Internet Appendix

"Waiting or Acting: The Effects of Environmental Regulatory Uncertainty on Green Innovation"

Jiawei (Brooke) Wang

Appendix A. Figures and Tables

Figure A1: **The climate policy uncertainty index**

The measurement of climate policy uncertainty involves searching for articles in eight leading U.S. newspapers from January 2000 to March 2021. Initially, articles containing the terms ("uncertainty" or "uncertain") and ("carbon dioxide" or "climate" or "climate risk" or "greenhouse gas emissions" or "greenhouse" or "CO2" or "emissions" or "global warming" or "climate change" or "green energy" or "renewable energy" or "environmental") and ("regulation" or "legislation" or "White House" or "Congress" or "EPA" or "law" or "policy"), including variants such as "uncertainties", "regulatory", "policies", etc. The eight newspapers are the Wall Street Journal, New York Times, USA Today, Chicago Tribune, the Boston Globe, Los Angeles Times, Miami Herald, and Tampa Bay Times. For each newspaper, the number of relevant articles per month is scaled by the total number of articles during the same month. Subsequently, these eight series are standardized with a unit standard deviation and then averaged across newspapers monthly. Finally, the averaged series are normalized to have a mean value of 100 from January 2000 to March 2021. The following figure presents the Climate Policy Uncertainty (CPU) index (the solid orange line) monthly during the sample period from 2000 to 2020. The dashed blue line represents the Economic Policy Uncertainty (EPU) index during the same period.

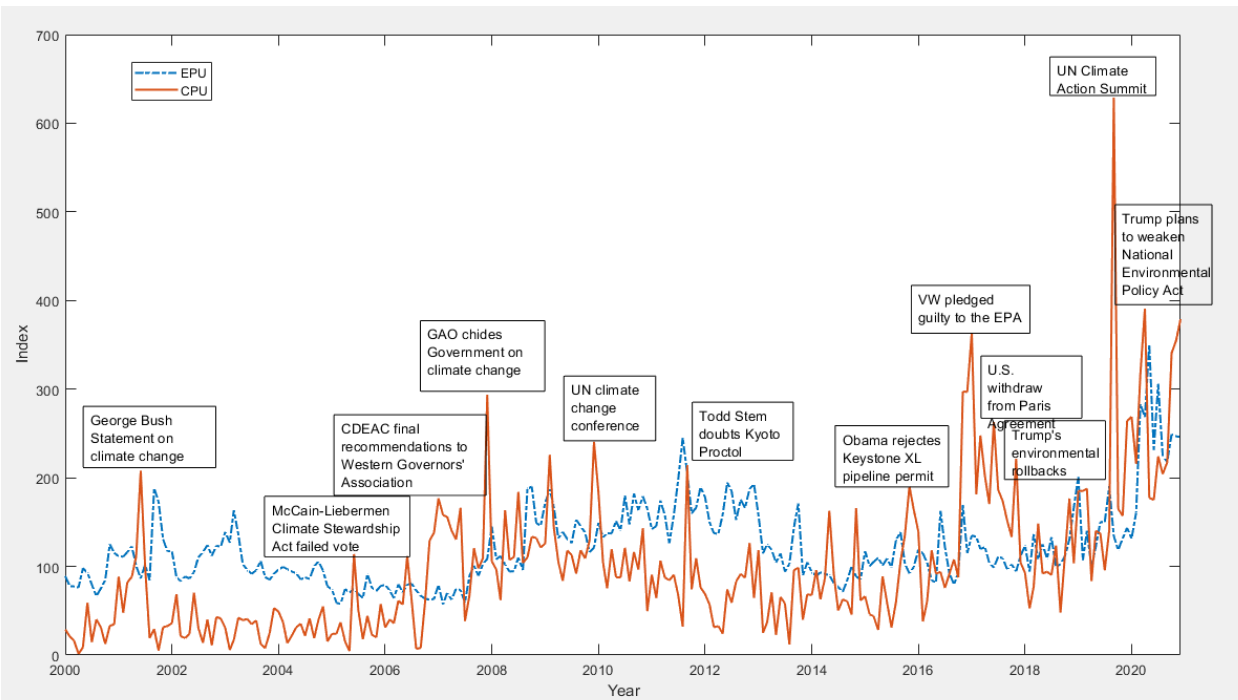

Figure A2: **The trend of emissions**

This figure shows the trend of emissions and the number of polluting plants in the U.S. firms. The blue solid line with circle markers represents the natural logarithm of pollutant emissions (y-axis in the left). The orange dashed line with square markers represents emission intensity (the natural logarithm of total on-site emissions scaled by total assets). The orange dotted line with diamond markers presents the natural logarithm of the number of polluting plants of a firm (ln(plants)).

