

**WP 2019-05**  
**May 2019**



# Working Paper

Department of Applied Economics and Management  
Cornell University, Ithaca, New York 14853-7801 USA

## **A NARRATIVE ON TWO WEAKNESSES OF THE TRI FOR RESEARCH PURPOSES**

**Neha Khanna**

It is the Policy of Cornell University actively to support equality of educational and employment opportunity. No person shall be denied admission to any educational program or activity or be denied employment on the basis of any legally prohibited discrimination involving, but not limited to, such factors as race, color, creed, religion, national or ethnic origin, sex, age or handicap. The University is committed to the maintenance of affirmative action programs which will assure the continuation of such equality of opportunity.

## **A Narrative on two weaknesses of the TRI for research purposes**

**Neha Khanna**

**May 2, 2019**

Key words: TRI, data quality, facility, firm, Dun and Bradstreet

JEL codes: Q53, Q50

### **Abstract**

I focus on two issues commonly encountered when working with facility-level data from the Environmental Protection Agency's Toxic Releases Inventory (TRI). First, is the issue of zero emissions and the second relates to the linkage between a facility and its parent firm as reported via the Dun and Bradstreet number in the TRI. My purpose is to make researchers aware of the issues and to suggest potential remedies where possible.

## 1. Introduction

The U.S. Environmental Protection Agency's (EPA) Toxics Release Inventory (TRI; <https://www.epa.gov/toxics-release-inventory-tri-program/tri-data-and-tools>) provides facility level annual information on the emissions and management of more than 600 toxic chemicals including several carcinogens, dioxins and persistent, bioaccumulative, toxins (PBTs), from 1987 onward.<sup>1</sup> These data are reported annually to the EPA by all facilities from key manufacturing and industry sectors that meet the basic reporting requirements. The reporting facilities self-estimate their emissions and management of the chemicals using various engineering models (such as mass balance equations). While there is no direct measurement of the emissions data, the EPA routinely performs data quality checks (<https://www.epa.gov/toxics-release-inventory-tri-program/tri-data-quality>) that assure users of the accuracy of the data. Consequently, the TRI is widely used by regulators, researchers, community organizations and other not-for-profit agencies and private individuals for information on toxic releases in the U.S.<sup>2</sup>

In this narrative, I highlight two issues in the TRI database that have prominently affected my work as an academic researcher. Unless otherwise stated, the examples are based on the universe of data reported in the TRI Basic Plus between 2000 and 2017 (accessed on Dec 4, 2018): 1,621,428 chemical-facility-year level observations and 416,735 facility-year level observations. There are 42,213 unique TRI facilities (identified by their TRI facility IDs) that are associated with 22,370 unique Dun & Bradstreet (D&B) identification numbers (variable D and B NR A in the TRI) and 7,272 unique parent firm D&B identification numbers (variable PARENT COMPANY D and B NR in the TRI).<sup>3</sup>

## 2. Zero total production

For any given year, the TRI includes information on the following categories: total releases, total energy recovery (on-site and off-site), total quantity recycled (on-site and off-site) and total quantity treated (on-site and off-site). The sum of these categories is 'total production'. There is a separate category for one-time accidental or catastrophic releases. A facility must report this information to the TRI if it meets *all three* of the following criteria (<https://www.epa.gov/toxics-release-inventory-tri-program/basics-tri-reporting>): (i) is in a TRI reporting sector (see <https://www.epa.gov/toxics-release-inventory-tri-program/tri-covered-industry-sectors>, but they generally cover manufacturing, mining and electric power generation), (ii) employs 10 or more full-time (or equivalent) employees, and (iii) manufactures, processes or otherwise uses a TRI reportable chemical above the specified threshold. The standard thresholds are manufacture or process above 25,000 lbs per year or otherwise use in excess of 10,000 lbs per year. There a few

---

<sup>1</sup> The full list of currently reportable chemicals, as well as a record of historical changes in the chemicals list, is available at <https://www.epa.gov/toxics-release-inventory-tri-program/tri-listed-chemicals>.

