Environmental Regulations, Air and Water Pollution, and Infant Mortality in India

Michael Greenstone

Rema Hanna

Online Appendix

Not for Publication

This online appendix discusses the one-stage approach applied in Greenstone and Hanna (2013) and its comparability to the two-stage approach.

The single-stage approach consolidates the two stage approach in Greenstone and Hanna (2013) into a single regression. The single-stage regression comes in three specifications. Similar to equation (2a), the first specification includes a policy dummy and captures the mean-shift in pollution levels as a result of the policy. Note that for comparability with the two-stage approach, we must limit the policy dummy to event years that are identified by the minimum number of city-year observations (20 city years for Catalytic Converter and NRCP; 15 city years for SCAP). For water pollution, the regression takes the following form:

$$Y_{ct} = \alpha + \theta_1 1(NRCPRange)_\tau + \theta_2 1(NRCP)_\tau * (NRCPRange)_\tau + \theta_3 1(\tau Left)_\tau + \theta_4 1(\tau Right)_\tau + \mu_t + \gamma_c + \beta X_{ct} + \epsilon_{ct} \quad (3a)$$

where $1(NRCPRange)_\tau$ is a dummy variable indicating that $-7 \leq \tau \leq 10$ and $1(NRCP)_\tau$ is a dummy variable indicating whether or not one of the NRCP suite of policies is in force (i.e. an adopting city with $\tau \geq 0$). Two additional dummy variables, $1(\tau Left)_\tau$ and $1(\tau Right)_\tau$, indicate that $\tau < -7$ or $\tau > 10$, respectively.

In the case of air pollution, we include both policies in the same regression:

$$Y_{ct} = \alpha + \theta_1 1(SCAPRange)_\tau + \theta_2 1(SCAP)_\tau * (SCAPRange)_\tau + \theta_3 1(\tau Left)_\tau + \theta_4 1(\tau Right)_\tau + \rho_1 1(CCRange)_\phi \\ ht)_\tau + \rho_1 1(CCRange)_\phi + \rho_2 1(CC)_\phi * (CCRange)_\phi + \rho_3 1(\phi Left)_\phi + \rho_4 1(\phi Right)_\phi + \mu_t + \gamma_c + \beta X_{ct} + \epsilon_{ct} \quad (3a)$$

where $1(SCAPRange)_\tau$ is a dummy variable indicating that $-7 \leq \tau \leq 3$ and $1(CCRange)_\phi$ is a dummy variable indicating that $-7 \leq \phi \leq 9$. $1(SCAP)_\tau$ and $1(CC)_\phi$ are dummies indicating whether or not the SCAP and Catalytic Converter policies, respectively, are in force (i.e. an adopting city with $\tau \geq 0$ and/or $\phi \geq 0$). The dummies $1(\tau Left)_\tau$ and $1(\tau Right)_\tau$ indicate that $\tau < -7$ or $\tau > 3$, respectively; $1(\phi Left)_\phi$ and $1(\phi Right)_\phi$ indicate that $\phi < -7$ or $\phi > 9$, respectively .

The second specification is analogous to Equation (2b) insofar as it includes a linear time trend variable. However, we must interact it with the Policy-Range dummy. In the case of water pollution:

$$Y_{ct} = \alpha + \theta_1 1(NRCPRange)_\tau + \theta_2 1(NRCP)_\tau * (NRCPRange)_\tau + \theta_3 1(NRCPRange)_\tau * \tau + \theta_4 1(\tau Left)_\tau + \theta_5 1(\tau Right)_\tau + \mu_t + \gamma_c + \beta X_{ct} + \epsilon_{ct} \quad (3b)$$