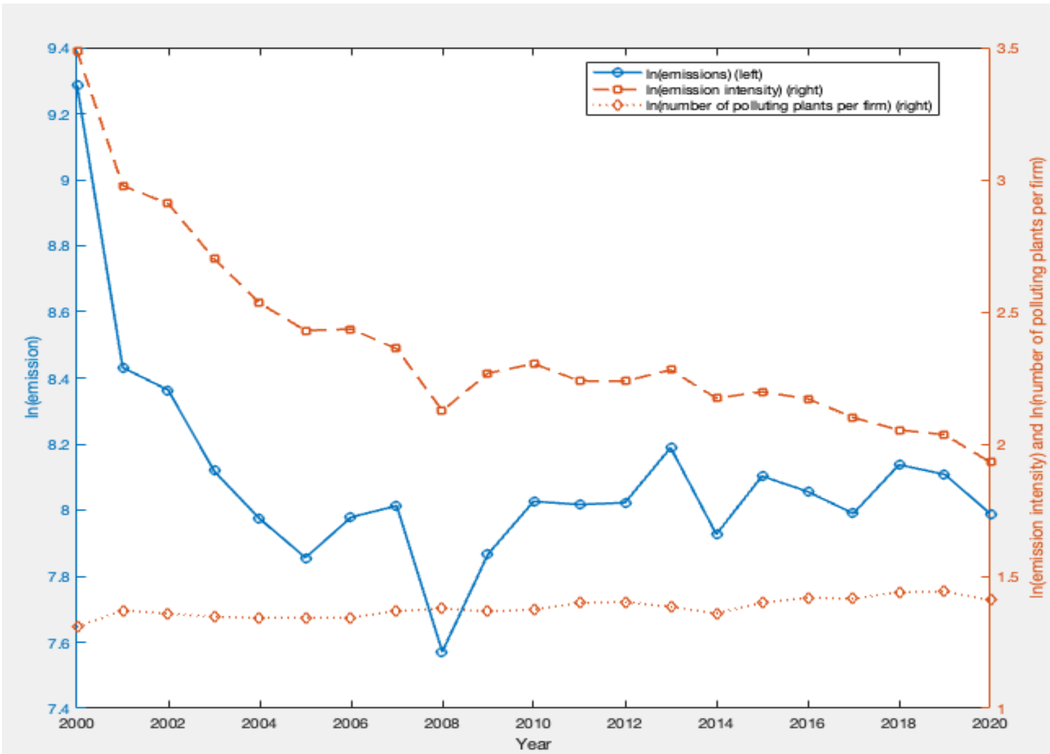

Figure A3: **The textual analysis of ACE rule**

This figure shows regulatory uncertainty in treated (orange, solid with markers) and control (blue) firms based on a textual analysis of regulatory uncertainty around the 2018 ACE rule from 2014 to 2020, with the post-treatment period (2018-2020) shaded. The treated and control firms are defined in subsubsection III.D.2. For each firm-year 10-K, I compute the sentence frequency in *Item 1A. Risk Factors*, which is defined as the fraction of Item 1A sentences that contain at least one uncertainty keyword and at least one regulation keyword. The uncertainty keyword list includes uncertain, uncertainty, uncertainties, cannot predict, unable to predict, not possible to determine, cannot estimate, difficult to predict, currently unknown, we cannot predict, we cannot assure, and subject to. The regulation keyword list includes regulation, regulations, regulatory, ACE, EPA, affordable energy rule, CPP, clean power plan, environmental protection agency, rulemaking, litigation, legal challenge, finalize rule, and proposed rule. All filings are converted to lower case before keyword search to ensure consistency. In the pre-2018 period both series trend upward slowly. After 2018 the treated series rises sharply through 2020, while the control series flattens slightly. The sentence frequency variable is winsorized at the 1% and 99% level.