<sup>2</sup> For example, the U.S. National Library of Medicine has developed a publicly accessible GIS tool, TOXMAP (<https://toxmap.nlm.nih.gov>) that draws upon the data reported in the TRI.

<sup>3</sup> This excludes 15 parent firm D&B numbers that do not correspond to the standard reporting format and appear to be erroneous.

exceptions like PBTs which have a far lower threshold of either 100 lb or 10 lb for manufacture, process or otherwise use.<sup>4</sup>

Given the reporting requirements, I expect that if a facility did not meet the three criteria simultaneously in a particular year, it would not report information for that year. Yet, out of the universe of 416,735 facility year observations, there are 73,212 observations where a facility reports zero total production. This does not meet the third reporting criterion and the facility should not be reporting information to the TRI.

Table 1 includes examples for three facilities that report zero total production.<sup>5</sup> Facility 1 reports a zero value for all variables that make up total production for 16 of the 18 years between 2000 and 2017; small positive values for total releases are reported in 2003 and 2004.<sup>6</sup> Similarly, Facility 2 reports zero values in 17 of the 18 years, and a small positive value for total releases in 2013.<sup>7</sup> Facility 3 reports zero values for 17 of the 18 years, while it does not report any values at all for 2010.

It is not clear to me why these facilities report zero values to the TRI. Indeed, even the positive quantities reported in some years by Facility 1 and Facility 2 are well below the reporting thresholds. Are these facilities reporting this information voluntarily to the TRI? Are the zero values truly zero values? Can a researcher assume that a facility is voluntarily reporting any quantities that fall below the TRI reporting thresholds? Likewise, why doesn't Facility 3 consistently report zero values for all 17 years; what makes 2010 different from the other years so that this facility does not report any information to the TRI?

In my own work, I have often trimmed the analysis sample to exclude observations with very low values. Thus, facilities like the four highlighted in Table 1 would normally be excluded from my analyses. However, for some analyses, zero values are critical and a researcher might want to know whether an observation with a zero reported value should be treated any differently from a facility that does not report to the TRI at all. Likewise, for facilities with releases below the reporting threshold, what distinguishes facilities that voluntarily report releases to the TRI from facilities that choose not to report?

---

<sup>4</sup> Dioxins and dioxin-like compounds have an even lower reporting threshold of 0.1 gms.

<sup>5</sup> In a relatively scarce fourth category, facilities report positive values for one-time accidental or catastrophic releases, while total production is zero. An example is TRI Facility ID = 99336CHVRNFINLE that reports accidental releases in 2000, 2001 and 2002.

<sup>6</sup> There is a discrepancy in the data reported for this facility in the TRI Basic and TRI Basic Plus data files. For example, in 2001 and 2002 the facility reports zero total releases in the TRI Basic Plus data files but small positive quantities for copper compounds, arsenic compounds, chromium compounds and lead in the TRI Basic data files. All of the positive quantities reported in the TRI Basic data files fall below the reporting thresholds for TRI chemicals as well as for lead, which is the only PBT. Furthermore, while the facility reports positive total releases in both data files in 2003 and 2004, the magnitudes are different.

<sup>7</sup> In the case of Facility 2, identical data are reported to the TRI Basic and TRI Basic Plus files.