For air pollution:

$$Y_{ct} = \alpha + \theta_1 1(SCAPRange)_\tau + \theta_2 1(SCAP)_\tau * (SCAPRange)_\tau + \theta_3 1(SCAPRange)_\tau * \tau + \theta_4 1(\tau Left)_\tau + \theta_5 1(\tau Right)_\tau + \rho_1 1(CCRange)_\phi + \rho_2 1(CC)_\phi * (CCRange)_\phi + \rho_3 1(CCRange)_\phi * \phi + \rho_4 1(\phi Left)_\phi + \rho_5 1(\phi Right)_\phi + \mu_t + \gamma_c + \beta X_{ct} + \epsilon_{ct} \quad (3b)$$

Lastly, the third specification includes a Policy*Time-Trend*Policy-Range interaction term (analogous to Equation (2c)). In the case of water pollution:

$$Y_{ct} = \alpha + \theta_1 1(NRCPRange)_\tau + \theta_2 1(NRCP)_\tau * (NRCPRange)_\tau + \theta_3 1(NRCPRange)_\tau * \tau + \theta_4 1(NRCP)_\tau * \tau * (NRCPRange)_\tau + \theta_5 1(\tau Left)_\tau + \theta_6 1(\tau Right)_\tau + \mu_t + \gamma_c + \beta X_{ct} + \epsilon_{ct} \quad (3c)$$

For air pollution:

$$Y_{ct} = \alpha + \theta_1 1(SCAPRange)_\tau + \theta_2 1(SCAP)_\tau * (SCAPRange)_\tau + \theta_3 1(SCAPRange)_\tau * \tau + \theta_4 1(SCAP)_\tau * \tau * (SCAPRange)_\tau + \theta_5 1(\tau Left)_\tau + \theta_6 1(\tau Right)_\tau + \rho_1 1(CCRange)_\phi + \rho_2 1(CC)_\phi * (CCRange)_\phi + \rho_3 1(CCRange)_\phi * \phi + \rho_4 1(CC)_\phi * \phi * (CCRange)_\phi + \rho_5 1(\phi Left)_\phi + \rho_6 1(\phi Right)_\phi + \mu_t + \gamma_c + \beta X_{ct} + \epsilon_{ct} \quad (3c)$$

As Equation 2c is our preferred estimation strategy, we provide the results from Equation 3c in Tables 3 and 4 for comparison with the two-stage approach.

Appendix Table 1: Summary Statistics, by Year

	Panel A: Air Quality			*Panel B: Water Quality*			*Panel C: Infant Mortality*
Year	PM	SO_2	NO_2	BOD	Ln(Fcoli)	DO	IMR
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
1986	--	--	--	4.0	6.1	7.4	--
	--	--	--	111	100	110	--
1987	279.3	21.0	29.3	3.8	6.5	7.2	35.9
	20	17	18	122	101	119	87
1988	243.6	18.7	23.2	3.2	6.3	7.2	30.5
	25	25	25	190	162	195	88
1989	245.7	18.2	26.2	3.3	6.7	7.2	28.2
	31	31	31	221	166	224	87
1990	249.1	19.9	24.9	3.9	6.6	7.2	24.4
	44	43	43	273	205	275	96
1991	256.2	20.3	28.8	3.8	5.6	7.2	26.9
	47	46	45	264	212	268	80
1992	223.5	20.9	28.5	4.4	6.1	7.2	26.2
	58	57	57	287	206	286	76
1993	257.2	22.5	27.9	4.6	5.9	7.3	25.3
	66	64	63	306	246	307	79
1994	249.3	22.2	28.5	4.1	5.7	7.3	22.1
	58	65	65	327	256	324	79
1995	216.0	22.2	23.3	4.7	5.1	7.1	21.8
	42	41	43	327	285	327	78
1996	244.9	21.9	25.9	4.7	4.9	7.2	22.4
	70	68	68	328	279	326	58
1997	213.6	23.6	30.0	4.4	4.9	7.1	19.6
	76	72	72	338	290	335	59
1998	229.8	21.7	27.7	4.1	4.9	7.1	18.7
	68	66	66	329	282	334	58
1999	221.5	22.1	29.1	4.4	4.9	7.1	18.1
	77	75	75	326	265	324	55
2000	193.5	16.7	27.8	4.0	4.9	7.1	18.3
	68	71	71	313	270	311	51
2001	223.0	14.7	27.3	4.2	5.1	7.1	17.8
	57	69	70	377	321	367	63
2002	189.4	13.0	26.7	4.6	5.2	7.1	15.7
	67	77	77	389	335	380	61
2003	230.3	11.9	26.5	4.2	5.1	7.0	17.6
	76	76	76	393	349	382	62
2004	221.4	11.6	26.1	3.7	5.1	6.9	14.6
	82	85	85	417	370	417	30
2005	210.6	12.7	26.7	4.1	5.9	7.1	--
	97	86	93	310	285	308	--
2006	202.9	12.0	25.0	--	--	--	--
	116	104	115	--	--	--	--
2007	206.7	12.4	24.9	--	--	--	--
	125	106	124	--	--	--	--

