EPIC INDIA
ENERGY POLICY INSTITUTE
AT THE UNIVERSITY OF CHICAGO

EVIDENCE FOR POLICY DESIGN
Center for International Development
at Harvard University

Odd-Even Program: Analysis Q&A

Summary

The Delhi government piloted the odd-even program during January 1-15, 2016 as a measure to reduce the alarming air pollution levels in the city. The program restricted odd-numbered cars to odd-numbered days, and even cars to even numbered days: in principle, halving the number of private cars on the roads. The program was accompanied by a few other measures such as pressing more private buses to service on Delhi roads to complement public transportation, and extending the winter break to schools.

We measure the impact of the odd-even program on PM2.5 concentrations in New Delhi using a difference-in-differences (DiD) approach. To do this, we compare air pollution levels before and after the program, within and outside Delhi in the National Capital Region, where the program was not implemented.

We use hourly data (from November 2015- January 2016) from CPCB and IndiaSpend monitors in Delhi and in Faridabad, Noida and Gurgaon for our analysis. In Delhi, we are relying on 7 monitors from CPCB and 16 from IndiaSpend. There are fewer monitors outside Delhi in the NCR- one CPCB monitor in Faridabad, and one each from IndiaSpend in Noida and Gurgaon.

We estimate that the program reduced concentrations by about 35-47 µg/ m^3, or an estimated 10-13% on average. A statistically significant additional 10% reduction in emissions is estimated on average between 8am-8pm; air quality levels in the evening after 7pm are relatively unaffected. These results are supported by another set of hourly regression models which show stark reductions in PM concentrations around the noon hours, which then taper off in the late afternoon and evening.

This is a working document and will be updated to incorporate additional analysis and data, as and when they become available. We invite comments and suggestions which we could respond to or include in future versions of this document. (Last updated: January 15, 2015)

Table of Contents

Q&A: Methodology

Pollution levels have increased in January 2016, as compared to end December 2015. How can the program have reduced pollution?

Although absolute levels have increased both within and outside Delhi in the NCR, the levels in Delhi have seen a smaller increase. We argue that this can be attributed to the program, and apply a popular impact evaluation method called differences-in-differences to quantify the reduction.

Particulate Matter (PM) has many different sources: like thermal power plants, construction activity, road dust, vehicles, burning of solid fuels for cooking and heating, trash burning. Particulate Matter concentrations are also affected by weather conditions: low temperatures and low wind speeds typically result in higher concentrations of PM. Ambient concentrations in Delhi are highest during November- January, and this year, there has been an increase from late December to January likely due to weather conditions (low windspeeds, temperature and precipitation). However, it is likely that this is still lower than what it would have been in the absence of a program restricting one source of pollution, i.e. cars, and the challenge is in separating out the component that has reduced.

We have estimated this using a statistical technique called difference-in-differences by comparing ambient monitoring data from within Delhi and outside Delhi in the NCR during November 1, 2015 to January 8, 2016. The NCR has very similar weather conditions throughout and are equally exposed to external factors affecting pollution, like crop burning in the nearby states. As a result, the program impact can be estimated as the differences in the changes in pollution levels, before and after the program.

What are the data sources?

We primarily rely on data from the continuous ambient air quality monitoring systems, maintained by the Central Pollution Control Board (CPCB) along with the Delhi Pollution Control Committee and the Haryana Pollution Control Board. These data are shared online by CPCB, and can be accessed at

In addition, we also use data from 18 monitoring devices that have been developed and deployed by IndiaSpend in and around the National Capital Region, including in Noida and Gurgaon. We are currently renting five of these monitors from them.

How does difference-in-differences work?

We use a standard statistical technique for impact evaluation called difference-in-differences. By looking at data before and after a program both in the area where the program has been implemented and a suitable area where the program has not been implemented, we can estimate the change *because* of the program.