**Table 1: Total Production (pounds)**

Reporting year	Total releases	Energy recovery (on site)	Energy recovery (off site)	Quantity recycled (on site)	Quantity recycled (off site)	Quantity treated (on site)	Quantity treated (off site)	Total produced	One time catastrophic releases
<u>Facility 1 (01007NRTHS201OL)</u>									
2000	0	0	0	0	0	0	0	0	0
2001	0	0	0	0	0	0	0	0	0
2002	0	0	0	0	0	0	0	0	0
2003	2393	0	0	0	0	0	0	2393	0
2004	211.78	0	0	0	0	0	0	211.78	0
2005	0	0	0	0	0	0	0	0	0
2006	0	0	0	0	0	0	0	0	0
2007	0	0	0	0	0	0	0	0	0
2008	0	0	0	0	0	0	0	0	0
2009	0	0	0	0	0	0	0	0	0
2010	0	0	0	0	0	0	0	0	0
2011	0	0	0	0	0	0	0	0	0
2012	0	0	0	0	0	0	0	0	0
2013	0	0	0	0	0	0	0	0	0
2014	0	0	0	0	0	0	0	0	0
2015	0	0	0	0	0	0	0	0	0
2016	0	0	0	0	0	0	0	0	0
2017	0	0	0	0	0	0	0	0	0
<u>Facility 2 (01007NVRSL149BA)</u>									
2000	0	0	0	0	0	0	0	0	0
2001	0	0	0	0	0	0	0	0	0
2002	0	0	0	0	0	0	0	0	0
2003	0	0	0	0	0	0	0	0	0
2004	0	0	0	0	0	0	0	0	0
2005	0	0	0	0	0	0	0	0	0
2006	0	0	0	0	0	0	0	0	0
2007	0	0	0	0	0	0	0	0	0
2008	0	0	0	0	0	0	0	0	0
2009	0	0	0	0	0	0	0	0	0
2010	0	0	0	0	0	0	0	0	0
2011	0	0	0	0	0	0	0	0	0
2012	0	0	0	0	0	0	0	0	0
2013	543	0	0	0	0	0	0	543	0
2014	0	0	0	0	0	0	0	0	0
2015	0	0	0	0	0	0	0	0	0
2016	0	0	0	0	0	0	0	0	0
2017	0	0	0	0	0	0	0	0	0

Facility 3 (31533GLDNPBUSHE)

2000	0	0	0	0	0	0	0	0	0	
2001	0	0	0	0	0	0	0	0	0	
2002	0	0	0	0	0	0	0	0	0	
2003	0	0	0	0	0	0	0	0	0	
2004	0	0	0	0	0	0	0	0	0	
2005	0	0	0	0	0	0	0	0	0	
2006	0	0	0	0	0	0	0	0	0	
2007	0	0	0	0	0	0	0	0	0	
2008	0	0	0	0	0	0	0	0	0	
2009	0	0	0	0	0	0	0	0	0	
2010										
			not reported in either TRI Basic or TRI Basic Plus							
2011	0	0	0	0	0	0	0	0	0	
2012	0	0	0	0	0	0	0	0	0	
2013	0	0	0	0	0	0	0	0	0	
2014	0	0	0	0	0	0	0	0	0	
2015	0	0	0	0	0	0	0	0	0	
2016	0	0	0	0	0	0	0	0	0	
2017	0	0	0	0	0	0	0	0	0	

---

### 3. Facility and parent Dun and Bradstreet (D&B) number

Each facility is required to report to the TRI identifying information on itself and its parent firm, which is the firm at the highest level of ownership. Each facility is identified by its name, location, and a D&B number; likewise each parent firm is identified by, among other things, a firm D&B number and firm name.

It is not unusual to find that in any given reporting year, multiple facilities, in different locations, report the same facility DUNS number. Based on my experience, this occurs when all facilities with the same facility D&B number belong to the same immediate parent firm (whose name is not reported to the TRI). D&B assigns these numbers to facilitate financial reporting, so it is reasonable that facilities belonging to same firm have the same facility DUNS number. This is useful information for researchers because it is a means for identifying facilities that belong to the same immediate parent firm. For example, in reporting year 2003, seven facilities, each with a unique facility name and location, report a facility D&B number of 7909013. A search in a third party database, Mergent Intellect, reveals that this D&B number is associated with a company named Pacific Corp; each of the unique facility names reported in the TRI also suggest that the facilities are different plants belonging to Pacific Corp.