Notes:

1. The figures depict annual mean pollution levels. There are no restrictions on the sample. Annual means are first taken across all monitors within a given city, and then across all cities in a given year.

2. Number of city-year observations is reported below the mean in each year.

3. Units are ug/m^3 for all air pollutants; mg/l for BOD and DO; Ln(MPN/100 ml) for Fcoli; and Deaths per 1000 Births for infant mortality.

4. Pollution data were drawn from Central Pollution Control Board's online and print sources. Infant mortality data were taken from the book *Vital Statistics of India* as well as various state registrar's offices.

Appendix Table 2: Pollution Policy Coverage in India

	Air		*Water*
Year	Supreme Court Action Plans	Catalytic Converters	National River Conservation Plan
	(1)	(4)	(6)
1986	--	--	25
1987	0	0	25
1988	0	0	25
1989	0	0	25
1990	0	0	25
1991	0	0	25
1992	0	0	25
1993	0	0	25
1994	0	0	97
1995	0	0	160
1996	0	4	160
1997	0	4	160
1998	1	4	160
1999	1	49	160
2000	1	49	160
2001	1	49	160
2002	1	49	160
2003	12	49	160
2004	17	49	160
2005	17	49	160
2006	17	49	--
2007	17	49	--

Notes:

1. This table describes the incidence of each policy across all of India.

2. Policy data were drawn from a variety of sources that are detailed in the text.

Appendix Table 3: Distribution of Air and Water Policies, by Relative Year

	Air		*Water*
Relative Year	Supreme Court Action Plans	Catalytic Converters	National River Conservation Plan
	(1)	(4)	(6)
-17	1	0	0
-16	6	0	0
-15	6	0	0
-14	10	0	0
-13	12	0	0
-12	12	0	0
-11	11	5	0
-10	14	8	0
-9	13	10	13
-8	10	19	14
-7	15	20	26
-6	16	22	30
-5	16	23	37
-4	16	18	37
-3	16	18	41
-2	16	26	39
-1	16	26	36
0	16	20	42
1	15	26	42
2	16	24	42
3	16	19	39
4	11	22	43
5	1	25	43
6	1	24	43
7	1	24	39
8	1	24	42
9	1	24	39
10	1	4	38
11	0	4	11
12	0	4	11

Notes:

1. This table describes the prevalence of policies in the dataset by year *relative to policy uptake* . Thus, a value of '-17' in the first column signifies '17 years before a policy is implemented'. We use relative year ('tau') as the year variable in event study regressions and figures.

2. Cities which enact a policy are subjected to the same inclusion rules as in event study analysis. That is, a city is only included here if it has at least one pollution data point 3 years or more before policy uptake as well as at least one data point 4 years or more after policy uptake.

3. In this table, a city is counted if it has *any* pollution data in that year. In analyses, a city is only used if it has pollution data for the specific dependent variable in a given regression. Thus, the above city counts must be interpreted as maximums among our regressions. Most city-years have available data for all pollutants studied here.

4. Since infant mortality data is only available up to the year 2004, the table above is not representative for analysis which uses infant mortality as the dependent variable.

