Codebook: The Politicization of Bureaucrats: Evidence from Brazil

Author: **Anderson Frey <anderson.frey@rochester.edu> & Rogerio Santarrosa <rogeriobs2@insper.edu.br>**

1. data_main.csv (observations at the municipality-coalition level)			
Variable	**Definition**	**Value & Label(s)**	**Source**
ibge	Municipality identifier (6-digit IBGE code)		IBGE
sta	State identifier (2-digit IBGE code)		IBGE
reg	Region identifier (1-digit IBGE code)		IBGE
t	Indicates the coalition status in 2009-2012	0=Runner-up 1=Incumbent	TSE
run	Margin of victory in the 2008 mayoral election, calculated as the vote difference between winner and runner-up, as a % of total votes	It has the same absolute value in each municipality for t=0 and t=1, but assumes negative values for the runner-up (t=0)	TSE
party	2-digit code of the party that leads the coalition (party of the mayoral candidate in 2008)		TSE
colig	2-digit code of all parties in the coalition		TSE
age	age of the mayoral candidate in 2008		TSE
gender	gender of the mayoral candidate in 2008	0=male 1=female	TSE
public_servant	binary variable that indicates whether the mayoral candidate was a bureaucrat before the election	0=no 1=yes	TSE
newcomer	binary variable that indicates whether the mayoral candidate ran in the previous (2004) mayoral or council elections	0=yes 1=no	TSE
education	binary variable that indicates whether the mayoral candidate has post-secondary education	0=no 1=yes	TSE
pt_mayor	binary variable that indicates whether the mayoral candidate belongs to the Worker's Party (PT, no 13)	0=no 1=yes	TSE
pt_federal	binary variable that indicates whether the mayoral candidate belongs to a party that supports PT at the national level (e.g. PT, PMDB, PDT, PCdoB, PSB, PR, PRB, PSC, PTC or PTN)	0=no 1=yes	TSE
pt_coalition	binary variable that indicates whether the mayoral candidate's coalition includes PT	0=no 1=yes	TSE
size	number of parties in the coalition		TSE
pot_enrollment	the difference between the new CadUnico target (in 2009) and the actual enrollment at the end of 2008, as a percentage of the 2008 population		IBGE, MDS
population	2008 population		IBGE
past_memb1	number of members of the coalition parties in 2008 among all voters (except bureaucrats, per 10,000 voting population)		MDS, RAIS, TSE2
past_memb2	number of members of the coalition parties in 2008 among all bureaucrats (except interviewers, per 10,000 voting population)		MDS, RAIS, TSE2
past_memb3	number of members of the coalition parties in 2008 among bureaucrats that were already doing CadUnico interviews in 2008 (per 10,000 voting population)		MDS, RAIS, TSE2
outcome	number of interviewers that joined a party in the coalition in 2009-2012		MDS, RAIS, TSE2
outcome.2	number of bureaucrats (except interviewers) that joined a party in the coalition in 2009-2012		RAIS, TSE2
outcome.3	number of voters (except bureaucrats) that joined a party in the coalition in 2009-2012		TSE2
total	total interviewers in the municipality	same value for both t=0 and t=1	MDS, RAIS
total.2	total bureaucrats (except interviewers) in the municipality	same value for both t=0 and t=1	RAIS
total.3	total voters (except bureaucrats) in the municipality	same value for both t=0 and t=1	TSE
total.m	total interviewers in the municipality, but only considering the ones that have a match in the other bureaucracy, as described in the text	same value for both t=0 and t=1 NA=interviewer not matched	MDS, RAIS, TSE2
outcome.m.i	number of matched interviewers that joined a party in the coalition in 2009-2012	NA=interviewer not matched	MDS, RAIS, TSE2
outcome.m.b	number of matched bureaucrats that joined a party in the coalition in 2009-2012	NA=interviewer not matched	MDS, RAIS, TSE2
mv3rd	the vote difference between runner-up and 3rd place in the mayoral election, as a % of total votes	NA=election only had 2 candidates	TSE
outcome.3rd	number of interviewers that joined a party in the coalition in 2009-2012 (here the variable t corresponds to the runner-up coalition (t=1) and the 3rd place (t=0))	NA=election only had 2 candidates	MDS, RAIS, TSE2
outcome2.3rd	number of bureaucrats (except interviewers) that joined a party in the coalition in 2009-2012 (here the variable t corresponds to the runner-up coalition (t=1) and the 3rd place (t=0))	NA=election only had 2 candidates	MDS, RAIS, TSE2

outcome3.3rd	number of voters (except bureaucrats) that joined a party in the coalition in 2009-2012 (here the variable t corresponds to the runner-up coalition (t=1) and the 3rd place (t=0))	NA=election only had 2 candidates	MDS, RAIS, TSE2
council_cand	number of council candidates running in 2012 that belong to one of the parties in the coalition (parties listed in *colig*)		TSE
inc	binary variable that indicates whether the incumbent mayor in 2005-2008 was a mayoral candidate in 2008	0=no 1=yes	TSE
cand_tot	total number of mayoral candidates in the 2008 election		TSE
hh	Herfindahl-Hirschman concentration index based on the mayoral vote in 2008		TSE