In some cases, facilities do not report any facility D&B number at all. For example, of the 42,213 unique facilities in the TRI Basic Plus between 2000 and 2017, 11,237 facilities (accounting for 104,963 facility-year observations: see Table 2 for a year-wise breakdown) do not report a facility D&B number. 5,171 of these facilities, however, report a parent firm D&B number. Among these, 1,889 are single facility firms and the parent firm DUNS is the only DUNS associated with the facility.

The TRI is well known for the facility-firm linkage and many researchers (including my collaborators and myself) have utilized this information. However, many experienced researchers remain wary of the reliability of this information. The main issue is that in *any given reporting year*, multiple facilities report the same facility D&B number but different ultimate parent firm names, or vice versa (same parent name but different facility D&B number).



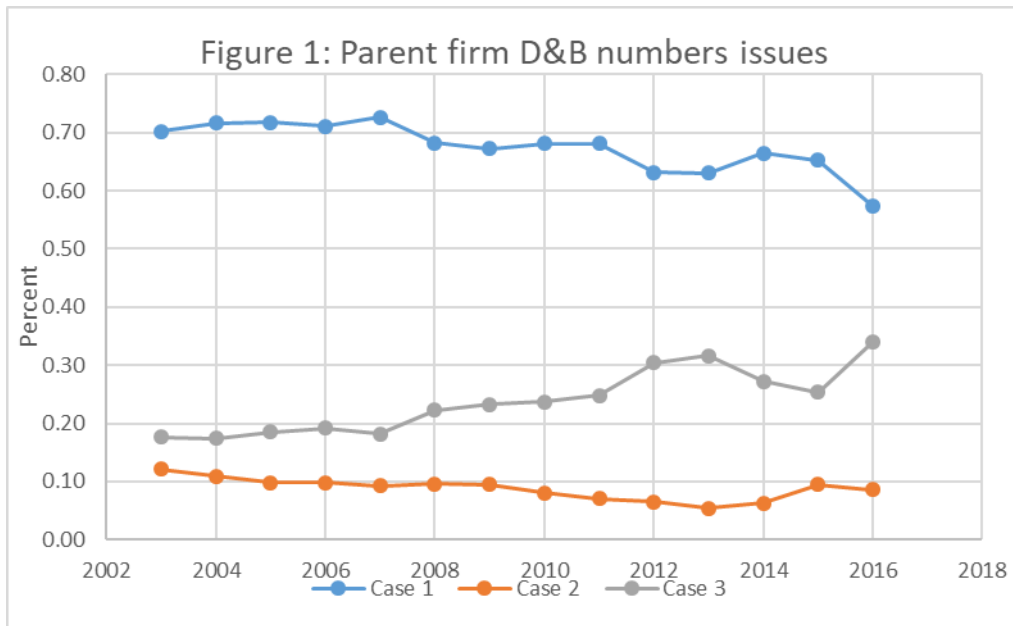
**Table 2: Missing facility and parent firm identifiers**

<b>Reporting year</b>	<b>Facility-year observations</b>	<b>No facility D&amp;B number</b>	<b>No parent D&amp;B number</b>	<b>No parent name</b>
2000	24194	5039	9554	6742
2001	25891	5695	10418	7243
2002	25240	5689	10164	6951
2003	24707	5689	10021	6737
2004	24491	5737	9939	6565
2005	24316	5781	9783	6354
2006	23870	5753	9538	6070
2007	23358	5802	9211	5681
2008	22784	5722	8847	5310
2009	21918	5600	8461	4993
2010	21809	5657	8443	5043
2011	21922	5740	8460	5059
2012	22109	5953	8472	5080
2013	22264	6134	8551	5138
2014	22305	6259	8603	5213
2015	22220	6353	8505	5191
2016	21881	6238	8438	5192
2017	21456	6122	8296	5098
<b>Totals</b>	<b>416735</b>	<b>104963</b>	<b>163704</b>	<b>103660</b>

Below I highlight four different types of data quality issues regarding D&B numbers that I have encountered in my work. The examples are drawn from the sample used in Rijal and Khanna (2019). This is a sub-sample of facility-year observations that excludes observations with no parent firm name or D&B number, zero total production and also all single facility firms from the universe of facility-year observations reported to the TRI between 2003 and 2016.<sup>8</sup> There are 63,464 facility-year level observations in this sample with 8,773 unique TRI facilities linked to 2,532 unique facility D&B numbers. 850 facility D&B numbers in this sample are linked to multiple parent firm names.