	Before Program	After Program	Change during the time where program is implemented
Area with program	B1	A1	(A1-B1)
Area without program	B2	A2	(A2-B2)

Change due to program in the area where program is implemented	**(A1-B1) – (A2-B2)**

We compare the changes in PM2.5 concentration levels before and after the program in Delhi monitors and outside Delhi in the NCR: Faridabad, Gurgaon and Noida .While the odd-even program was in place for commuters from these cities to Delhi, the odd-even program was not directly implemented there. If anything, the impact on the commuters makes our estimates lower bounds to the true impact.

Why can't we compare the pollution levels in Delhi in the weeks before and during the program, or last January with this January?

As already mentioned, particulate matter levels are affected by many factors, including the weather. While comparing before and after in the same area, we risk not accounting for all these factors appropriately and hence not being able to separate the impacts of different factors.

Why can't we compare the pollution levels between Delhi and Faridabad (or Gurgaon, Noida) directly? Why difference-in-differences?

There are typically some local factors that could change pollution levels. In fact, even within Delhi, Anand Vihar for instance tends to have a much poorer air quality than say, Dwarka.

Difference in differences therefore relies purely on the difference in changes as opposed to a difference in absolute levels. This way, we can estimate the impact of external factors affecting a larger region such as the NCR or Delhi.

In addition, we separately account for local effects and temperature in our regression models.

What are the assumptions being made while using differences-in-differences?

The major assumption with this method is that of 'parallel trends'. That is, in the absence of the program, the area with and without the program would have seen a similar change in the outcome variable (here, PM concentrations). In other words, we
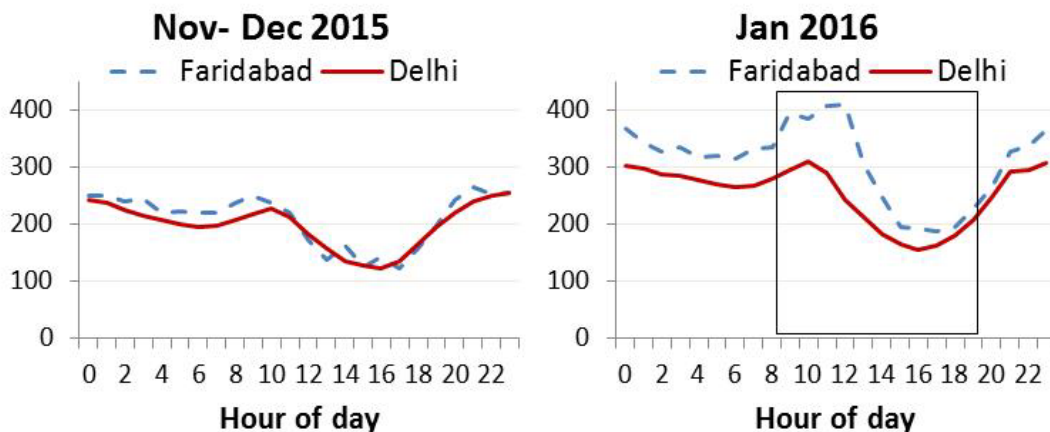
are assuming that pollution levels within and outside Delhi should change in a similar way due to external factors like weather. That way, any additional change that is seen can be attributed to dissimilar factors like the odd-even program being implemented only in Delhi.

