Title: Fetal Exposures to Toxic Releases and Infant Health

Authors:
- *Corresponding*: Janet M. Currie
  (janet.currie@columbia.edu; phone: 212-854-4520)
- Johannes F. Schmieder
  (jfs2106@columbia.edu; phone: 631-903-5646)

Address: Department of Economics
      Columbia University
      420 W 118th St.
      New York, NY 10025

Session Title: Air Pollution and Health around the World

Session Chair: Janet Currie (Columbia University)

Session Discussants:
- Heather Royer (Case Western Reserve University)
- Matthew Neidell (Columbia University)
- Lucas Davis (University of Michigan)

Fetal Exposures to Toxic Releases and Infant Health

Janet Currie and Johannes F. Schmieder[*]

Every year, millions of pounds of toxic chemicals thought to be linked to developmental problems in fetuses are released into the air. Yet, we have only limited information about the health effects of these releases. A 1998 Environmental Protection Agency (EPA) review found that complete screening data about toxicity was available for only 7 percent of 3,000 chemicals released in large quantities in the U.S. (U.S. EPA, 1998). Even for chemicals that have been studied, there is little information about how levels found in the environment affect human health. Laboratory data on toxicity may be of limited value given that tests are typically conducted on animals, and do not take human behaviors (such as staying inside on high pollution days) into account.

Moreover, it is quite difficult to draw a relationship between a disease such as cancer and toxic exposures in a particular location given that cancer develops over a long period, and people are mobile. In contrast, birth outcomes are likely to be highly affected by conditions during the brief interval of pregnancy (though of course they might also be influenced by factors affecting the mother before conception). Hence, infant health outcomes are an ideal place to look to see if existing environmental releases have detectable negative effects for human health.

This study uses data from the EPA's Toxics Release Inventory (TRI) matched to data from national Vital Statistics Natality and Mortality files to examine the effects of fetal exposure on health at birth and subsequent infant mortality. Exposure to toxic chemicals may be linked to many other characteristics of families and neighborhoods, and to swings in economic activity.

In an effort to identify the effect of toxic exposures, we compare the estimated effects of chemicals that are thought to be developmental toxicants to those which are not known to have developmental effects. We also compare the effects of "fugitive" air releases to the effects of "stack" air releases. Emissions that go up a smoke stack are more likely to be treated in some fashion (e.g. with scrubbers), and travel further than those that do not. Hence, they should be less likely to affect those in the immediate vicinity of the plant. Finally, we look at several of the most common known developmental toxicants separately.

II. Data and Methods

Information about pregnancy outcomes comes from the Vital Statistics Natality data. They cover virtually all births and include information about characteristics of the mother, characteristics of the child, and health at birth. Information on infant deaths is taken from Vital Statistics Mortality files. We focus on birth weight, gestational age, and infant mortality in the first year of life, since there is considerable variation across counties in these outcome measures.

Data on toxic releases comes from the Environmental Protection Agency's Toxic Release Inventory, which was created by the Emergency Planning, Community Right to Know Act (EPCRA) in 1986. EPCRA was a legislative response to the 1984 Bhopal disaster, in which a cloud of deadly methyl isocyanate escaped from a Union Carbide plant and killed thousands of people. A Union Carbide plant in West Virginia had a serious chemical release a short time later. These incidents added urgency to claims that communities had a "right to know" about hazardous chemicals that were being used or produced in their midst.

EPCRA required manufacturing plants (SIC=2000 to 3999) with more than 10 full-time employees that either use or produce more than threshold amounts of listed toxic substances to report releases to the EPA for public disclosure. Plants are required to file a separate form for

each substance and must identify whether the release was to ground, water, or air. We focus on air-borne releases because people living close to a plant may be more likely to be exposed to them than to water or ground releases. The previous calendar year's toxic releases are required to be reported by July 1. Data from the TRI are publicly available from the EPA on CD-Rom or on the internet.

The data are quite extraordinary and clearly the best available for our project.[1] Several studies have examined compliance to the reporting regulations and data quality (see Gerald V. Poje and Daniel M. Horowitz (1990), John Brehm and James T. Hamilton (1996), Thomas E. Natan and Catherine G. Miller (1998), Scott de Marchi and James T. Hamilton (2006), and Dinah A. Koehler and John D. Spengler (2007)). While these papers point to some underreporting, overall compliance was high and the decrease in reported releases corresponds to changes in plant operations and production levels.