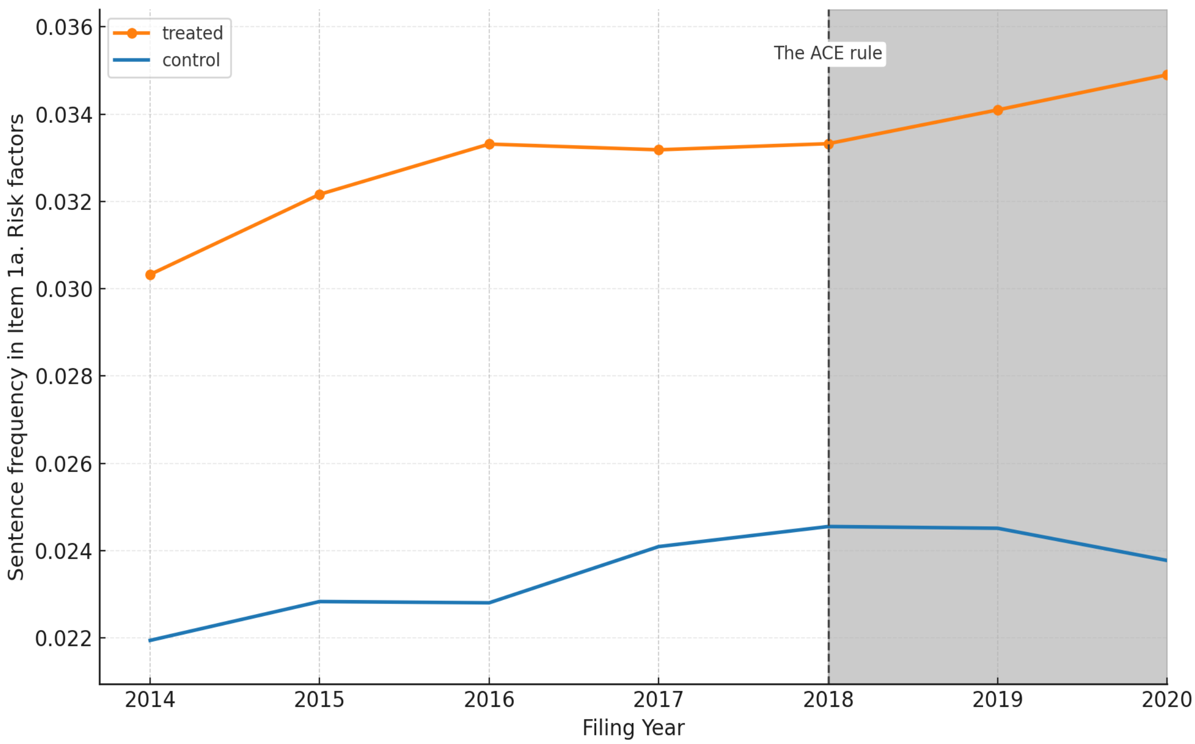

Table A1: Public attention

This table presents the results of the IV 2SLS regressions on green patents with an additional control for public attention. Public attention is measured using the state-level Google Search Volume Index (SVI), available from 2004 onward with 2000–2003 values back-filled using 2004 data. The sample period spans from 2000 to 2020. All continuous variables are winsorized at the 1st and 99th percentiles and standardized. The regressions control for fixed effects as indicated. Standard errors are clustered at the firm and state levels. For brevity, I only report the first-stage results for adjusted green patents in Column (1). T-statistics are in parentheses. Cragg-Donald Wald F-statistics are reported at the bottom of the table. Coefficients that are statistically significant at the 10%, 5%, and 1% levels are denoted by *, **, and ***, respectively.

VARIABLES	(1) $\Delta lnCPU$	(2) $\Delta GreenPat\_Adj_{avg(1,2)}$	(3) $\Delta citation\_Adj_{avg(1,2)}$	(4) $\Delta Value_{avg(1,2)}$
$\Delta\widehat{lnCPU}$		0.031**	-0.063	-0.017
		(2.61)	(-1.40)	(-0.28)
$\Delta Polarization_{t-2}$	0.147***			
	(30.37)			
Public attention	0.163***	-0.006	-0.014	0.000
	(8.30)	(-1.05)	(-1.00)	(0.03)
$\Delta Corresponding$	0.042	0.423***	0.011	0.090**
$OtherPat_{avg(1,2)}$	(0.86)	(4.91)	(0.46)	(2.28)
ΔlnEPU	0.135***	-0.009***	0.041**	-0.013
	(15.10)	(-3.02)	(2.69)	(-0.82)
Size	0.216***	-0.016	0.007	-0.044
	(8.52)	(-1.42)	(0.16)	(-1.32)
Firm age	-0.501***	0.028***	-0.069	-0.006
	(-10.94)	(3.23)	(-1.56)	(-0.15)
MTB	0.094***	0.006	0.021	-0.023**
	(5.99)	(0.78)	(1.49)	(-2.15)
ROA	0.098**	0.003	-0.024	0.006
	(2.36)	(0.36)	(-1.19)	(0.32)
Capex	0.030**	-0.003	0.002	0.008
	(2.18)	(-0.56)	(0.14)	(0.79)
Tangibility	0.007	-0.009	-0.043	-0.019
	(0.33)	(-1.38)	(-1.26)	(-1.06)
Leverage	0.024*	0.001	0.005	-0.005
	(1.70)	(0.20)	(0.44)	(-0.51)
Cash Flow	-0.088**	-0.001	0.015	-0.014
	(-2.33)	(-0.22)	(0.55)	(-0.70)
GDP Growth	-0.046***	0.000	0.007	0.024***
	(-10.44)	(0.13)	(0.75)	(2.96)
Observations	6,639	6,639	6,639	6,639
R-squared	0.152			
Cragg-Donald Wald F statistics	300.28			
Firm FE	Y	Y	Y	Y

Table A2: Governmental environmental spending and R&D tax credit

This table presents the results of the IV 2SLS regression on green patents with additional controls: government environmental spending, tax deductions for environmental investment, and state R&D tax credits. The state R&D tax dummy equals one if a state provides tax credits for R&D expenditures. The REE tax cut dummy captures the effect of the Renewable Energy and Energy Conservation Tax Act of 2007. Federal environmental spending is proxied by the log of the total annual budget of the EPA and DOE (ln(EPA and DOE budget)). The dependent variable and other control variables follow those in Table II. The sample period is between 2000 and 2020. All continuous variables are winsorized at the 1st and 99th percentiles and standardized. The regressions control for fixed effects as shown in the bottom of the table. Standard errors are clustered at the firm and state levels. For brevity, I only report the first-stage results for adjusted green patents in Column (1). T-statistics are in parentheses. Cragg-Donald Wald F-statistics are reported at the bottom of the table. Coefficients that are statistically significant at the 10%, 5%, and 1% levels are denoted by *, **, and ***, respectively.

