Expanding coverage by 1 percentage point would raise Lula’s vote share in .12 percentage points. For this municipality, the direct effect estimated by the model reported in Figure 4 would be 0.29, or 2.3 times larger than the extrapolation of the individual level effect.

While these estimates are by no means the same, they are definitely in the same order of magnitude. More importantly, we did not really expect them to be the same. After all, the individual level effects we estimated above would be the BF effect on *otherwise similar* individuals. Our previous extrapolation applied this difference to all individuals in the hypothetical municipalities, even though the overall differences between beneficiaries and non-beneficiaries must be considerably larger. In addition, we already knew that the aggregate effect was probably an overestimation of the individual effects, as it is an average of the effects found in all different types of municipalities — even rich ones where beneficiaries and non-beneficiaries differ markedly¹⁷. Hence, we tentatively conclude that despite the shortcomings of our individual level data and the limitation of the aggregate data analysis, the two sets of figures are at least roughly compatible.

6 Conclusion & Future Research

That support for the incumbent candidate in Brazil was higher among beneficiaries of the *Bolsa Familia* program is not too surprising. However, one noteworthy result of this paper is that government’s support among non-beneficiaries of the program is very high in the less developed regions of the country. Municipalities in these regions rely heavily on the *Bolsa Familia* program, and in places like this, there is often very little private-sector economic activity going on, and much of what exists depends on some type of government transfer. Under these circumstances, voters’ perceptions about the economy are likely to be very much influenced by the cash handouts themselves, even for those that do not benefit directly. In fact, we also show that the program has a non-negligible indirect electoral effect through the economic stimulus it provides. This indirect effect ranges between 1/4 and 1/5 of the direct effect of the program, and is stronger in the poorest municipalities.

While there should be little doubt that the program was responsible for at least part of Lula’s new voting patterns, we have not achieved, as of yet, a proper estimate of the impact

¹⁷One variation that could help get at this would be to include an interaction term in the first regression allowing the effect of the BF to vary with the level of development of each municipality.

a program such as the Bolsa Familia on the individual’s voting decision. From the policy and from the purely electoral standpoint, it might not make too much sense to attempt estimate such effects. After all, it might not matter why the program is effective, but simply whether or not it is effective. Hence, it is not surprising that in the last election all four candidates — which spanned the spectrum from the extreme left to the center-right — advocated the expansion of the program. Electoral effectiveness is particularly relevant as the economic outlook becomes gloomier, budget cuts are considered, and priorities need to be reset. Effective results in fighting malnutrition, in keeping children in school, and in improving the lives of the neediest might make the Bolsa Familia a “good” policy. But good policies have fallen, and will fall, to the imperatives of the governments’ budgets. A good policy that is also politically viable has more chances of surviving, but to show effectiveness, in this arena, it suffices to show the aggregate effects.

However, from a theoretical standpoint the question of individual level effects is a very relevant one, for at least two different reasons. First, if recipients effectively vote for the incumbent government because of the hand-outs, one needs to explain why they do so even if the benefit is not conditional on voting. Second, if individual voters can actually be swayed in this way, without the need of costly monitoring of voting behavior, why then had not previous governments implemented such a program?

Previous analysis we carried out in past electoral results suggest that incumbent party candidates always tended to do better in poorer municipalities — as would be expected considering that the government controls large amounts of politically invaluable resources, and these are more fruitful when targeted to poorer voters. However, ecological inference techniques applied to aggregate electoral results suggests that these incumbents did not always manage to reach the poorest voters. What might lie at the bottom of this combination is simply that while the basic mechanics of distributional politics has always been clear, incumbents might have lacked the technology to actually reach the poorest voters, wherever they may be.

To answer these questions, more research, including better data quantitative data analysis and an effort to collect in depth qualitative evidence, lies ahead.

References

- Almeida, Alverto Carlos. 2006. *Por que Lula? O contexto e as estratégias políticas que explicam a eleição e a crise*. Rio de Janeiro: Record.
- Almeida, Jorge. 1996. *Como Vota o Brasileiro*. 2 ed. São Paulo: Xamã.
- Bonomo, Marco & Cristina Terra. 1999. “The Political Economy of Exchange Rate Policy in Brazil: 1964–1997.” *Revista Brasileira de Economia* 53(4):411–432.

- Canêdo-Pinheiro, Mauricio. 2009. “Bolsa Família ou Desempenho da Economia? Determinantes da Reeleição de Lula em 2006.” Unpublished manuscript, Fundação Getúlio Vargas, Rio de Janeiro.
- Carraro, André, Ari Francisco Araújo Jr, Otávio Menezes Damé, Lenoardo Monteiro Monasterio & Cláudio Djissey Shikida. 2007. “‘É a economia, companheiro!’: uma análise empírica da reeleição de Lula com base em dados municipais.” Ibemec MG Working Paper - WP41.
- Carvalho Filho, Irineu de & Marcos Chamon. 2008. “A Micro-Empirical Foundation for the Political Economy of Exchange Rate Populism.” *IMF Staff Papers* 55(3):481–510.
- Datafolha. 2002. Intenção de voto para presidente 2002 — Ibope/BR2006.SET-01692. In *Banco de Dados do Centro de Estudos de Opinião Pública*, ed. CESOP-UNICAMP. <http://www.cesop.unicamp.br/site/htm/busca/php>, accessed on 09/01/2008: .
- Datafolha. 2006a. Intenção de voto para presidente 2006 — PO 613364A Datafolha. In *Banco de Dados do Datafolha*, ed. Datafolha.
- Datafolha. 2006b. “Opinion Poll Data.” Available at <http://datafolha.folha.uol.com.br/>.
- De La O, Ana Lorena. 2008. “Do Poverty Relief Funds Affect Electoral Behavior? Evidence from a Randomized Experiment in Mexico.” Unpublished manuscript, Yale University.
- Dornbusch, Rudiger & Sebastian Edwards. 1990. *The macroeconomics of Populism in Latin America*. Chicago: Univesrity of Chicago Press.
- Hunter, Wendy & Timothy Power. 2007. “Rewarding Lula: Executive Power, Social Policy and the Brazilian Elections of 2006.” *Latin American Politics and Society* 49(1):1–30.
- Imai, Kosuke, Ying Lu & Aaron Strauss. 2008. “Bayesian and Likelihood Inference for 2 x 2 Ecological Tables: An Incomplete Data Approach.” *Political Analysis* 16(1):41–69.
- King, Gary. 1997. *A solution to the ecological inference problem : reconstructing individual behavior from aggregate data*. Princeton: Princeton University Press.
- Licio, Elaine, Henrique Castro & Lucio Rennó. 2009. “Bolsa Família e Voto nas Eleições Presidenciais de 2006: Em Busca do Elo Perdido.” Unpublished mansucrypt, Univesidade de Brasília.
- Nicolau, Jairo & Vitor Peixoto. 2007. “As bases municipais da votação de Lula em 2006.” Published online in Fórum Internet as Position Paper 2 <http://www.forumnacional.org.br/forum/pforum62a.asp>.
- Samuels, David. 2006. “Sources of Mass Partisanship in Brazil.” *Latin American Politics and Society* 48(2):1–27.
- Zucco Jr., Cesar. 2008. “The President’s ‘New’ Constituency: Lula and the Pragmatic Vote in Brazil’s 2006 Presidential Election.” *Journal of Latin American Studies* 40(1):29–49.