5. Policy data are drawn from a variety of sources detailed in the text.

Appendix Table 4: Robustness Checks with Varying Taus

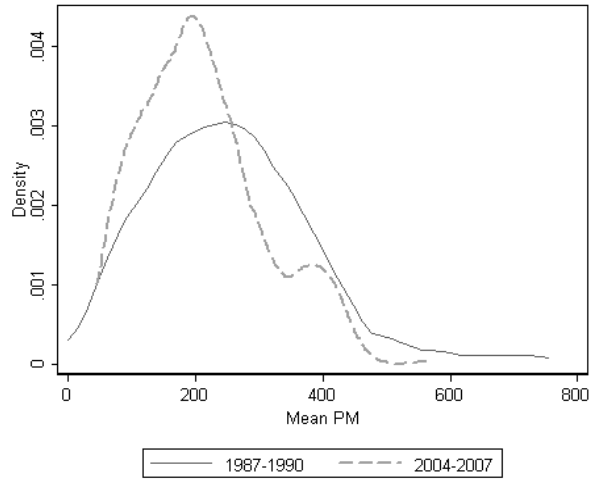
	Narrow Tau Range			Wide Tau Range		
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. Supreme Court Action Plans						
	PM	SO_2	NO_2	PM	SO_2	NO_2
5-Year Effect	23.6	-0.5	-20.8***	-61.2**	1.4	-4.1
p-value	0.38	0.88	0.00	0.01	0.62	0.22
N	9	9	9	19	19	19
Panel B. Catalytic Converters						
	PM	SO_2	NO_2	PM	SO_2	NO_2
5-Year Effect	-54.1*	-9.2*	-6.5	-30.0	-14.9***	-3.2
p-value	0.09	0.08	0.27	0.17	0.00	0.32
N	11	11	11	19	19	19
Panel C: National River Conservation Plan						
	BOD	Ln(Fcoli)	DO	BOD	Ln(Fcoli)	DO
5-Year Effect	2.0	0.1	-0.4	15.5***	1.3	-0.4
p-value	0.56	0.95	0.20	0.00	0.11	0.20
N	11	11	11	22	22	22

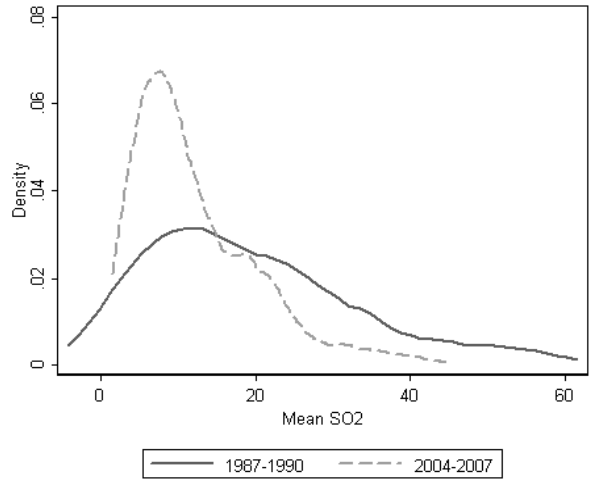
Notes:

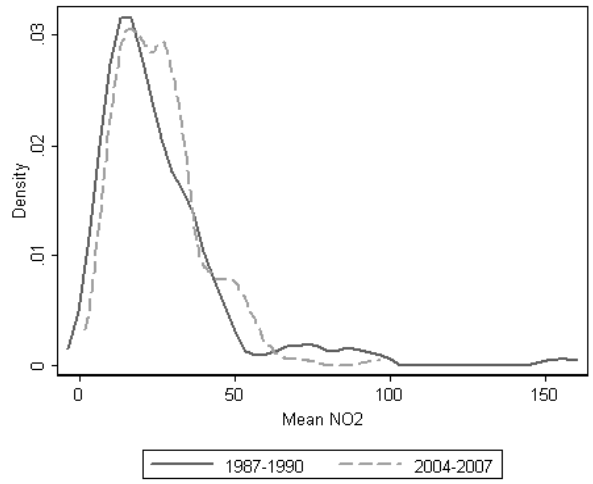
1. This table provides robustness checks for the prefereed econometric specification (Equation 2c).