*Case 1: different parent name reporting protocols*

In 2005, both “GE IONICS BRIDGEVILLE” (TRIFID: 15017NCSNC3039W) operating in Allegheny, PA, and “GE IONICS” (TRIFID: 02172NCSNC65GRO) operating in Middlesex, MA, reported “1046325” as their facility D&B number. While the former reported “GE CO” as its parent company name, the latter reported “GENERAL ELECTRIC CO INC” as its parent company. In this case, the difference in the parent names can be explained by the differing nomenclature protocols followed by the two facilities and should not be considered an anomaly *per se*. The EPA seems to be aware of such cases and in recent years it has made an effort to standardize parent firm names (EPA 2016). Figure 1 shows the declining trend of such cases between 2003 and 2016 based on the sample used by Rijal and Khanna (2019).



<sup>8</sup> Rijal and Khanna exclude an additional 1,640 observations for reasons unique to their analysis including missing NAICS information, missing FRS identification numbers, and facilities located in US territories.

### *Case 2: wholly owned subsidiaries*

Under the TRI reporting instructions (EPA 2016), a facility should report its ultimate parent company's name and D&B number. So, for example, if a facility is wholly owned by firm A which in turn is a wholly owned subsidiary of firm B, then the facility should report firm B as its parent (not firm A). Yet, there are several cases that do not appear to follow the instructions. For example, in 2007 both "MACDERMID PRINTING SOLUTIONS" (TRIFID: 92069NPPSY360SO) operating in San Diego, CA, and "MACDERMID INC" (TRIFID: 48220MCDRM1221F) operating in Oakland, MI, report "1164599" as their facility D&B number. While the former reports "MACDERMID INC" as its parent company name, the latter reports "PLATFORM SPECIALTY PRODUCTS CORP" as its parent company. A search for the reported parent company names in a third party database, Mergent Intellect, reveals that Macdermid Inc is a wholly owned subsidiary of Platform Specialty Product Corp. If my understanding of the reporting instructions is correct, Platform Specialty Product Corp should be the reported parent company in both cases. There are 1,809 such cases in the TRI sub-sample used in Rijal and Khanna (2019), and Figure 1 shows the slowly declining trend over time. A variation of this case occurs when, in a given reporting year, the same parent firm name is associated with different facility D&B numbers. Table 3 reports three examples, extracted from the TRI Basic Plus for reporting year 2003, that highlight this issue.

Consider the first case: 24 facilities in the TRI report Berkshire Hathaway Inc. as their parent firm name; 5 different facility D&B numbers are associated with this parent firm in the TRI. Mergent Intellect data show that the facility D&B numbers reported in the TRI are associated with different company names, each of which is owned by a different ultimate parent firm. However, in each case the global ultimate parent firm is Berkshire Hathaway Inc. Similarly, 17 facilities report Bridgestone Americas Inc. as their parent firm in the TRI. This parent firm is associated with 4 different facility D&B numbers in the TRI. Mergent Intellect reports a different company name with each of the facility D&B numbers reported in the TRI. The ultimate parent firm for three out of four of these companies is Bridgestone Americas Inc., whereas the global ultimate parent firm is Bridgestone Corporation (Japan). The remaining facility D&B number is associated with a company named Keefer, Leelan but with no associated parent or global parent firm information reported in Mergent Intellect. In the final case, 17 facilities report Koch Industries Inc. as the parent firm name in TRI, associated with 7 different facility D&B numbers. Two of these D&B numbers are not associated with any company in Mergent Intellect; the remaining 5 have unique ultimate parent firms which are owned by the global parent Koch Industries Inc.