For example, if we look at the monthly averages in the Delhi monitors and the Faridabad monitor from June 2015 (from when data from the Faridabad monitor are regularly available), we find that the two are very similar to each other. It is only in January 2016 that the Faridabad monitors have clearly higher levels of pollution.
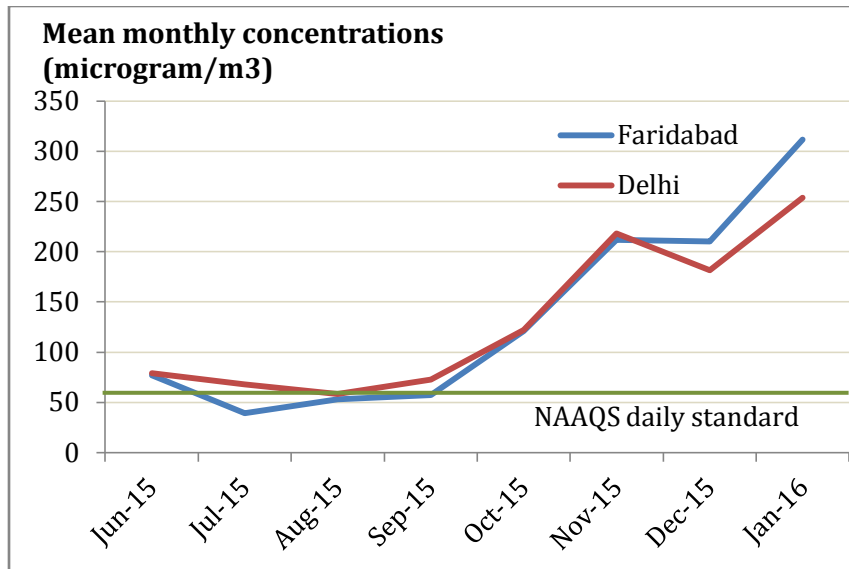
We could go further and look at hourly averages during November-December 2015 and January 2016. While the Delhi and Faridabad columns are virtually identical before the program, we would find that

1) Pollution levels in January are higher for both
2) Faridabad's levels have increased more than those in Delhi

Faridabad (minus) Delhi

(Chart: Mean PM2.5 concentrations in microgram/m3 vs Hour of day, comparing Nov- Dec 2015 and Jan 2016)
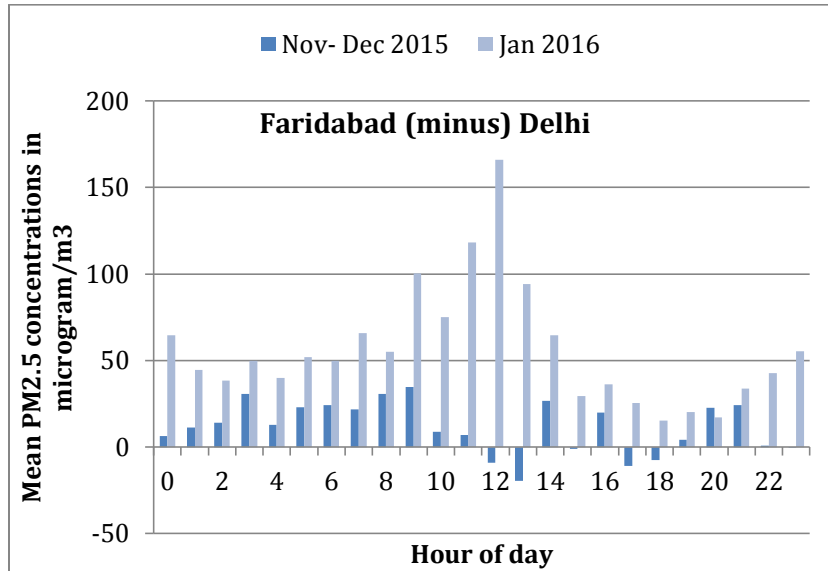
The results would be similar if we consider Noida and Gurgaon in addition too.

The estimate seems too high. There are reports that say vehicles account for only 25% of the air pollution in Delhi of which only a fraction are due to cars—how could reducing half the cars make so much difference?

The odd-even program reduces pollution in two key ways
i)   Fewer cars on the road—thereby, directly removing some of the polluting sources
ii)   Reduced congestion would reduce idling and slow moving traffic across the city, thereby reducing pollution for everybody.

This second factor is especially not well understood, and may well have been a major factor here, but this question merits further investigation. In addition, there are several estimates

Are EPIC and EPOD doing any additional research about the program?

Yes, we are.

Collaborating with researchers from other groups—the Council on Energy, Environment and Water (New Delhi), the University of British Columbia and IIM Ahmedabad— we are conducting traffic surveys to measure changes in vehicle composition and volumes on the roads.