VARIABLES	(1) $\Delta lnCPU$	(2) $\Delta GreenPat\_Adj_{avg(1,2)}$	(3) $\Delta citation\_Adj_{avg(1,2)}$	(4) $\Delta Value_{avg(1,2)}$
$\Delta\widehat{lnCPU}$		0.079***	-0.170	-0.003
		(3.11)	(-1.61)	(-0.02)
$\Delta Polarization_{t-2}$	0.063***			
	(20.76)			
ln(EPA and DOE budget)	0.342***	-0.027***	0.063	-0.007
	(36.97)	(-3.27)	(1.58)	(-0.16)
State R&D tax dummy	0.017	0.013	-0.045	-0.031*
	(0.48)	(1.05)	(-1.02)	(-1.70)
REE tax cut dummy	-0.911***	0.082***	-0.205*	0.031
	(-109.88)	(2.77)	(-1.99)	(0.25)
$\Delta Corresponding$	0.099***	0.415***	0.011	0.089**
$OtherPat_{avg(1,2)}$	(3.47)	(4.88)	(0.44)	(2.25)
ΔlnEPU	0.108***	-0.014***	0.045**	-0.013
	(21.06)	(-4.18)	(2.31)	(-0.77)
Size	0.277***	-0.034***	0.035	-0.048
	(12.08)	(-2.71)	(0.62)	(-0.90)
Firm age	0.089***	-0.008	-0.018	-0.013
	(3.34)	(-0.92)	(-0.68)	(-0.46)
MTB	0.049***	0.006	0.020	-0.023**
	(4.36)	(0.72)	(1.34)	(-2.17)
ROA	0.056	0.001	-0.020	0.005
	(1.58)	(0.06)	(-0.95)	(0.25)
Capex	-0.008	-0.001	-0.001	0.008
	(-0.99)	(-0.24)	(-0.04)	(0.83)
Tangibility	0.046***	-0.012*	-0.037	-0.019
	(3.28)	(-1.82)	(-1.03)	(-1.00)
Leverage	0.061***	-0.004	0.013	-0.005
	(7.51)	(-0.76)	(0.84)	(-0.39)
Cash Flow	-0.079**	0.004	0.006	-0.013
	(-2.19)	(0.54)	(0.22)	(-0.56)
GDP Growth	-0.180***	0.014**	-0.024	0.029
	(-67.71)	(2.20)	(-1.31)	(1.20)
Observations	6,639	6,639	6,639	6,639
R-squared	0.387			
Cragg-Donald Wald F statistics	72.96			
Firm FE	Y	Y	Y	Y

Table A3: Abatement activities

This table presents the Logit regression results examining the likelihood that a firm undertakes source reduction abatement activities, which are categorized into four types — raw materials, equipment, operating, and process, as detailed in Table A8. The dependent variable is a dummy variable indicating whether a firm engaged in at least one abatement activity in each category. The independent variable of interest is lnCPU_residual, which is obtained by regressing lnCPU on a linear time trend. The sample period spans from 2000 to 2020. All continuous variables are winsorized at the 1st and 99th percentiles and standardized. The regressions control for firm fixed effects. The reported standard errors are Jackknife standard errors. T-statistics are in parentheses. Coefficients that are statistically significant at the 10%, 5%, and 1% levels are denoted by *, **, and ***, respectively.

	(1)	(2)	(3)	(4)
VARIABLES	$Raw\ material_{t+1}$	$Equipment_{t+1}$	$Operating_{t+1}$	$Process_{t+1}$
lnCPU_residual	0.120*	-0.110	-0.112*	0.061
	(1.89)	(-1.51)	(-1.96)	(0.96)
lnEPU_residual	0.080	0.033	0.155**	0.037
	(1.01)	(0.33)	(2.34)	(0.64)
Size	-0.059	-0.381	-0.439	-0.499
	(-0.16)	(-0.64)	(-1.35)	(-1.50)
Firm age	-1.457***	-0.737**	-0.864***	-0.845***
	(-4.70)	(-2.17)	(-2.97)	(-3.26)
MTB	0.041	-0.072	0.058	0.006
	(0.28)	(-0.42)	(0.54)	(0.04)
ROA	-0.311	-0.054	0.055	0.010
	(-1.25)	(-0.13)	(0.24)	(0.05)
Capex	-0.064	0.149	0.141*	0.016
	(-0.52)	(0.99)	(1.79)	(0.20)
R&D	0.114	0.089	-0.150	0.307
	(0.39)	(0.34)	(-0.75)	(1.49)
Tangibility	0.024	-0.021	0.047	-0.352*
	(0.12)	(-0.06)	(0.22)	(-1.88)
Leverage	-0.194	-0.083	0.093	-0.058
	(-1.44)	(-0.53)	(0.76)	(-0.59)
Cash Flow	0.237	0.267	0.084	0.064
	(0.91)	(0.67)	(0.35)	(0.30)
GDP Growth	0.145*	0.139	0.043	0.071
	(1.92)	(1.40)	(0.64)	(1.02)
Observations	3,035	2,330	3,591	3,750
Firm FE	Y	Y	Y	Y

