Internet Appendix for "Not in My Backyard: Intrinsic Motivation and Corporate Pollution Abatement"

Angie Andrikogiannopoulou, Alexia Ventouri, and Scott Yonker*

*Angie Andrikogiannopoulou and Alexia Ventouri are at King's College London, Scott Yonker is at Cornell University.

In this appendix, we present some additional information and tables that support the results in the paper. In Section IA.A, we discuss in detail our method for matching TRI data with NETS and YTS data. In Section IA.B, we show information on the industries represented in our sample. And, in Section IA.C, we present some additional results.

IA.A Matching TRI with NETS and YTS

The TRI database has 636,300 observations for the period 1992–2018, of which 531,222 are in the period 1992–2013 (which overlaps with the NETS database) and 507,471 are in the period 1997–2018 (which overlaps with the YTS database). To match TRI with NETS and with YTS, we proceed sequentially in the following steps for each year in the data:

First, we match observations in which the facility street address and state are the same across the two databases and either (i) the facility city name is the same, or (ii) the facility zip code is the same, or (iii) the geodesic distance between the facility latitude-longitude in the two databases is less than 1km. In this step, we match with NETS 38% of the observations in TRI for the period 1992–2013. In matching TRI with YTS we skip this step, because YTS does not provide facility street addresses.

Second, we match observations in which the geodesic distance between the facility latitude-longitude in the two databases is less than 1km and the facility or parent name (or DUNS number when matching TRI with NETS) are identical across the two databases. In this step, we match with NETS 10.5% of the observations in TRI for the period 1992–2013. We also match with YTS 65% of the observations in TRI for the period 1997–2018.

Third, we match observations in which the geodesic distance between the facility latitude-longitude in the two databases is less than 1km and the facility or parent name—excluding the city name, if part of the name—are very similar (i.e., a fuzzy name match) across the two databases. In this step, we match with NETS 0.7% of the observations in TRI for the period 1992–2013. We also match with YTS 14% of the observations in TRI for the period 1997–2018.

Fourth, we match observations in which the geodesic distance between the facility latitude-longitude in the two databases is less than 1km and the facility or parent name—excluding the city name, if part of the name—starts with the same word across the two databases. In this step, we match with NETS 0.3% of the observations in TRI for the period 1992–2013. We also match with YTS 3.7% of the observations in TRI for the period 1997–2018.

Fifth, we match observations in which the geodesic distance between the facility latitude-longitude in the two databases is less than 1km and the facility or parent name are short and very similar (i.e., a fuzzy name match using criteria appropriate for small names) across the two databases. In this step, we match with NETS 0.002% of the observations in TRI for the period 1992–2013. We also match with YTS 0.2% of the observations in TRI for the period 1997–2018.

We subsequently check for false positives, and verify that they are quite rare. At the end of this procedure, we have matched with NETS 49.9% of the observations in TRI for the period 1992–2013, and we have matched with YTS 82.8% of the observations in TRI For the period 1997–2018.

We note that, for some of the matched observations, NETS and YTS are missing information for the number employees and the U.S. dollar value of sales. From NETS, we have information on the number of employees for 235,344 observations (i.e., 44.3% of the observations in TRI for the period 1992–2013) and we have information on the U.S. dollar value of sales for 235,316 observations (i.e., 44.3% of the observations in TRI for the period 1992–2013). From YTS, we have information on the number of employees for 309,540 observations (i.e., 61% of the observations in TRI for the period 1997–2018) and we have information on the U.S. dollar value of sales for 301,224 observations (i.e., 59.4% of the observations in TRI for the period 1997–2018). Overall, we have information on the number of employees from either NETS or YTS for 449,585 observations (i.e., 70.7% of the observations in TRI for the period 1992–2018) and we have information on the U.S. dollar value of sales for 445,276 observations (i.e., 70% of the observations in TRI for the period 1992–2018).

IA.B Industries represented in our sample

Table IA.B1: Facility industries in our sample

This table lists the facility industries (defined at the 3-digit NAICS level) included in our final sample of 104,067 facility-years in the period 1992–2018. The first column shows an industry's 3-digit NAICS code and the second column shows its description. The third column shows the distribution of facility-year observations in our sample across industries. The last column shows the distribution of total toxic emissions in our sample across industries. Industries that represent less than 0.1% of our sample are not shown in the table.