**Table 3: Examples of same parent firm name, multiple facility D&B numbers (2003)**

TRI Basic Plus				Mergent Intellect				
Parent firm name	Parent firm D&B	Facility D&B	# facilities reporting D&B	Company name	Ultimate parent firm	Ultimate parent firm D&B	Global ultimate parent firm	Global ultimate parent firm D&B
Berkshire Hathaway Inc	7909013	7909013	4	Pacific Corp	PPW Holdings LLC	79730265	Berkshire Hathaway	1024314
	8018772	8018772	10	Justin Industries	Berkshire Hathaway	1024314	Berkshire Hathaway	1024314
	45840055	45840055	6	Shaw Industries Group	Berkshire Hathaway	1024314	Berkshire Hathaway	1024314
	181837238	181837238	3	Benjamin Moore	Benjamin Moore	1210715	Berkshire Hathaway	1024314
	2227528	2227528	1	Owl Wire & Cable	Marmon Group	79587181	Berkshire Hathaway	1024314
Bridgestone Americas Inc	101479322, 1288109	1288109	5	Bridgestone Americas Tire Operations	Bridgestone Americas	101479322	Bridgestone Corporation (Japan)	48963558
	101479322, 111482506	809622447	6	Firestone Building Products Company	Bridgestone Americas	101479322	Bridgestone Corporation (Japan)	48963558
	101479322	868917949	2	Bridgestone Americas Tire Operations	Bridgestone Americas	101479322	Bridgestone Corporation (Japan)	48963558
	101479322	5268758	4	Keefer, Leelan	No ultimate parent information		No global parent information	
Koch Industries Inc	6944334	5246673	2	Molex LLC	Koch Industries	6944334	Koch Industries	6944334

6944334	6944334	3	Koch Industries	No ultimate parent information		Koch Industries	6944334
6944334	9020777	1	Georgia-Pacific LLC	Georgia Pacific Holdings	80226064	Koch Industries	6944334
6944334	48494033	2		DUNS not found in Mergent Intellect			
6944334	69634525	1		DUNS not found in Mergent Intellect			
6944334	143369416	3	Flint Hills Resources Alaska LLC	Flint Hills Resources LLC	102456447	Koch Industries	6944334
6944334	602084618	5	Koch Fertilizer	Koch Ag and Energy Solutions	79725044	Koch Industries	6944334

Table 4 provides a year-wise breakdown of such cases in the sub-sample used by Rijal and Khanna (2019). In any given year, well above a third of facility-year observations have a many-to-one mapping from facility D&B number to parent firm name.

**Table 4: Same parent firm name, multiple facility D&B numbers: 2003 to 2016**

Year	Facility-year observations	Many-one observations	%
2003	4773	1827	0.38
2004	4784	1896	0.40
2005	4697	1890	0.40
2006	4780	2007	0.42
2007	4818	2002	0.42
2008	4870	2101	0.43
2009	4683	2056	0.44
2010	4633	1969	0.42
2011	4615	1920	0.42
2012	4631	1952	0.42
2013	4611	1960	0.43
2014	4522	1869	0.41
2015	3586	1519	0.42
2016	3461	1463	0.42

*Case 3: unrelated firms*

Consider the following example. In 2000, six facilities report “1306992” as their D&B number. Those facilities are:

- “ILEVEL BY WEYERHAEUSER” (TRIFID: 9852WPCFCV1NRTH) operating in Crow Wing, MN
- “WEYERHAEUSER VALDOSTA” (TRIFID: 9852WPCFCV1NRTH) operating in Lowndes, GA
- “WEYERHAEUSER NR - ARCADIA OSB” (TRIFID: 9852WPCFCV1NRTH) Operating in Lincoln, LA
- “PACIFIC VENEER” (TRIFID: 9852WPCFCV1NRTH) Grays Harbor, WA
- “INTERNATIONAL PAPER CO” (TRIFID: 9852WPCFCV1NRTH) operating in McCurtain, OK
- “COSMO SPECIALTY FIBERS INC” (TRIFID: 9852WPCFCV1NRTH) operating in Grays Harbor, WA