There have been several changes in the TRI requirements for reporting. In 1995, the list of chemicals tracked was expanded; in 1998, the type of facilities required to report was expanded, and in 2000, the thresholds for the reporting of chemicals that persist in the environment were lowered and a few chemicals have been de-listed. We focus on a set of chemicals and industries where reporting requirements were consistent between 1988 and 1999 and exclude releases are to off-site facilities (where the exact location is not known).

In 1999, about 23,000 facilities submitted reports describing the releases of more than 2.3 billion pounds of toxic substances. This was however, a decline of approximately 40 percent from 1988 levels. There is considerable variation in the quantities of toxic releases across coun-

---

[1] The National Air Toxics Assessment has more complete information about air toxics. However, it is currently available only for 1996 and 1999, and the two waves are not strictly comparable because of changes in chemicals tracked and in the dispersion models used to form estimates of the toxics present in each location.

ties of the U.S., with the southern states accounting for a disproportionate and increasing share of releases over time.

One common criticism of the TRI is that it tracks only a subset of the many chemicals in widespread use in the U.S. Hence, it is impossible to know if the overall declines in reported releases track total releases of toxics over time, or if, for example, companies simply substitute from listed to unlisted chemicals where possible. We can however, examine the estimated effects of the most commonly released individual chemicals that are thought to affect reproductive success or to affect fetal, infant, or child development (developmental toxicants). Estimates of the effects of individual chemicals are of considerable interest in their own right.

We use information from the California Office of Environmental Health Hazard Assessment's (OEHHA) to identify TRI chemicals that are known developmental toxicants. This determination is made on the basis of the available evidence, for example from studies of animals exposed to high levels of these substances. Eighty of the chemicals on the OEHHA list are tracked in the TRI. This list enables us to distinguish between developmental toxicants and other toxic chemicals (which might, for example, be carcinogens). In addition to looking at these 80 chemicals as a group, we focus on 10 important developmental toxicants and divide them into two broad classes that could be expected to have very different actions in the body: Volatile organic compounds (VOCs), which generally diffuse quickly into the air, and heavy metals, which are highly persistent in the environment.[2] Turning to individual chemicals, we examine toluene and lead, which accounts for the vast majority of the VOC and heavy metal air emissions, as well as epichlorohydrin and cadmium, which are thought to be among the most dangerous develop-

---

[2] The heavy metals we examine are lead, cadmium, arsenic and mercury. The VOCs are benzene, carbon disulphide, dibromoethane, epichlorohydrin, ethylene oxide, and toluene. We chose these chemicals on the basis of frequency of releases and known toxicity.

mental chemicals. We have classified compounds of these chemicals together with the main chemical.

We do not know the date of TRI releases, only that they occurred in a particular calendar year. This raises the question of how best to merge the TRI data with the Natality data, where we know the month of birth. We focus on births in January of each year. For these births, most of the pregnancy occurred in the previous calendar year.

Because very low birth weight births and infant deaths are rare events, when we look at these outcomes, we expand the sample to include all births from January to March. For all of these children, at least six months of their time in utero occurred in the previous calendar year, and so they have a greater than 50 percent chance of having been exposed to the releases reported for that year. Focusing on January births has the additional advantage of controlling for any seasonality in birth/death outcomes.

Having made these selections, we

Table 1: Means for Sample and Unmerged Births, and Maximum Emissions

Outcomes	Merged Birth & TRI	Unmerged
Gestation (weeks)	38.99	39.08
	[0.305]	[0.241]
Birth weight (grams)	3328.4	3345.2
	[77.59]	[71.86]
Low birth weight	0.0748	0.0695
(<2500 grams)	[0.0235]	[0.0161]
Very low birth weight	0.0137	0.0111
(<1500 grams)	[0.00873]	[0.00495]
Deaths per 1,000	8.111	7.822
(in 1st Year)	[6.311]	[3.498]
Selected Controls		
Mother Age	26.79	26.07
	[1.438]	[1.049]
Mother Education	12.82	12.58
	[0.766]	[0.508]
Mother Black	0.148	0.0768
	[0.158]	[0.121]
Mother Hispanic	0.108	0.0682
	[0.158]	[0.0968]
Smoking	0.165	0.195
	[0.109]	[0.0550]
TRI Fugitive Air Releases in pounds per square mile		
Developmental	212.4	88.59
	[791.1]	[893.7]
Non Developmental	840	375.4
	[1993.9]	[3909.3]
VOCs	200.8	83.5
	[775.0]	[821.3]
Toluene	176.5	69.21
	[742.6]	[721.7]
Epichlorohydrin	0.391	0.142
	[4.319]	[4.552]
Heavy Metals	1.691	0.356
	[12.33]	[8.648]
Lead	1.518	0.334
	[11.82]	[8.610]
Cadmium	0.116	0.00538
	[2.273]	[0.103]
# County*Year Cells	5279	19898