We have also rented 5 monitors from IndiaSpend, who are also supporting the collaborative effort by sharing their data from all the monitors in the NCR.

Appendix: Regression Analysis

Model Specification for Point Estimates of program impact

We use standard specifications for difference-in-difference models to conduct the regressions to estimate the percentage change in PM2.5 concentrations due to the program. There is an indicator variable for the area where the program is implemented, for the time when the program is implemented and the interaction variable for program area and time. The coefficient of the interaction term ("e" in the equation below) is used to measure the impact of the program.

PM2.5Conc = a + b. Delhi + c. Year2016 + d. eight_to_eight + **e.Delhi_Year2016**

Where,
- PM2.5Conc are hourly concentrations on $PM_{2.5}$ from each monitor,
- Delhi is a dummy variable for whether the monitor is in Delhi (Delhi=1) or not (Delhi=0),
- Year2016 is a dummy variable =1, if the observation is in the year 2016
- eight_to_eight is a dummy variable =1, if the observation is between 8am and 8 pm during which hours the program implemented.
- Delhi_Year2016 is an interaction term between Delhi and Year2016 (Equal to 1, if Delhi=1 and Year2016=1)

In addition to the variables in the basic model above, other controls are included as shown in the table while running the regressions.

	(1)	(2)	**(3)**	(4)	**(5)**
Monitor Fixed effects	N	Y	**Y**	Y	**Y**
Month dummy	N	Y	**Y**	Y	**Y**
Triple difference	N	N	**N**	Y	**Y**
Weather	N	N	**Y**	N	**Y**
Hour dummy	N	Y	**Y**	Y	**Y**

The triple difference variable is included to investigate if there is a reduction in the hours of the day (8am-8pm) when the program is active. The triple difference variable is therefore the interaction term between Delhi, Year2016 and eight_to_eight, and forms an alternative way to measure the program impact.
Models 3 and 5 are preferred specifications with the weather controls having been included.

Each observation in the dataset is for one hour from a single monitor. Model results here involve observations only from **November 1, 2015 to January 8, 2016**. Weather parameters as measured at the CPCB monitors are also included in a few models.

We run analysis on the data from the CPCB monitors alone, and from CPCB and IndiaSpend monitors together. Within Delhi, the major CPCB monitors in the data set with high data availability are
1. Anand Vihar
2. Mandir Marg
3. IHBAS
4. Punjabi Bagh
5. Shadipur
6. RK Puram
7. Dwarka

Outside Delhi, we are using the Faridabad ambient air quality monitor—which is also maintained by the regulators. Note that there are a few other government monitors in Delhi as well as monitors in Gurgaon and Noida, but $PM_{2.5}$ data from these are not available on the CPCB website[1] for the period of interest.

Table 1 Results from the regression on the model using only data from CPCB monitors between November 1, 2015- January 8, 2016. t statistics in parentheses, * p<0.05, ** p<0.01, ***- p<0.001.

	(1)	(2)	(3)	(4)	(5)
Delhi	-10.18***	0	0	0	0
	(-3.39)	(.)	(.)	(.)	(.)
Year2016	102.7***	89.85***	69.66***	76.72***	59.03***
	(11.86)	(20.42)	(16.57)	(17.31)	(13.21)
Eight_to_eight	-53.76***	-35.22**	-58.01**	-44.17**	-55.07*
	(-23.89)	(-5.06)	(-4.54)	(-3.97)	(-2.92)
delhi_year2016	-47.96***	-54.19**	**-35.35 #**	-37.35*	**-16.63**
	(-5.10)	(-5.21)	**(-2.29)**	(-3.17)	**(-1.01)**
year2016_eighttoeight				25.90***	20.41***
				(124.40)	(6.72)
delhi_eighttoeight				11.18	-0.904
				(0.94)	(-0.10)
delhi_eighttoeight_year2016				-33.25**	**-36.24*****
				(-3.90)	**(-7.62)**
Constant	236.3***	264.6***	260.1*	264.6***	254.8*
	(81.68)	(34.93)	(3.31)	(34.93)	(3.21)
Number of observations	11194	11194	8909	11194	8909

#- The p-value for this result is 0.055 and hence this is still quite statistically significant.