They report the following parent company names, respectively:

- “WEYERHAEUSER”
- “WEYERHAEUSER CO”

- “WEYERHAEUSER NR CO”
- “WILLIS ENTERPRISES INC.”
- “INTERNATIONAL PAPER CO”
- “INTERNATIONAL PAPER CO”

The first three instances are examples of Case 1. The last two instances are examples of Case 2: in 2000, Weyerhaeuser Co, formed a 28% joint-venture with other equity members, including International Paper Co, to develop and operate a business-to-business marketplace for the forest products industry. This can be confirmed by searching for Weyerhaeuser or International Paper Co in Mergent Archives. However, the remaining case is a third type of anomaly because there appears to be no link between Willis Enterprises Inc and either Weyerhaeuser or International Paper Co in either Mergent Archives or Mergent Intellect. So this raises the question whether facility Pacific Veneer has reported an incorrect facility D&B number or an incorrect parent firm name, or whether the information reported to the TRI is correct and that the information in Mergent Archives/Intellect is incomplete.<sup>9</sup>

A variation on Case 3 occurs when two facilities report the same facility D&B number with different parent firm names, but Mergent Archives and Mergent Intellect do not contain information on any of the parent firm names reported to the TRI. This case is highlighted in the following example. In 2001, “ARMACELL ENGINEERED FOAMS” (TRIFID: 28613HLSTD1004K) operating in Catawba, NC, and “RBX INDUSTRIES COLT ARKANSAS PLANT” (TRIFID: 72326HLSTDHWY1) operating in St. Francis, AR, report “1217504” as their facility D&B number. While the former reports “ARMACELL US HOLDINGS INC” as its parent company name, the latter reports “RBX INDUSTRIES INC” as its parent company. However, neither Armacell US Holdings Inc nor RBX Industries are listed as companies in the Mergent Archive database; nor is the D&B number associated with any company in Mergent Intellect.

Figure 1 summarizes the yearly occurrence of both variations of Case 3 between 2003 and 2016 in the trimmed sample from Rijal and Khanna (2019).

*Case 4: missing parent name and/or parent DUNS number*

A final issue is the lack of parent firm identifiers. Out of the universe of 416,735 facility-year observations reported to the TRI Basic Plus files there are 163,704 observations with no parent firm D&B number and 103,660 observations with no parent firm name (with some overlap between these two categories). Table 2 provides a year-by-year breakdown. As many as 39-41% of facilities do not report a mandatory parent firm D&B identifier in any given year. Out of the 253,031 facility-year observations where parent firm D&B numbers are reported, 143 observations seem to report invalid parent firm D&B numbers.

---

<sup>9</sup> In Mergent Intellect, D&B number 1306992 is associated with Weyerhaeuser Company, which is also the global ultimate parent.

#### 4. Conclusion

In light of the issues highlighted above, it behooves researchers to be very careful when using TRI data to separate facilities that meet the reporting requirements for TRI from those that do not, report zero values for various reportable categories and for establishing facility-firm linkages. In many, though not all, cases facilities with very low levels of reportable emission categories are likely to (or should) be trimmed from the final analysis sample for statistical reasons unrelated to the TRI itself. However, when using the TRI to establish a facility-firm linkage, there are a few possible options to consider.