2. Note that for the Supreme Court Action Plan, Narrow Tau Range, the effect is calcuated at 4 years out (rather than 5 years) due to limited data

Appendix Figure 1A: Kernel Density Graphs of Air Quality







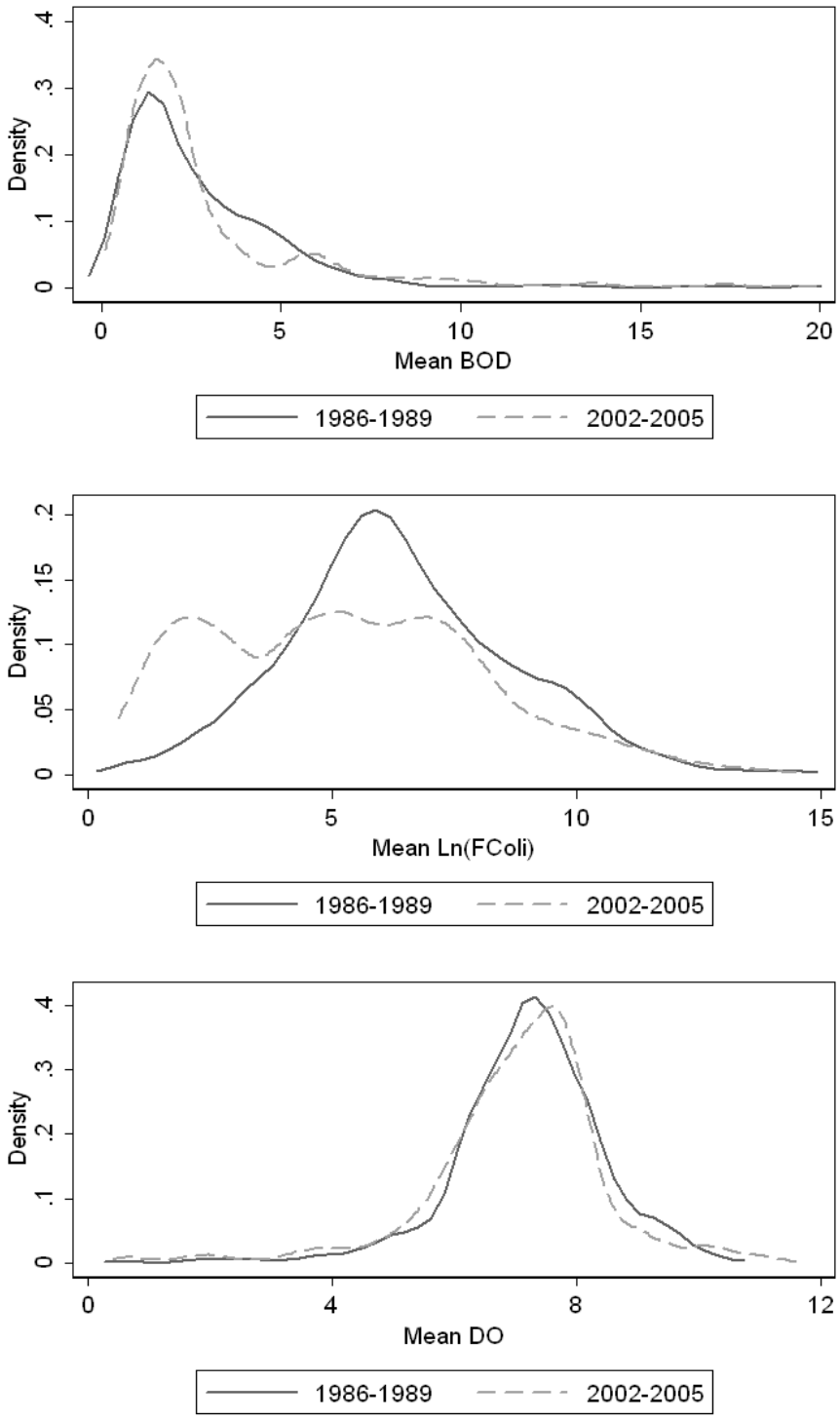
Notes:

1. The graphs provide the distribution of air pollution for the early (1987-1990) and later (2004-2007) periods of our sample. They are constructed using an Epanechnikov kernel function.

2. Units are $\mu g/m^3$ for all pollutants.

3. The data were drawn from Central Pollution Control Board's online and print sources.

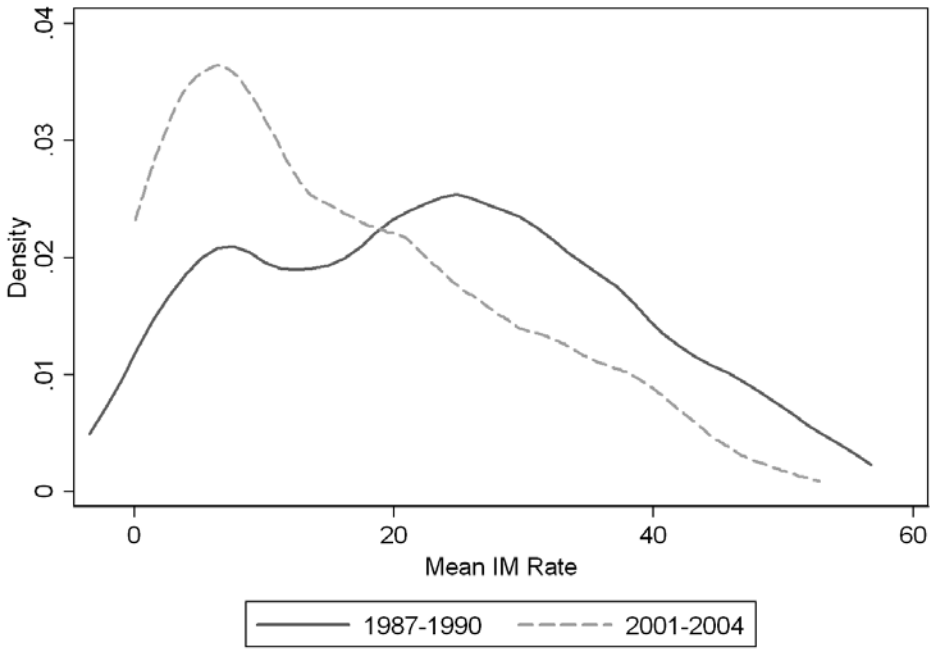
Appendix Figure 1B: Kernel Density Graphs of Surface Water Quality



Notes:

1. The graphs provide the distribution of water pollution for the early (1986-1989) and later (2002-2005) periods of our sample. They are constructed using an Epanechnikov kernel function.
2. Units are mg/l for BOD and DO, and Ln(MPN/100 ml) for FColi.
3. The right tail of the BOD distribution extends to 100 mg/l but is truncated to provide a more detailed picture of the distribution.
4. The data were drawn from Central Pollution Control Board's online and print sources.

Appendix Figure 1C: Kernel Density Graphs of Infant Mortality



Notes:

1. The graph provides the distribution of infant mortality for the early (1987-1990) and later (2001-2004) periods of our sample. They are constructed using an Epanechnikov kernel function.
2. Infant mortality data are restricted to those cities which have at least one air or water pollution measurement in the full sample.
3. Units of infant mortality rate are Deaths per 1000 Births.
4. The right tail of the IM distrbution extends above 100 but is truncated to provide a more detailed picture of the distribution.
5. The data were drawn from the book *Vital Statistics of India* as well as various state registrar's offices.