The percentage reduction can be estimated by dividing the sum of the coefficient of delhi_Year2016 and the sum of the coefficients obtained for constant, Delhi and Year2016 terms in the table above.

This yields estimates the following point estimates for percentage reduction of PM2.5 concentrations.

[1] www.cpcb.gov.in/caaqm/frmUserAvgReportCriteria.aspx Last accessed on 10 January, 2016

	(1)	(2)	(3)	(4)	(5)
Estimated average reduction all day	-15%	-15%	-11%	-11%	-5%
Estimated average additional reduction between 8am-8pm				-11%	-14%

Models 3 and 5 are the preferred specifications. Model 3 provides an estimated reduction of mean daily concentrations by **35 µg/ m^3 (or an 11% reduction)**. Much of this seems to be through reductions during the morning and afternoon hours. As per model 5, there is a statistically significant reduction of about 36 µg/ m^3 between 8am-8pm.

With the same specifications, the results remain similar when IndiaSpend monitors[2] are included in addition to the CPCB. Here we include an additional dummy variable for the source of the data (= 0 for CPCB, 1 for IndiaSpend). IndiaSpend data are primarily between December 24, 2015 and January 8, 2016. There are two IndiaSpend monitors in Gurgaon and Noida in our dataset which makes the results less dependent on Faridabad for the difference-in-differences assumptions

Table 2 Results from the regression on the model using data from CPCB and IndiaSpend monitors between November 1, 2015- January 8, 2016. t statistics in parentheses, * p<0.05, ** p<0.01, *- p<0.001.**

	(1)	(2)	(3)	(4)	(5)
Delhi	-6.338*	0	0	0	0
	(-2.44)	(.)	(.)	(.)	(.)
Year2016	98.06***	84.71***	69.66***	84.71***	59.03***
	(17.91)	(10.91)	(16.57)	(10.91)	(13.21)
Eight_to_eight	-56.66***	-96.37***	-58.01**	-96.37***	-55.07*
	(-30.96)	(-6.84)	(-4.54)	(-6.84)	(-2.92)
delhi_year2016	-27.39***	**-36.86*** **	-35.35	-36.86***	-16.63
	(-4.64)	**(-3.91)**	(-2.29)	(-3.91)	(-1.01)
year2016_eighttoeight				18.89**	20.41***
				(3.05)	(6.72)
delhi_eighttoeight				4.395	-0.904
				(0.48)	(-0.10)
delhi_eighttoeight_year2016				**-26.16** **	-36.24***
				(-2.94)	(-7.62)
Constant	232.2***	270.7***	260.1*	270.7***	254.8*
	(91.14)	(44.01)	(3.31)	(44.01)	(3.21)
Number of observations	17865	17865	8909	17865	8909

With the IndiaSpend monitors included, the coefficients remain very similar. As IndiaSpend data did not have weather data, models 3 and 5 will yield almost identical results as in the previous table. Model 2 and 4 results are most interesting in this case (without the weather controls). There is an estimated reduction of **37 µg/ m^3 (or a 13% reduction)**. Between 8am-8pm, an additional reduction of 27 µg/ m^3 is estimated.

---

[2] Real time data from IndiaSpend can be accessed at aqi.indiaspend.org/aq/analysis

	(1)	(2)	(3)	(4)	(5)
Estimated average reduction all day	-10%	-13%	-10%	-9%	-5%
Estimated average additional reduction between 8am-8pm				-10%	-11%

Model Specification for hourly estimates of program impact

We have also run a series of 24 models for each hour of the day separately to estimate the change in PM2.5 concentration, if any, at each hour of the day. While this reduces the statistical power by restricting the number of observations on which the regression is run, the model becomes more flexible in accounting for weather and other factors during the day.