Table A4: Poisson-pseudo maximum likelihood regressions for green patents

This table presents the Poisson-Pseudo Maximum Likelihood (PPML) regression results. The dependent variables are the counts of green patents filed in year (t+1) and (t+2) and their total forward citation counts, respectively. The independent variable of interest is lnCPU_residual, which obtained from the regression of lnCPU on a time trend. The sample period spans from 2000 to 2020. All continuous variables are winsorized at the 1st and 99th percentiles and standardized. The regressions control for fixed effects as indicated. Standard errors are clustered at the firm and state levels. T-statistics are in parentheses. Coefficients that are statistically significant at the 10%, 5%, and 1% levels are denoted by *, **, and ***, respectively.

VARIABLES	(1) $GreenPat_{t+1}$	(2) $GreenPat_{t+2}$	(3) $Citation_{t+1}$	(4) $Citation_{t+2}$
lnCPU_residual	0.042**	0.071***	0.120***	0.140***
	(2.50)	(3.45)	(7.00)	(5.87)
Corresponding OtherPat	0.001	0.001**	0.000***	0.000***
	(1.49)	(2.35)	(6.57)	(8.05)
lnEPU_residual	0.008	0.009	-0.060**	-0.054
	(0.34)	(0.43)	(-2.37)	(-1.37)
Size	-0.030	-0.078	-0.249**	-0.195
	(-0.36)	(-0.95)	(-2.14)	(-1.34)
Firm age	-0.454***	-0.758***	-2.473***	-2.871***
	(-3.46)	(-3.99)	(-6.39)	(-6.42)
MTB	-0.113**	-0.108**	-0.163***	-0.245***
	(-2.40)	(-2.33)	(-2.90)	(-2.80)
ROA	0.011	0.050	-0.120	-0.015
	(0.19)	(0.54)	(-1.10)	(-0.10)
Capex	0.050	0.043	-0.093	-0.154
	(0.87)	(0.93)	(-0.90)	(-1.29)
Tangibility	-0.289***	-0.312**	-0.155	-0.068
	(-2.69)	(-2.54)	(-1.62)	(-0.61)
Leverage	-0.190***	-0.194***	0.082	0.067
	(-3.71)	(-3.08)	(0.62)	(0.51)
Cash Flow	0.008	-0.022	0.207**	0.160
	(0.10)	(-0.19)	(2.17)	(1.22)
GDP Growth	0.064***	0.040*	0.084**	0.089
	(3.16)	(1.79)	(2.26)	(1.22)
Constant	3.123***	3.169***	7.980***	8.142***
	(10.80)	(9.61)	(27.36)	(21.74)
Observations	4,050	3,944	3,668	3,538
Pseudo R-squared	0.800	0.794	0.865	0.863
Firm FE	Y	Y	Y	Y

Table A5: Regressions with firm-year fixed effects

This table presents the OLS and IV 2SLS regression results on green patents with the firm-by-year fixed effects. Column (1) reports OLS results; Columns (2) and (3) present IV 2SLS estimates. The sample period spans from 2000 to 2020. The control variables follow Table II. All continuous variables are winsorized at the 1st and 99th percentiles and standardized. Standard errors are clustered at the firm and state levels. T-statistics are in parentheses. Cragg-Donald Wald F-statistic is reported for the IV regression. Coefficients statistically significant at the 10%, 5%, and 1% levels are denoted by *, **, and ***, respectively.