2. data_cadunico.csv (observations at the interviewer level (not identified))			
Variable	**Definition**	**Value & Label(s)**	**Source**
ibge	Municipality identifier (6-digit IBGE code)		IBGE
new_party	2-digit code of the party of affiliation in 2009-2010	NA=did not join any party	MDS, TSE2
old_party	2-digit code of the party of affiliation in 2008 or before	NA=did not join any party	MDS, TSE2
coalition_side	Indicates the coalition of affiliation in 2009-2012	-1=Other 0=Runner-up 1=Incumbent NA=did not join any party	MDS, TSE, TSE2
party_win	2-digit code of the party that leads the incumbent coalition in 2009-2012		MDS, TSE, TSE2
party_los	2-digit code of the party that leads the runner-up coalition in 2009-2012		MDS, TSE, TSE2
date_int	date of the first interview on record for the bureaucrat	YYYY-MM-DD	MDS
date_new	date of party affiliation in 2009-2012	YYYY-MM-DD, NA=did not join any party	RAIS, TSE2
int_200912	total number of interviews in 2009-2012		MDS
int_2012	total number of interviews in 2012		MDS
int_elig	total number of interviews in 2009-2012 with households with BF-eligible income		MDS
int_new	total number of interviews in 2009-2012 with households entering CadUnico for the first time		MDS
int_exact1	total number of interviews in 2009-2012 with households with income exactly at R$70		MDS
int_exact2	total number of interviews in 2009-2012 with households with income exactly at R$140		MDS
median1	median number of interviews in 2009-2012 in the municipality		MDS
median2	median number of interviews in 2012 in the municipality		MDS
salary	average monthly wages in 2008	in R$	RAIS
salary12	average monthly wages in 2012	in R$, NA=not available	RAIS
ocup	occupation code on record in 2008		RAIS
ocup12	occupation code on record in 2012	NA=not available	RAIS
education	education level of the interviewer in 2008	categorical variable with levels between 1 (illiterate) and 11 (doctoral degree)	RAIS
education12	education level of the interviewer in 2012	categorical variable with levels between 1 (illiterate) and 11 (doctoral degree)	RAIS
year_born	year of birth of the interviewer		RAIS
year_hired	year that the interviewer was hired into the bureaucracy		RAIS
gender	gender	0=male, 1=female	RAIS
coalition_side2	Indicates the coalition of affiliation in 2009-2012 for interviewers that ran for local council in 2012 (based on the party of the candidacy and the 2009-2012 coalitions)	-1=Other 0=Runner-up 1=Incumbent NA=did not run	MDS, TSE, TSE2
remains_bureaucrat	binary variable that indicates whether the interviewer remained in the bureaucracy in 2012	0=no 1=yes NA=not found in 2012 RAIS	RAIS
admin	binary variable that indicates whether the interviewer has a clerical occupation	0=no 1=yes	RAIS
social	binary variable that indicates whether the interviewer is a social worker	0=no 1=yes	RAIS
health	binary variable that indicates whether the interviewer has an occupation in health care	0=no 1=yes	RAIS
lowskill	binary variable that indicates whether the interviewer has a low-skill occupation	0=no 1=yes	RAIS

manag	binary variable that indicates whether the interviewer has a managerial position	0=no 1=yes	RAIS
civil	binary variable that indicates whether the interviewer was hired under the civil service code (job security), as of 2008	0=no 1=yes	RAIS
civil12	binary variable that indicates whether the interviewer was hired under the civil service code (job security), as of 2012	0=no 1=yes	RAIS
coalition_side_past	Indicates the coalition status based on the party affiliation in 2008 or before	-1=Other 0=Runner-up 1=Incumbent NA=not a party member before 2009-2012	MDS, TSE, TSE2

3. data_bureau.csv (observations at the bureaucrat level (not identified), excludes CadUnico interviewers)			
Variable	**Definition**	**Value & Label(s)**	**Source**
ibge	Municipality identifier (6-digit IBGE code)		IBGE
new_party	2-digit code of the party of affiliation in 2009-2010	NA=did not join any party	RAIS, TSE2
old_party	2-digit code of the party of affiliation in 2008 or before	NA=did not join any party	RAIS, TSE2
coalition_side	Indicates the coalition of affiliation in 2009-2012	-1=Other 0=Runner-up 1=Incumbent NA=did not join any party	RAIS, TSE, TSE2
party_win	2-digit code of the party that leads the incumbent coalition in 2009-2012		RAIS, TSE, TSE2
party_los	2-digit code of the party that leads the runner-up coalition in 2009-2012		RAIS, TSE, TSE2
date_new	date of party affiliation in 2009-2012	YYYY-MM-DD, NA=did not join any party	RAIS, TSE2
salary	average monthly wages in 2008	in R$	RAIS
ocup	occupation code on record in 2008	NA=not available	RAIS
education	education level of the interviewer in 2008	categorical variable with levels between 1 (illiterate) and 11 (doctoral degree)	RAIS
year_born	year of birth of the interviewer		RAIS
year_hired	year that the interviewer was hired into the bureaucracy		RAIS
gender	gender	0=male, 1=female	RAIS
coalition_side2	Indicates the coalition of affiliation in 2009-2012 for interviewers that ran for local council in 2012 (based on the party of the candidacy and the 2009-2012 coalitions)	-1=Other 0=Runner-up 1=Incumbent NA=did not run	RAIS, TSE, TSE2
admin	binary variable that indicates whether the interviewer has a clerical occupation	0=no 1=yes	RAIS
social	binary variable that indicates whether the interviewer is a social worker	0=no 1=yes	RAIS
health	binary variable that indicates whether the interviewer has an occupation in health care	0=no 1=yes	RAIS
lowskill	binary variable that indicates whether the interviewer has a low-skill occupation	0=no 1=yes	RAIS
manag	binary variable that indicates whether the interviewer has a managerial position	0=no 1=yes	RAIS
civil	binary variable that indicates whether the interviewer was hired under the civil service code (job security), as of 2008	0=no 1=yes	RAIS

