Web Appendix to "Sweetening the Deal? Political Connections and Sugar Mills in India"

Sandip Sukhtankar*

September 22, 2011

1 Data

The table at the end of this section summarizes the yearly availability and sources for each variable. Below I describe the different sets of data I use.

1.1 Mill outcomes data

Many technical details on mill operations are available from the bulletins of the Sugar Technologists Association of India. Annual data from 1993-94 up to 2005-06 on the amount of cane crushed and sugar produced, amounts of lime and sulphur used in production, recovery rates, days in operation, hours lost due to machinery breakdowns and cane shortages, and yearly capacity on almost 200 cooperative mills in Maharashtra are hence available from various sources, described in the table A1. In 2005-06 132 mills operated; the number of mills in existence is typically much higher than the number of mills in operation (180 in 2005-06), since some mills may remain closed in a particular year. All price data are deflated using the Consumer Price Index for Agricultural Labor in Maharashtra, 1969-1994 from a dataset compiled by Ozler, Datt and Ravallion (1996) at the World Bank, 1995 onwards from the Statistical Abstract of India, which is the original source for the Ozler et al. (1996) dataset.

1.2 Climate data

Monthly rainfall data on a half degree by half degree grid from the Global Precipitation Climatology Center were interpolated spatially using the exact geographical coordinates of the mill to approximate rainfall in the cane growing command area of the mill. The "krig" method for interpolation was used. For each mill-year observation, I use rainfall and the squared deviation from average rainfall in that month for all months of the year. Hence each mill-year observation for rainfall consists of 24 variables; a variable recording rainfall (in millimeters) and a second variable that includes the squared deviation from the average rainfall in that month for the twelve months of the year.

*Dartmouth College, Department of Economics, 326 Rockefeller Center, Hanover, NH 03755; sandip.sukhtankar@dartmouth.edu.

1.3 Elections data

Elections data for state and national elections include the date of elections, names and party affiliations of candidates, number of candidates, and votes polled. The election year timing was also adjusted to the cane season. For example, if an election took place in May 1994, this corresponds to the 1993-94 crushing season (year 1993 in the data), whereas if it took place in December 1996 it would correspond to the 1996-1997 crushing season (year 1996 in the data). Elections take place in different months of the year as determined by the State or National Election Commission. Some take place during the time that the mills are operating, while others take place in the summer when they are closed.

1.4 Political connections data

Names of chairmen of mills were matched to the names of electoral candidates as listed by the Election Commission of India. Due to the abundance of common last names (e.g. Patil, Deshmukh, Pawar), only exact matches were considered. An exact match implies the first name, middle name or initial, and last name matched, with allowances for common misspellings. To further reduce spurious correlations, only candidate names from constituencies containing or adjacent to the chairman's sugar mill were considered. It is extremely unlikely that a chairman contests from a constituency very far away from the mill, while it is possible that someone with the same name contesting from another constituency is actually *not* the same person. The political party of the chairman-candidate was then considered to be the party affiliated with the sugar mill.

1.5 Village Public Goods data

Village public goods data are published in the 2001 Village Directory series published by the Census of India, and village locations are mapped in GIS by MapInfo (available at the Harvard Map Collection). I considered a village to be in the catchment area of a mill if it was located within a 15 kilometer radius of the mill.

2 Cane planted and growing

All satellite data were obtained through Landsat 7, and downloaded from a United States Geological Survey website: earthexplorer.usgs.gov. The data are 30m resolution multi-spectral images, i.e. they contain various wavelengths including those outside the visible range. The images are taken in September/ early October, when all sugarcane that will be crushed in the season has been planted and is growing but not yet harvested.

These were mathematically transformed using ArcGIS into an index called the Normalized Difference Vegetation Index (NDVI). NDVI uses the near infra-red and red wavelengths of the satellite images, using a standard algorithm to transform multiple spectral bands into a single dimension corresponding to physical vegetation parameters. Each pixel of the image becomes a value in between and including -1 and 1. The values represent the vegetation of the particular pixel; different crops correspond to different ranges within this range. It is safe to say that values zero and below do not represent vegetation.

I follow standard procedures in remote sensing to determine sugarcane planted in the areas surrounding particular mills. (See, for example, Rao et al. (2002); Rehman et al. (2004); Mehta et al. (2006)) Crops in general fall between (0,1] in the NDVI range. Although this range includes vegetation that may not be crops, I assume that this forms the denominator for our calculation as it applies to potential land available for growing sugarcane. By referencing coordinates of over 20 sample sugarcane fields in Maharashtra and Tamil Nadu, I calibrated sugarcane to lie between 0.3 and 0.6 in the NDVI range.[1]. This range covers healthy growing sugarcane, and as such captures a measure of quality of the crop as well as mere existence.

Given these ranges of all crops and sugarcane, I created a catchment area of a 15 km radius circle around the exact location of the mill. This 15 km radius is appropriate since for many years this was the buffer zone around a mill in which other mills were not allowed to locate. I calculated the number of pixels in the 15 km radii that were crops in general and sugarcane, and then the proportion of crops that were sugarcane. Because there were instances of overlap and cutting off of the buffer areas when I overlayed them onto the NDVI images, I picked the NDVI images that included the greatest portion of the cut off buffer mills. Areas of cloud cover were ignored.