Notes: Standard deviations of the county means (for unmerged birth data of the state means) in brackets. Maximum values of toxic releases in parentheses.

aggregate the TRI data, the birth data, and the death data to the county level. Since counties are

of wildly different sizes we rescale the toxic releases by dividing them by county area. We esti-

mate models in which the outcome (gestation, birth weight or the infant death rate) depends on

toxic releases and control variables obtained by aggregating the Natality data. All of our models

control for county-year level means of indicators for maternal age, race, ethnicity, education,

whether the mother smoked during pregnancy and how many cigarettes per day, whether the

mother drank during the pregnancy and the number of drinks per day, as well as child gender.

We also control for year and county fixed effects in order to deal with overall time trends and

differences between counties. Furthermore using the County Business Pattern data[3] we control

for county employment as a proxy for the business cycle. Finally, we weight the regressions us-

ing the average number of births in each county over the sample period and cluster standard er-

rors on the county level to control for serial correlation.

Counties with populations of less than 100,000 are not identified in the Natality data.

Hence, our sample consists of relatively large counties, which represent about 75 percent of all

US births.   Table 1 presents summary statistics for both the merged TRI/birth sample and for

those counties which have TRI data, but have populations less than 100,000 so that the county is

not identified and cannot be merged with the TRI data. It should therefore be kept in mind that

the results discussed below apply to a relatively urban population and not necessarily to rural

counties.

III. Results

Our main results are shown in Table 2. Table 2 shows strong evidence that within-county varia-

tions in reported toxic releases are related to infant health outcomes: Nearly all coefficients indi-

cate a negative effect on birth outcomes and are quite precisely estimated. The only wrong

---

[3] See http://www.census.gov/epcd/cbp/index.html

signed coefficient is the estimate of the effect of lead on infant death which has a very large

standard error. Moreover, a comparison of sections 2 and 3 indicates that the estimated effects

are much larger for developmental releases than for non-developmental releases as one would

expect if the estimates truly reflect effects of the chemicals and not the effects of other factors

(such as economic cycles) that might be linked to similar variations in emissions of both types of

chemicals.

Table 2: Effects of Fugitive Air Toxic Releases on Infant Health

	[1] Gestation (weeks)	[2] Birth Weight (grams)	[3] Low Birth Weight	[4] Very Low Birth Weight	[5] Infant Death per 1000 births
1. All Releases	-0.0115	-1.458	0.398	0.0829	0.106
	[0.00210]	[0.342]	[0.0727]	[0.0377]	[0.0248]
2. Developmental Chemicals	-0.0246	-2.845	0.861	0.229	0.248
	[0.00603]	[1.097]	[0.203]	[0.0871]	[0.0543]
3. Non Developmental Chemicals	-0.0146	-1.928	0.503	0.0914	0.129
	[0.00371]	[0.491]	[0.117]	[0.0576]	[0.0433]
4. Volatile Organic Compounds	-0.0245	-2.837	0.874	0.235	0.247
	[0.00627]	[1.136]	[0.202]	[0.0881]	[0.0556]
4.a) Toluene	-0.0245	-3.19	0.977	0.25	0.276
	[0.00703]	[1.125]	[0.191]	[0.0884]	[0.0509]
4.b) Epichlorohydrin	-0.927	-172.6	12.57	19.65	13.81
	[0.676]	[73.57]	[27.70]	[5.628]	[7.204]
5. Heavy Metals	-0.879	-180	56.26	10.44	8.725
	[0.475]	[98.02]	[37.13]	[9.813]	[10.55]
5.a) Lead	-0.362	-78.82	16.02	3.696	-1.725
	[0.269]	[53.05]	[10.33]	[8.716]	[7.394]
5.b) Cadmium	-2.679	-531.5	211.7	41	45.28
	[0.232]	[40.24]	[15.61]	[3.945]	[3.228]

Notes: Each coefficient is from a separate regression. Pollution units are in thousand pounds per square mile. Standard errors in brackets, clustered on county level. Coefficients and standard errors on LBW and VLBW multiplied by 1000. There are 5279 observations.