PM2.5Conc = a + b. Delhi + c. Year2016 + **d. Delhi_Year2016** + e. Temp + f. TempSquared + g. WindSpeed

With Monitor Fixed Effects, where Temp and WindSpeed are Temperature and Wind Speed are measured at the CPCB monitors. We assume a quadratic functional form for the impact of temperature on PM2.5 concentrations.
The results of the 24 hour-wise regressions are plotted below, along with the estimated standard errors. The results are most significant in the morning and afternoon hours—however, this is not the case in the rest of the day where the results are by and large not statistically significant.
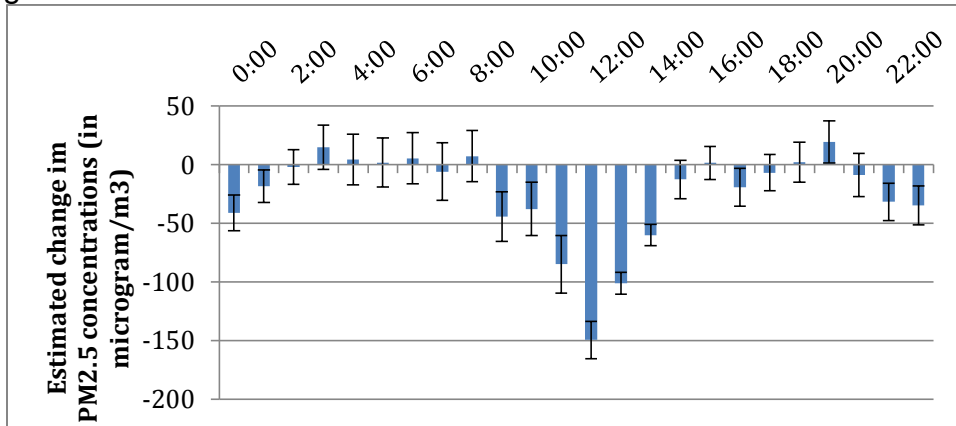

Figure 1 Estimated reductions in PM concentration by time of day using CPCB monitors. Standard errors have been plotted for each column

Once again, we redo the analysis including the IndiaSpend monitors, which include locations outside NCR, in Noida and Gurgaon. The results follow similar trends and we have statistical significant reductions seen over a larger part of the morning hours.
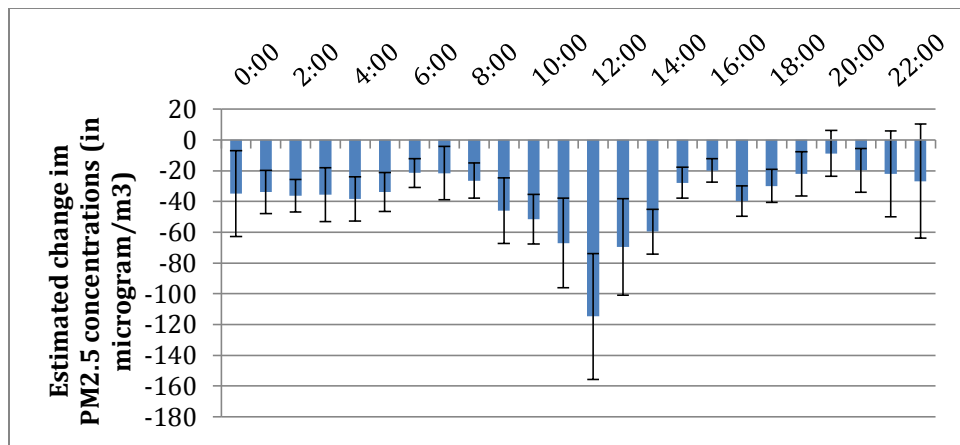
Figure 2 Estimated reductions in PM concentration by time of day using CPCB and IndiaSpend monitors. Standard errors have been plotted for each column

Further investigation is required to understand why the hourly distribution looks the way it does and why the pollution levels in the evening seem broadly unaffected. One hypothesis could be that dispersion over the course of the day within the NCR area cancels out any local reductions due to the program; but this requires further study.