NAICS code	Description	Percent of facility-years	Percent of total pollution
325	Chemical manufacturing	17.13%	21.80%
311	Food manufacturing	10.52%	4.36%
332	Fabricated metal product manufacturing	10.05%	1.49%
336	Transportation equipment manufacturing	8.34%	2.75%
333	Machinery manufacturing	6.64%	0.41%
334	Computer and electronic product manufacturing	5.51%	0.26%
331	Primary metal manufacturing	4.72%	7.94%
327	Nonmetallic mineral product manufacturing	4.62%	1.07%
326	Plastics and rubber products manufacturing	4.42%	1.45%
424	Merchant wholesalers, nondurable goods	4.34%	0.25%
335	Electrical equipment, appliance, and component manufacturing	3.78%	0.38%
322	Paper manufacturing	3.20%	8.74%
221	Utilities	3.15%	24.73%
324	Petroleum and coal products manufacturing	2.70%	2.70%
321	Wood product manufacturing	2.13%	0.49%
339	Miscellaneous manufacturing	2.04%	0.28%
562	Waste management and remediation services	1.47%	4.09%
312	Beverage and tobacco product manufacturing	1.43%	0.39%
337	Furniture and related product manufacturing	0.88%	0.23%
323	Printing and related support activities	0.81%	0.48%
313	Textile mills	0.58%	0.09%
314	Textile product mills	0.36%	0.01%
212	Mining (except Oil and Gas)	0.35%	15.45%
316	Leather and allied product manufacturing	0.19%	0.04%

IA.C Additional results

In Table IA.C1, we use data from the NETS and YTS databases to determine the years in which each facility is operating, and we study the effect of CEO birthplace proximity on whether a facility is operating in a given year. Across all specifications using different sets of fixed effects and controls, we find that proximity to CEO birthplaces has no effect on facility operating status. This suggests that the results documented in the paper capture the overall effect of CEO favoritism on emission changes near their birthplaces and are not affected by differential survivorship bias due to CEOs being less (more) likely to close down (open) a working facility near their birthplace. This is also consistent with our result in Table 13 of the paper that proximity to CEO birthplace has no effect on the scale of the facility.

Along similar lines, in Table IA.C2, we repeat our baseline analysis (Table 4 of the paper) at the firm-city level instead of the facility level. Our results suggest that cities with a higher number of facilities near the CEO's birthplace are more likely to experience emission reductions relative to those with a lower number of such facilities. This provides further support that the effect of CEO birthplace proximity on facility emissions is not counteracted by CEOs being less (more) likely to close down (open) a facility near their birthplace. It also supplements our finding in Section 5.1 of the paper that CEOs do not shift emissions to other facilities.

Finally, in Table IA.C3, we study whether our documented effect is different for firms that have made commitments to reduce their future emissions versus those that have not made such commitments. We obtain data on firm commitments from firm disclosures to the Carbon Disclosure Project (CDP) from 2011 to 2018. Firm commitments feature in about 21% of our facility-year observations. We repeat our baseline analysis (Table 4 of the paper) where we interact the hometown facility dummy with a variable that is equal to 1 for facilities whose parent firms have declared an emission reduction target to the CDP and 0 otherwise (i.e., firms that have declared no CDP commitment or have not provided data to the CDP at all). The coefficient on the hometown facility dummy captures the hometown effect for facilities without firm commitments, and the coefficient on the interaction term captures the differential hometown effect for facilities with commitments. We find that the hometown effect is not statistically different for the facilities with commitments than for those without commitments. We conclude that the CEO behavior we document is not primarily driven by their desire to deliver on certain environmental commitments on behalf of firms.