First, assume that the D&B numbers (facility and parent firm) reported to the TRI are accurate. Except in a small fraction of cases, this assumption seems to be generally borne out by the cross-checking against information available in Mergent Intellect/Archives. This implies that unique TRI facilities that report the same facility D&B number belong to the same immediate parent firm; likewise, facilities with the same parent firm D&B number always belong to same ultimate global parent firm, regardless of differences in reported parent firm name. Conversely, facilities with different facility D&B numbers belong to different immediate parent firms even if the reported ultimate parent firm name is the same (implying the same ultimate global parent).<sup>10</sup> Second, assuming erroneous reporting, eliminate all cases where there is a discrepancy between parent firm name and facility D&B number that cannot be explained by a simple difference in reporting nomenclature as in Case 1. Given the frequency of the issues noted above, this would eliminate a large fraction of facility-year observations and probably yield an unrepresentative sample. Third, and this is what I have done in much of my work, consult a third party database such as Mergent Archives and/or Mergent Intellect to establish firm history and family tree. While the third option yields the most meticulous dataset, it is also the most painstaking since each firm must be cross-checked manually by firm name or D&B number in Mergent Archives/Intellect. Case 4 presents an insurmountable problem because there is no identifying parent firm information. In my own work where the link between a facility and its parent firm was essential to the analysis, I have typically excluded any facility-year observations with missing parent firm identifiers.

---

<sup>10</sup> Over the years, I have made several attempts to contact Dun and Bradstreet with clarifications regarding the association between firm name and firm D&B number. To date, I have not received any response.



## **Acknowledgement**

I thank Binish Rijal for his research assistance and for providing examples from the TRI Basic Plus data files. All errors are mine.

## **References**

Environmental Protection Agency, 2016. *Toxic Chemical Release Inventory Reporting Forms and Instructions*. EPA 740-B-16-001, December.

Rijal B. and N. Khanna, 2019. Pollution Leakage and Voluntary Pollution Prevention. Working paper.

**OTHER A.E.M. WORKING PAPERS**

WP No	Title	Fee (if applicable)	Author(s)
2019-05	A Narrative on Two Weaknesses of the TRI for Research Purposes		Khanna N.
2019-04	Village in the City: Residential Segregation in Urbanizing India		Bharathi N., Malghan D., Rahman A.
2019-03	Inequality in a Global Perspective		Kanbur R.
2019-02	Impacts of Minimum Wage Increases in the U. S. Retail Sector: Full-time versus Part Time Employment		Yonezawa K., Gomez M., McLaughlin M.,
2019-01	Minimum Wages and Labor Supply in an Emerging Market: the Case of Mauritius		Asmal Z., Bhorat H., Kanbur R., Ranzani M., Paci P.
2018-17	Improving Economic Contribution Analyses of Local Agricultural Systems: Lessons from a study of the New York apple industry		Schmit, T., Severson, R., Strzok, J., and Barros, J.
2018-16	Public Goods, and Nested Subnational Units: Diversity, Segregation, or Hierarchy?		Bharathi, N., Malghan, D., Mishra, S., and Rahman, A.
2018-15	The Past, Present and Future of Economic Development		Chau, N., and Kanbur, R.
2018-14	Commercialization of a Demand Enhancing Innovation by a Public University		Akhundjanov, S. B., Gallardo, K., McCluskey, J., Rickard, B
2018-13	Sustainable Development Goals and Measurement of Economic and Social Progress		Kanbur, R., Patel, Ebrahim, and Stiglitz, J.
2018-12	Sustainable Development Goals and Measurement of Economic and Social Progress		Kanbur, R., Patel, E, and Stiglitz, J.
2018-11	Migrants, Towns, Poverty and Jobs: Insights from Tanzania		Christiaensen, L., Weerdt, J.D., and R. Kanbur
2018-10	Inequality and the Openness of Borders		Kanbur, R.
2018-09	Measuring Unfair Inequality: Reconciling Equality of Opportunity and Freedom from Poverty		Hufe, P., Kanbur, R., and A. Peichl
2018-08	Isolated by Caste: Neighbourhood-Scale Residential Segregation in Indian Metros		Bharathi, N., Malghan, D. and A. Rahman

Paper copies are being replaced by electronic Portable Document Files (PDFs). To request PDFs of AEM publications, write to (be sure to include your e-mail address): Publications, Department of Applied Economics and Management, Warren Hall, Cornell University, Ithaca, NY 14853-7801. If a fee is indicated, please include a check or money order made payable to Cornell University for the amount of your purchase. Visit our Web site (<http://dyson.cornell.edu/research/wp.php>) for a more complete list of recent bulletins.