CODEBOOK FOR THE DATASET ON THE ELECTORAL PERFORMANCE AND CRIMINAL STATUS OF CANDIDATES CONTESTING THE 2004 AND 2009 PARLIAMENTARY ELECTIONS TO THE LOK SABHA (INDIA)

Miriam A. Golden, P.I.
Department of Political Science
University of California, Los Angeles
golden@ucla.edu

With assistance of
Devesh K. Tiwari
Department of Political Science
University of California, San Diego

This version created: 8 November 2014

This dataset contains information on the electoral performance and criminal status of all candidates contesting elections to the lower house of the Indian Parliament, the Lok Sabha, in the elections of 2004 and 2009.

These were the first national elections conducted after the 2002 Supreme Court ruling mandating that all candidates running for public office file affidavits with the Election Commission of India prior to the election in which they report criminal histories or pending criminal charges for any offense punishable with imprisonment of two years or more and lodged prior to six months before the individual filed candidate papers. Affidavits were compiled and digitized by the Liberty Institute of India, a New Delhi-based NGO, and posted on its website http://www.empoweringindia.org/new/home.aspx. We input only whether the candidate reported being charged or not, excluding information such as the statute number.

Each row in this dataset corresponds to a candidate who is running for parliament in an electoral constituency in a given election year.

All electoral data included in this dataset is from the Election Commission of India. The 2004 data was downloaded October 2007 and the 2009 data was downloaded October 2009 (both from http://eci.nic.in).

Data on self-reported criminal status of candidates is from the Liberty Institute of India (downloaded May 2009 from http://www.empoweringindia.org/).

This codebook reflects the output generated by the “codebook” command when the dataset is loaded in Stata.

VARIABLES

---

year — Election year (2004/2009)

---

type: numeric (byte)

range: [4,9] units: 1
unique values: 2 missing .: 0/13505

tabulation:

Freq.	Value
5435	4
8070	9

---

cand_name — Candidate name

---

type: string (str70)

unique values: 12141 missing "": 0/13505

examples:
"Dau Ram Ratnakar"
"Kandarpa Lahkar"
"P.K. Raman"
"Saravanan.V"

warning: variable has embedded blanks

---

gotseat — Won or lost seat

---

type: numeric (float)
label: won

range: [0,1] units: 1
unique values: 2 missing .: 0/13505

tabulation:

Freq.	Numeric	Label
12419	0	Lost
1086	1	Won

---

incum — Incumbent

---

type: numeric (float)
label: incum

range: [0,1] units: 1
unique values: 2 missing .: 0/13505

tabulation:	Freq.	Numeric	Label
	12803	0	Not incumbent
	702	1	Incumbent

---

pc_name — Parliamentary constituency name

---

type: string (str25)

unique values: 699 missing "": 0/13505

examples: "Chamarajanagar"
"Howrah"
"Mahbubnagar"
"Rajnandgaon"

warning: variable has embedded blanks

pc_name:
1. Boundaries changed between 2004 and 2009.

---

pc_no — Parliamentary constituency number

---

type: numeric (float)

range: [1,80] units: 1
unique values: 80 missing .: 0/13505

mean: 19.8306
std. dev: 17.5315

percentiles:	10%	25%	50%	75%	90%
	3	6	15	28	42

---

pid — Party ID

---

type: string (str56), but longest is str10

unique values: 453 missing "": 0/13505

examples: "CGVP"
"IND"
"IND"
"JKN"

pid:
1. Independent candidates labeled as IND. See party abbreviations file for full list.

---

rank — Vote rank

---

type: numeric (float)

range: [1,43] units: 1
unique values: 43 missing .: 0/13505

mean: 8.34543
std. dev: 6.22982

percentiles:	10%	25%	50%	75%	90%
	2	4	7	12	16

rank:
1. The rank of each candidate in the parliamentary constituency in order of vote totals. Rank of 1 indicates that the candidate won the seat.

---

res — Reserved

---

type: numeric (float)
label: reserv

range: [0,1] units: 1
unique values: 2 missing .: 0/13505

tabulation:	Freq.	Numeric	Label
	11079	0	Not reserved
	2426	1	Reserved for SC/ST

res:
1. SC = Scheduled Caste; ST = Scheduled Tribe

---

st_code — State code (ECI)

---

type: string (str3)

unique values: 35 missing "": 0/13505

examples: "S06"
"S13"
"S22"
"S24"

---

st_name — State name

---

type: string (str35), but longest is str25

unique values: 35 missing "": 0/13505

examples: "Gujarat"
"Madhya Pradesh"
"Punjab"
"Uttar Pradesh"

warning: variable has embedded blanks

---

votes — Number of votes received

---

type: numeric (float)

range: [45,855543] units: 1
unique values: 9731 missing .: 0/13505

mean: 59751.1
std. dev: 117573

percentiles:

10%	25%	50%	75%	90%
850	1791	4795	27123	259834

---

charge — Criminal charge

---

type: numeric (float)
label: charge

range: [0,1] units: 1
unique values: 2 missing .: 0/13505

tabulation:	Freq.	Numeric	Label
	12137	0	reports no charge
	1368	1	reports charge

---

party_type — National Party (ECI)

---

type: numeric (float)
label: party_typel

range: [0,2] units: 1
unique values: 3 missing .: 0/13505

tabulation:	Freq.	Numeric	Label
	9013	0	unrecognized or unaffiliated
	1562	1	state
	2930	2	national

party_type:
1. National parties in 2004 were BJP, BSP, CPI, CPM, INC, and NCP. National parties in 2009 were BJP, BSP, CPI, CPM, INC, NCP, and RJD.