Turning to types of chemicals, and estimates of the effects of individual chemicals, Table

1 shows that toluene accounts for the 78 percent of the fugitive air emissions of developmental

chemicals we focus on, as well as 83 percent of the VOC fugitive air releases that we focus on.

It is not surprising then that toluene is estimated to have negative effects which are similar in

magnitude, though slightly larger than those for all developmental chemicals. The estimated ef-

fects of epichlorohydrin and heavy metals, including lead and cadmium are much larger. While the estimates for heavy metals and lead are relatively noisy, epichlorohydrin and cadmium have highly statistically significant negative effects on gestation, birth weight, and the probability of infant death.

However, taken at face value, the magnitude of some of the effects is small. For example, the point estimate of the effect of toluene on gestation in row 5 implies that an additional one thousand pounds per square mile of toluene emissions in a county would reduce gestation by only 0.024 weeks. Given Table 1, this is about a third more than a one standard deviation change. Similarly, the point estimate of the effects of toluene on birth weight implies a 3.2 gram change in birth weight per thousand pound per square mile change in emissions. The coefficients on heavy metals suggest that a two standard deviation change in lead emissions would reduce gestation by 0.008 weeks, and would reduce birth weight by 1.8 grams on a mean of 3,300 grams. Finally, although the estimated coefficients on cadmium and epichlorohydrin are very large, they reflect the fact that relatively small amounts are released. A two standard deviation change in cadmium releases would decrease gestation by 0.012 weeks, and would decrease birth weight by 2.4 grams, while a two standard deviation change in epichlorohydrin would decrease gestation by 0.01 weeks and birth weight 1.8 grams.

These modest effects on the overall means of gestation and birth weight mask the fact that there are sizeable effects on the probability of a newborn child being of low (< 2500 grams) and very low (< 1500 grams) birth weight: a two standard deviation change in cadmium releases would increase the probability of low birth weight by 1.2 percent and the probability of very low birth weight by 1.4 percent. For toluene (epichlorohydrin) the effect of a two standard deviation

increase in releases would be to increase the incidence of low and very low birth weight by 1.9 (0.2) and 2.7 (1.5) percent respectively.

The effects of releases on infant death are of comparably large magnitude. The estimated effect of toluene on deaths implies that a two standard deviation change in toluene emissions would increase deaths by 0.405 on a baseline of about 8 deaths per 1,000 live births, while a two standard deviation change in cadmium would increase the death rate by about 5 percent. More-over reported toluene releases decreased from an average of 340 pounds per square mile to about 75 pounds per square mile over the sample period. For the counties in the sample, this decrease accounts for 3.2 percent of the overall reduction in infant mortality and translates to about 220 fewer infants deaths in 2000. Similarly the reductions in lead and cadmium imply 9 and 40 fewer infant deaths. Reductions in these three chemicals alone can account for about 3.9 percent of the reduction in infant mortality during the late 80s and 90s from 9.2 to 6.9 deaths per 1000 live births.

The distribution of releases is extremely skewed, as is indicated by very large standard deviations and maximum values relative to the means. For example, the maximum release of lead is 433.3 pounds per square mile, nearly 300 times the mean. It may not be very plausible to assume the effect to be linear over such a wide range of releases. Our estimates may be driven by severe effects on health from large releases. With the small number of such releases in the sample, it is unfortunately not feasible to estimate this nonlinearity with any precision.

Table 3 shows an important specification check. As predicted, fugitive air releases have larger negative effects than stack air releases and the inclusion of the stack air coefficients has essentially no impact on the fugitive air coefficients (this holds for all specifications in Table 2,

available in the online Appendix Table 3). We view this as additional evidence that we are detecting an actual health effect, rather than the effect of omitted variables correlated with emissions. Presumably both fugitive and stack air emissions vary with fluctuations in economic activity, but fugitive air would be expected to have greater effects on health. Indeed, the effect of stack air emissions is sometimes positive, as one might expect if they picked up the effect of upswings in economic activity, for example. This is never the case with fugitive air emissions, which always have negative estimated effects on health. The contrasting patterns between fugitive and stack air emissions suggest once again that we are underestimating the effect of fugitive air releases on health. If positive economic conditions improve health but also increase emissions in industrial areas, then the true health effect might more accurately be measured by the difference between the stack air and fugitive air coefficients. As an additional check we estimated the model controlling for state specific time trends. The results are very similar (available in Appendix Table 4).