VARIABLES	(1) $\Delta GreenPat\_Adj_{avg(1,2)}$	(2) $\Delta lnCPU$	(3) $\Delta GreenPat\_Adj_{avg(1,2)}$
ΔlnCPU	0.011***		
	(3.51)		
$\Delta\widehat{lnCPU}$			0.022*
			(1.76)
$\Delta Polarization_{t-2}$		0.140***	
		(21.17)	
$\Delta Corresponding$	0.330***	0.170**	0.328***
$OtherPat_{avg(1,2)}$	(4.53)	(2.68)	(4.55)
ΔlnEPU	-0.001	0.127***	-0.003
	(-0.34)	(15.53)	(-0.96)
Size	-0.018**	0.577***	-0.025*
	(-2.06)	(9.96)	(-1.90)
Firm age	0.015**	-0.250***	0.018*
	(2.03)	(-7.20)	(1.79)
MTB	0.012	0.078***	0.011
	(1.59)	(3.03)	(1.45)
ROA	-0.006	0.142**	-0.008
	(-0.97)	(2.63)	(-1.26)
Capex	-0.001	0.026	-0.002
	(-0.42)	(1.55)	(-0.51)
Tangibility	-0.003	0.016	-0.003
	(-0.52)	(0.50)	(-0.62)
Leverage	-0.006	0.086***	-0.007
	(-1.49)	(5.39)	(-1.63)
Cash Flow	0.006	-0.146***	0.007
	(0.96)	(-3.05)	(1.22)
GDP Growth	-0.001	-0.064***	-0.001
	(-0.28)	(-11.47)	(-0.30)
Constant	-0.000		
	(-0.46)		
Observations	6,091	6,091	6,091
R-squared	0.351	0.182	0.061
Cragg-Donald Wald F-statistics		246.13	
Firm × Year FE	Yes	Yes	Yes

Table A6: Facility emissions

This table presents the OLS results (Columns 1-2) and the IV 2SLS results (Columns 3-4) examining the impact of environmental regulatory uncertainty on the facility-level emissions. Both ln(emission) and lnCPU are used in either first-differenced or residualized format, where residuals are obtained by regressing each variable on a linear time trend. The sample period spans from 2000 to 2020. The sample includes 39,042 facility-year observations and 4,009 unique facilities. All continuous variables are winsorized at the 1st and 99th percentiles and standardized. The regressions control for fixed effects as indicated. Standard errors are clustered at the facility and state levels. T-statistics are in parentheses. Cragg-Donald Wald F-statistic is reported at the bottom of the table. Coefficients statistically significant at the 10%, 5%, and 1% levels are denoted by *, **, and ***, respectively.

VARIABLES	(1) $\Delta ln(emission)_{avg(1,2)}$	(2) ln(emission) $residual_{avg(1,2)}$	(3) $\Delta lnCPU_t$	(4) $\Delta ln(emission)_{avg(1,2)}$
ΔlnCPU	-0.017***			
	(-8.11)			
lnCPU_residual		-0.012***		
		(-4.34)		
$\Delta \widehat{lnCPU}$				-0.038***
				(-3.55)
$\Delta Polarization_{t-2}$			0.127***	
			(29.96)	
ΔlnEPU	0.005***		0.155***	0.009***
	(3.42)		(27.42)	(3.40)
lnEPU_residual		0.012***		
		(3.89)		
Size	-0.001	0.012	0.299***	0.006
	(-0.19)	(0.76)	(13.63)	(0.99)
Firm age	-0.037***	-0.027	-0.259***	-0.043***
	(-4.94)	(-1.44)	(-10.53)	(-4.75)
MTB	0.005**	0.002	0.135***	0.008***
	(2.13)	(0.44)	(7.01)	(2.83)
ROA	0.329***	-0.050	5.071***	0.447***
	(3.21)	(-1.55)	(4.51)	(3.55)
Capex	-0.031	-0.120	3.220***	0.052
	(-0.16)	(-1.24)	(3.83)	(0.25)
R&D	-0.004	-0.007	-0.030	-0.004
	(-0.93)	(-0.78)	(-0.91)	(-1.17)
Tangibility	-0.010***	0.007	-0.021*	-0.010***
	(-3.04)	(0.76)	(-1.68)	(-3.04)
Leverage	-0.004	0.009	0.019**	-0.003
	(-1.54)	(1.48)	(2.11)	(-1.24)
Cash Flow	-0.006*	0.003	-0.067***	-0.008**
	(-1.86)	(0.83)	(-4.51)	(-2.44)
GDP Growth	-0.005***	0.012***	-0.062***	-0.005***
	(-3.40)	(4.34)	(-35.14)	(-3.39)
Constant	-0.015***	0.003***		
	(-7.86)	(6.20)		
Observations	30,429	34,108	30,429	30,429
R-squared	0.092	0.903	0.132	
Cragg-Donald Wald F-statistics			991.85	
Facility FE	Y	Y	Y	Y

Table A7: Environmental regulatory uncertainty and production activities

This table presents the OLS and the IV 2SLS regression results examining the effect of environmental regulatory uncertainty on production activities. The dependent variables are the natural logarithm of number of employees and production waste measured in year (t+1). Based on the ADF tests, ln(employees) is non-stationary (p = 0.466) and is used in first-differenced format. ln(production waste) is stationary (p = 0.011) and used in levels in the regressions. The sample period spans from 2000 to 2020. All continuous variables are winsorized at the 1st and 99th percentiles and standardized. The regressions control for fixed effects as indicated. Standard errors are clustered at the firm and state levels. *T*-statistics are in parentheses. Cragg-Donald Wald F-statistic is reported at the bottom of the table. Coefficients that are statistically significant at the 10%, 5%, and 1% levels are denoted by *, **, and ***, respectively.