Table IA.C1: Effect of CEO Birthplace Proximity on Facility Operating Status

This table presents the effect of CEO birthplace proximity on the operating status of a facility. The dependent variable equals 1 if the facility is operating in a given year and 0 otherwise. *Hometown facility* is an indicator that equals 1 if the CEO's birthplace is located within 20 miles from the facility's location. All specifications include firm-by-year fixed effects. The specifications in columns 2 – 5 also include facility fixed effects. The specification in column 3 additionally includes facility-state-by-year fixed effects, and the specification in column 4 additionally includes facility-industry-by-year fixed effects. The specification in column 5 additionally includes facility-level time-varying controls for the lagged log number of employees and the number of toxic chemicals used by the facility. *t*-statistics from standard errors clustered at the parent-firm level are reported. */**/*** indicate significance at the 10%/5%/1% levels.

	Facility Operating Status					
	(1)	(2)	(3)	(4)	(5)	
Hometown facility	0.024	0.003	0.006	-0.002	-0.010	
	1.499	0.114	0.200	-0.092	-0.969	
log(Facility employment)					0.004	***
					3.686	
# of chemicals					0.005	***
					5.186	
Facility FE	No	Yes	Yes	Yes	Yes	
Firm-year FE	Yes	Yes	Yes	Yes	Yes	
Facility State-year FE	No	No	Yes	Yes	Yes	
Facility Industry-year FE	No	No	No	Yes	Yes	
# of Observations	156,878	156,407	155,389	155,145	124,647	
R^2	0.102	0.628	0.635	0.640	0.369	

Table IA.C2: CEO Birthplace Proximity and Toxic Emissions – Firm-City Analysis

In Column 1 (2), the dependent variable is the percentage annual change in pounds of air, water, and ground (just air) toxic emissions released on-site by all facilities operating by the same parent firm in the same city. *# of hometown facilities* is the total number of facilities operating by the same firm in the same city that are within 20 miles of the CEO's birthplace. The analysis includes firm-city, firm-by-year, state-by-year, and firm-industry-by-year fixed effects. It also includes time-varying controls for the lagged log total number of employees and the total number of toxic chemicals at all facilities operating by the same firm in the same city. *t*-statistics from standard errors clustered at the parent-firm level are reported. */**/*** indicate significance at the 10%/5%/1% levels.

	Δ (Total Emissions)	Δ (Air Emissions)
	(1)	(2)
# of hometown facilities	-0.093 ***	-0.086 **
	-2.622	-2.401
log(employment)	-0.006	-0.007
	-1.112	-1.345
# of chemicals	-0.058 ***	-0.058 ***
	-3.888	-4.132
Firm-city FE	Yes	Yes
Firm-year FE	Yes	Yes
State-year FE	Yes	Yes
Firm Industry-year FE	Yes	Yes
# of Observations	62,606	60,983
R^2	0.256	0.259

Table IA.C3: CEO Birthplace Proximity and Toxic Emissions – Firm Commitments

In Column 1 (2), the dependent variable is the percentage annual change in pounds of air, water, and ground (just air) emissions released on-site at the reporting facility. *Hometown facility* is an indicator that equals 1 if the CEO's birthplace is located within 20 miles from the facility's location. *Commitment* is an indicator variable that equals 1 if the facility's parent firm has declared an emission reduction target to the Carbon Disclosure Project (CDP) and 0 otherwise, i.e., if the parent firm has declared no commitment or has not provided data to the CDP at all. All specifications include facility-level fixed effects and facility-level time-varying controls for the lagged log number of employees and the number of toxic chemicals used by the facility. They also include firm-by-year fixed effects, facility-state-by-year fixed effects, and facility-industry-by-year fixed effects. *t*-statistics from standard errors clustered at the parent-firm level are reported. */**/*** indicate significance at the 10%/5%/1% levels.

	Δ (Total Emissions)	Δ (Air Emissions)
	(1)	(2)
Hometown facility	-0.143 ***	-0.141 ***
	-2.995	-2.922
Hometown facility	0.047	-0.050
x Commitment	0.356	-0.324
log(Facility employment)	-0.008	-0.007
	(-1.608)	(-1.435)
# of chemicals	-0.064 ***	-0.064 ***
	(-3.568)	(-3.812)
Facility FE	Yes	Yes
Firm-year FE	Yes	Yes
Facility State-year FE	Yes	Yes
Facility Industry-year FE	Yes	Yes
# of Observations	72,062	70,215
R^2	0.281	0.284