4. data_matched.csv (observations at the interviewer level (not identified), only for the 86% that had a match in the other bureaucracy)			
Variable	**Definition**	**Value & Label(s)**	**Source**
ibge	Municipality identifier (6-digit IBGE code)		IBGE
gender	gender of the interviewer		MDS, RAIS
admin_int	binary variable that indicates whether the interviewer has a clerical occupation	0=no 1=yes	MDS, RAIS
social_int	binary variable that indicates whether the interviewer is a social worker	0=no 1=yes	MDS, RAIS
health_int	binary variable that indicates whether the interviewer has an occupation in health care	0=no 1=yes	MDS, RAIS
lowskill_int	binary variable that indicates whether the interviewer has a low-skill occupation	0=no 1=yes	MDS, RAIS
manag_int	binary variable that indicates whether the interviewer has a managerial position	0=no 1=yes	MDS, RAIS

salary_int	average monthly salary of the interviewer in 2008	in R$	MDS, RAIS
education_int	education level of the interviewer in 2008		MDS, RAIS
year_born_int	year of birth of the interviewer	categorical variable with levels between 1 (illiterate) and 11 (doctoral degree)	MDS, RAIS
year_hired_int	year that the interviewer was hired into the bureaucracy		MDS, RAIS
The following variables have the characteristics of the bureaucrat that was matched to each interviewer in the dataset, according to the procedure described in the main text			
admin_bur	binary variable that indicates whether the bureaucrat has a clerical occupation	0=no 1=yes	RAIS
social_bur	binary variable that indicates whether the bureaucrat is a social worker	0=no 1=yes	RAIS
health_bur	binary variable that indicates whether the bureaucrat has an occupation in health care	0=no 1=yes	RAIS
lowskill_bur	binary variable that indicates whether the bureaucrat has a low-skill occupation	0=no 1=yes	RAIS
manag_bur	binary variable that indicates whether the bureaucrat has a managerial position	0=no 1=yes	RAIS
salary_bur	average monthly salary of the bureaucrat in 2008	in R$	RAIS
education_bur	education level of the bureaucrat in 2008		RAIS
year_born_bur	year of birth of the bureaucrat	categorical variable with levels between 1 (illiterate) and 11 (doctoral degree)	RAIS
year_hired_bur	year that the bureaucrat was hired into the bureaucracy		RAIS

5. data_partisans.csv (observations at the municipality-coalition-affiliation date level)			
Variable	**Definition**	**Value & Label(s)**	**Source**
ibge	Municipality identifier (6-digit IBGE code)		IBGE
coalition_side	binary variable that indicates whether the voters became members of an incumbent or opposition party in 2009-2012	0=runner-up, 1=incumbent	TSE, TSE2
late_term	binary variable that indicates the date of the party membership	0=2009-2010, 1=2011-2012	TSE, TSE2
all	number of voters that joined each coalition, each period		TSE, TSE2
mayor	number of voters that joined each coalition, each period, but only the ones that joined the party that led the coalition		TSE, TSE2

6. data_partisanship.csv (observations at the party-year level)			
Variable	**Definition**	**Value & Label(s)**	**Source**
year	year of party affiliation	1980-2012	TSE2
party	2-digit code of the party of affiliation		TSE2
members	number of voters that joined each party, each year		TSE2
cancel	number of voters that left each party, each year		TSE2

Data References

IBGE, Estimativas de População. 2008. Data for each municipality.
https://sidra.ibge.gov.br/pesquisa/estimapop/tabelas

MDS, Cadastro Unico. 2012. Data on individual households enrolled in the Bolsa Família registry.
https://www.gov.br/mds/pt-br/servicos/sagi/microdados
Data only available upon request.

RAIS, Relação Anual de Informações Sociais. 2008-2012. *Data only available upon request.*
Data only available upon request.

TSE, Portal de Dados Abertos. 2020. Election results and profile of candidates for all elections in Brazil in 2008 and 2012.
https://dadosabertos.tse.jus.br/

TSE2, Relação de Filiados. 2016. Membership rolls of all Brazilian parties.
http://www.tse.jus.br/partidos/filiacao-partidaria/relacao-de-filiados
Data only available upon request.