VARIABLES	(1) Δln(Employees)	(2) ln(prod_waste)	(3) $\Delta lnCPU$	(4) Δln(Employee)	(5) ln(prod_waste)
ΔlnCPU	0.000	-0.001			
	(0.13)	(-0.25)			
$\Delta\widehat{lnCPU}$				-0.011	0.046
				(-1.45)	(1.46)
$\Delta Polarization_{t-2}$			0.141***		
			(26.87)		
ΔlnEPU	-0.012***	0.001	0.175***	-0.009***	-0.009
	(-6.06)	(0.18)	(26.36)	(-4.01)	(-0.96)
Size	-0.004	0.145***	0.351***	0.000	0.127***
	(-0.66)	(4.41)	(12.22)	(0.08)	(3.80)
Firm age	-0.007	-0.173***	-0.234***	-0.010	-0.162***
	(-1.22)	(-3.11)	(-8.47)	(-1.65)	(-2.76)
MTB	0.022***	-0.008	0.103***	0.024***	-0.013
	(7.07)	(-0.45)	(5.77)	(6.73)	(-0.79)
ROA	0.011**	0.012	0.114***	0.012**	0.007
	(2.28)	(0.57)	(3.09)	(2.37)	(0.32)
Capex	-0.004	-0.011	0.019	-0.003	-0.012
	(-1.33)	(-1.52)	(1.32)	(-1.16)	(-1.61)
R&D	-0.002	-0.025	0.016	-0.001	-0.026
	(-0.44)	(-0.82)	(0.69)	(-0.39)	(-0.83)
Tangibility	-0.004	0.018	0.015	-0.004	0.016
	(-1.00)	(0.81)	(0.71)	(-0.90)	(0.73)
Leverage	-0.012***	0.009	0.051***	-0.011***	0.005
	(-4.62)	(0.57)	(3.16)	(-4.06)	(0.34)
Cash Flow	0.000	-0.009	-0.112***	-0.001	-0.004
	(0.02)	(-0.36)	(-3.20)	(-0.20)	(-0.15)
GDP Growth	-0.003*	0.001	-0.073***	-0.004*	0.001
	(-1.76)	(0.15)	(-19.54)	(-1.78)	(0.19)
Constant	0.011***	0.061***			
	(20.06)	(11.42)			
Observations	6,305	6,418	6,305	6,305	6,418
R-squared	0.178	0.884	0.133		
Cragg-Donald Wald F statistics			259.29		
Firm FE	Y	Y	Y	Y	Y

Table A8: The categories of source reduction abatement activities

Code	Abatement Activities
Raw Materials	
W42	Substitute raw materials.
W49	Other raw materials modifications.
Operating	
W13	Improved maintenance scheduling, recordkeeping, or procedures.
W14	Changed production schedule to minimize equipment and feedstock changeovers.
W15	Introduced in-line product quality monitoring or other process analysis systems.
W19	Other changes in operating practices.
Equipment Modifications	
W59	Modified stripping/cleaning equipment.
W60	Changed to mechanical stripping/cleaning devices (from solvents or others).
W61	Changed to aqueous cleaners (from solvents or other materials).
W63	Modified containment procedures for cleaning units.
W64	Improved draining procedures.
W65	Redesigned parts racks to reduce drag-out.
W66	Modified or installed rinse systems.
W67	Improved rinse equipment design.
W68	Improved rinse equipment operation.
W71	Other cleaning and degreasing modifications.
W72	Modified spray systems or equipment.
W73	Substituted coating materials used.
W74	Improved application techniques.
W75	Changed from spray to other systems.
W78	Other surface preparation and finishing modifications.
Process	
W50	Optimized reaction conditions or otherwise increased efficiency of synthesis.
W51	Instituted recirculation within a process.
W52	Modified equipment, layout, or piping.
W53	Use of a different process catalyst.
W54	Instituted better controls on operating bulk containers to minimize discarding of empty containers.
W55	Changed from small volume containers to bulk containers to minimize discarding of empty containers.
W56	Reduced or eliminated the use of an organic solvent.
W57	Used biotechnology in the manufacturing process.
W58	Other process modifications.

Appendix B. The Empirical Strategy

In this section, I provide the major regression specifications.