References

Mehta, B.S., R. Prabhakar, and E. Somanathan, "How Degraded are Himalayan Forests?," *Current Science*, 2006, *91* (1), 61–67.

Ozler, Berk, Gaurav Datt, and Martin Ravallion, "A Database on Poverty and Growth in India," Dataset, Development Research Group, World Bank 1996.

Rao, P. V. Krishna, V. Venkateswara Rao, and L. Venkataratnam, "Remote sensing: A technology for assessment of sugarcane crop acreage and yield," *Sugar Tech*, 2002, *4* (3-4), 97–101.

Rehman, Md. Rejaur, A.H.M. Hedayatul Islam, and Md. Ataur Rehman, "NDVI Derived Sugarcane Area Identification and Crop Condition Assessment," *Planplus*, 2004, *2.*

[1]Exact coordinates of fields available on request

Table A.1: Data Sources

Variable	Years	Variation	Sources
Cane Price	1993-2004	Mill-year	Banerjee et al; MSSKS; YCI
Names of Chairmen	1993-2005	Mill-year	Mill annual reports (MSSKS); NFCSF
Recovery Rate	1993-2005	Mill-year	STAI; Banerjee et al; VSI
Cane Crushed	1993-2005	Mill-year	STAI; VSI
Sugar Produced	1993-2005	Mill-year	STAI; VSI
Cane Planted	2003-2005	Mill-year	See Appendix B
Capacity	1993-2005	Mill-year	STAI; Banerjee et al; VSI; MSSKS
Operational Status	1993-2005	Mill-year	STAI; VSI; MSSKS
Actual Days Worked	1993-2005	Mill-year	STAI; VSI
Actual Hours Worked	1993-2005	Mill-year	STAI; VSI
Hours Lost to Breakdowns	1993-2005	Mill-year	STAI; VSI
Hours Lost to Cane Shortage	1993-2005	Mill-year	STAI; VSI
Lime added	1993-2005	Mill-year	STAI
Sulphur added	1993-2005	Mill-year	STAI
Rainfall	1993-2005	Mill-month	GPCC
Election dates	1993-2005	Year	ECI
Election Candidates	1993-2005	Constituency-year	ECI
Consumer Price Index	1993-2005	Year	Ozler et al; SAI
Public goods	2001	Village	Census of India Village Directory data

Description of sources

STAI	Yearbooks of the Sugar Technologists Association of India
VSI	Yearbooks of the Vasantdada Sugar Institute
MSSKS	Maharashtra Sugar Cooperatives Federation
YCI	Yashwant Chavan Institute
NFCSF	Yearbooks of the National Federation of Cooperative Sugar Factories
GPCC	Global Precipitation Climatology Centre, http://gpcc.dwd.de
ECI	Election Commission of India, www.eci.gov.in
SAI	Statistical Abstract of India

Table A2
Separating National and State Politicians

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
National politician chairman	-6.622	-23.35			-6.336	-23.94		
	(23.89)	(17.38)			(24.30)	(17.25)		
State politician chairman			9.405	12.00			8.232	9.966
			(17.36)	(15.16)			(17.61)	(15.20)
National politician * national election	-15.31**	-13.76						
	(7.602)	(9.027)						
State politician * state election			-5.802	-6.194				
			(9.875)	(12.14)				
Chairman contests national election					-16.37*	-12.08***		
					(8.838)	(2.822)		
Chairman contests state election							-0.957	2.609
							(16.62)	(17.18)
Mill fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Rainfall, capacity	No	Yes	No	Yes	No	Yes	No	Yes
N	1,151	1,151	1,151	1,151	1,151	1,151	1,151	1,151
Adj R-squared	0.86	0.87	0.86	0.87	0.86	0.87	0.86	0.87

The table reports coefficients from estimations of equation 1 with the cane price as dependent variable, as well as a modified version of equation 1 to consider the case where sitting chairmen actually contest elections.
"National [state] politician chairman" refers to someone who was a candidate for national [state] elections during 1993-2005.
"Chairman contests" election is when a sitting chairman contested election in a given election year.
Standard errors are multi-way clustered by year-region as well as by mill. p-values: *** < 0.001, ** < 0.05, * < 0.01

Table A3
Elections during and out of season

	Cane Price			Recovery Rate		
	(1)	(2)	(3)	(4)	(5)	(6)
Election during season	-18.94			-0.0436		
	(16.91)			(0.0570)		
Election out of season	-22.48**			-0.0768		
	(9.155)			(0.0500)		
National election during season		-19.89			0.0189	
		(15.29)			(0.0612)	
National election out of season		-20.18**			-0.0664	
		(9.016)			(0.0503)	
State election during season			-7.984			-0.0760
			(29.67)			(0.0683)
State election out of season			-1.197			-0.0430
			(7.178)			(0.0493)
Mill fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Rainfall, capacity	Yes	Yes	Yes	Yes	Yes	Yes
N	1,151	1,151	1,151	1,413	1,413	1,413
Adj R-squared	0.87	0.87	0.87	0.75	0.75	0.75

The table reports coefficients from an estimation of equation 1 with the cane price and recovery rate as dependant variables, with elections split by whether they take place during the cane crushing season. Standard errors are clustered by year-region as well as by mill. p-values: *** < 0.001, ** < 0.05, * < 0.01