Our results are quite robust to the exact specification chosen. We also estimated models where we include the toxic releases in absolute values, rather than divided by the county area. This may be more appropriate if the county area is a bad approximation for what fraction of births in a county is actually af-

Table 3: Comparison of Effects of Fugitive and Stack Air Toxic Releases

	[1] Gestation (weeks)	[2] Birth Weight (grams)	[5] Infant Death per 1000 births
Dev. Chemicals Fugitive Air	-0.0243 [0.00621]	-2.865 [1.194]	0.238 [0.0587]
Dev. Chemicals Stack Air	-0.00064 [0.00458]	0.0487 [0.979]	0.0248 [0.0674]
VOC's Fugitive Air	-0.0244 [0.00642]	-2.87 [1.223]	0.239 [0.0598]
VOC's Stack Air	-0.00026 [0.00466]	0.0892 [1.012]	0.0218 [0.0699]
Heavy Metals Fugitive Air	-0.867 [0.470]	-176.9 [96.36]	8.439 [10.34]
Heavy Metals Stack Air	-0.252 [0.375]	-64.1 [61.40]	5.89 [4.411]

Notes: Each coefficient is from a separate regression. Pollution units are in thousand pounds per square mile. Standard errors in brackets, clustered on county level. There are 5279 observations.

fected by a given release. The results from these regressions, (available in the online Appendix Table 1), are fairly similar to our main results. One marked difference is that in this specification heavy metals and lead have a very clear negative effect on all health measures. All quantitative predictions (such as the explained part of the decrease in infant mortality) are very similar.

We also estimated models (available as Appendix Table 2 online) using data from December births (December to February for the rare outcomes) to show that our results are not driven by focusing on January births. While the effects tend to be slightly smaller, the overall pattern is quite similar. The discrepancy could arise because Toxic Releases may be concentrated at the end of the year and are perhaps particularly harmful at certain periods during fetal development. Finally, we asked whether positive correlations between toxic releases could cause an upward bias for the coefficients on the toxics in regressions that only include one toxicant. To check for this we estimated models controlling for all of the individual chemicals simultaneously. This change had little impact on either the levels of the effects or the precision of the estimates.

There is good reason to view the estimates discussed above as extreme lower bounds on the effects of toxic releases given the measurement issues discussed above. These estimates reflect the mean effect over children who may have been exposed to large doses of toxic chemicals at critical periods, and other children who may not have been exposed at all, or who may have been exposed at times that they were not vulnerable to injury. This logic suggests that if we could measure actual exposures delivered to particular children at critical periods while they were in utero, it is likely that the estimated effects would be much larger. This is consistent with the fact that we find relatively small effects on mean birth weight and gestation but larger effects on the more extreme outcomes of low birth weight, very low birth weight, and infant death.

11

Brehm, John and James T. Hamilton. "Noncompliance in Environmental Reporting: Are Violators Ignorant, or Evasive, of the Law?*" American Journal of Political Science*, 1996, 40(2), pp. 444-477.

De Marchi, Scott and James T. Hamilton. "Assessing the Accuracy of Self-Reported Data: An Evaluaton of the Toxics Release Inventory," *Journal of Risk and Uncertainty*, 32, 2006:57-76.

Dutzik, Tony, Jeremiah Baumann, and Meghan Purvis. 2003. "Toxic Releases and Health: A Review of Pollution Data and Current Knowledge on the Health Effects of Toxic Chemicals." *U.S. PIRG Education Fund*,

http://www.nypirg.org/enviro/pests/Toxic%20Releases%20final.pdf

Koehler, Dinah A. and John D. Spengler. "The Toxic Release Inventory: Fact or Fiction? A Case Study of the Primary Aluminum Industry," *Journal of Environmental Management*, 85, 2007: 296-307.

Natan, Thomas E. and Catherine G. Miller. "Are Toxic Release Inventory Reductions Real?" *Environmental Science and Technology*, 32(15), 1998: 368A-374A.

Poje, Gerald V. and Daniel M. Horowitz. "Phantom Reductions: Tracking Toxic Trends", National Wildlife Federation, Washington, DC (1990).

U.S. Environmental Protection Agency, Office of Pollution Prevention and Toxics. April 1998. *Chemical Hazard Data Availability Study: What Do We Really Know About the Safety of High Production Volume Chemicals?* http://www.epa.gov/HPV/pubs/general/hazchem.pdf

Woodruff, Tracey J., Daniel A. Axelrad, Jane Caldwell, Rachel Morello-Frosch, and Arlene Rosenbaum. 1998. "Public Health Implications of 1990 Air Toxics Concentrations Across the United States." *Environmental Health Perspectives*, 106(5): 245-251.