The IV 2SLS estimation specification in Table II Panel B that examines the effects of environmental regulatory uncertainty on green innovation in polluting firms is as below.

$$\begin{aligned}\Delta lnCPU_t = {} & \psi \Delta Polarization_{t-2} + \bar{\mu}\bar{X}_{i,t} + \bar{\omega}\bar{Z}_t \\ & + \xi \Delta CorrespondingOtherPat_{i,(avg(t+1,t+2))} + \alpha_i + \epsilon_{i,t}.\end{aligned} \tag{B1}$$

$$\begin{aligned}\Delta GreenPat_{i,(avg(t+1,t+2))} = {} & \beta \Delta \widehat{lnCPU}_t + \bar{\gamma}\bar{X}_{i,t} + \bar{\lambda}\bar{Z}_t \\ & + \theta \Delta CorrespondingOtherPat_{i,(avg(t+1,t+2))} + \alpha_i + \epsilon_{i,t}.\end{aligned} \tag{B2}$$

where $\Delta GreenPat_{i,avg(t+1,t+2)}$ denotes the first-differenced average adjusted number of green patents (citations or economic values) filed by firm i in the subsequent years $(t+1)$ and $(t+2)$. The instrumental variable, $\Delta Polarization_{t-2}$, is the two-year lagged first-differenced political polarization on climate and environmental-related votes in the U.S. House of Representatives. $\Delta CorrespondingOtherPat_{i,(avg(t+1,t+2))}$ represents the first-differenced average adjusted non-green patents (citations or economic values) filed by firm i in the subsequent two years, which proxies for firms' overall innovation capacity. Additional control variables, $\bar{X}_{i,t}$ and $\bar{Z}_t$, are specified in the table legend. α_i denotes firm fixed effect.

The estimation specification in Table II Panel C to examine the interaction effects between environmental regulatory uncertainty and past R&D or green patenting activities on green innovations is as below:

$$\begin{aligned}\Delta GreenPat_{i,(avg(t+1,t+2))} = {} & \beta \Delta lnCPU_t \times Dummy_{i,t} + \xi Dummy_{i,t} \\ & + \theta \Delta CorrespondingOtherPat_{i,(avg(t+1,t+2))} \\ & + \bar{\gamma}\bar{X}_{i,t} + \bar{\lambda}\bar{Z}_t + \alpha_i + \tau_t + \epsilon_{i,t}.\end{aligned} \tag{B3}$$

where $Dummy_{i,t}$ represents the binary variables, past GreenPat high (in Column (1)) or past R&D high (in Column (2)). The detailed variable definitions are reported in Appendix section D. $\bar{X}_{i,t}$ and $\bar{Z}_t$ are the same as in Equation B2 and are specified in the corresponding table legend. α_i and τ_t denote firm fixed effect and year fixed effect, respectively.

Next, I present the estimation specification of Table III Panel A that examines the differential effects of environmental regulatory uncertainty on R&D and Capex below.

$$Y_{i,(avg(t+1,t+2))} = \beta \Delta lnCPU_t + \bar{\gamma}\bar{X}_{i,t} + \bar{\lambda}\bar{Z}_t + \alpha_i + \epsilon_{i,t}, \tag{B4}$$

where $Y_{i,avg(t+1,t+2)}$ denotes either average R&D or Capex in years $(t+1)$ and $(t+2)$ in their levels (Columns (1) to (2)) or first-differenced format (Columns (3) to (4)). Control variables, $\bar{X}_{i,t}$ and $\bar{Z}_t$, are specified in the corresponding table legend. α_i denotes firm fixed effect.

The estimation specification in Panel B of Table III that examines the substitution effects between R&D and Capex is as below:

$$Y_{i,(avg(t+1,t+2))} = \beta \Delta lnCPU_t \times Dummy_{i,t} + \xi Dummy_{i,t} + \bar{\gamma}\bar{X}_{i,t} + \bar{\lambda}\bar{Z}_t + \alpha_i + \tau_t + \epsilon_{i,t}, \qquad \text{(B5)}$$

where $Y_{i,avg(t+1,t+2)}$ represents either average R&D or Capex in years $(t+1)$ and $(t+2)$ in their level format. $Dummy_{i,t}$ represents the binary variables, Capex decrease (Column (1)) or High redeployability (Column (2)). $\bar{X}_{i,t}$ and $\bar{Z}_t$ are the same as in Equation B4 and are specified in the corresponding table legend. α_i and τ_t denote firm fixed effect and year fixed effect, respectively.

Lastly, the estimation specification for the ACE rule analysis in Table IV and Table V is as below.

$$Y_{i,t} = \beta Post - ACE_t \times Treated_i + \bar{\gamma}\bar{X}_{i,t-1} + \bar{\lambda}\bar{Z}_{t-1} + \alpha_i + \tau_t + \epsilon_{i,t}, \qquad \text{(B6)}$$

I also estimate the following dynamic difference-in-differences specification:

$$Y_{i,t} = \sum_{\substack{t=2014 \\ t\neq 2018}}^{2020} \beta_t Year_t \times Treated_i + \bar{\gamma}\bar{X}_{i,t-1} + \bar{\lambda}\bar{Z}_{t-1} + \alpha_i + \tau_t + \epsilon_{i,t}, \qquad \text{(B7)}$$

where $Y_{i,t}$ represents R&D, log(GHG), or log(GHG_intensity). *Post-ACE* equals one in 2019 and 2020, and zero for 2014–2017 (2018 is excluded to avoid potential confounding effects). $Treated$ equals one if a firm had at least one fossil-fuel power plant that was affected by the 2018 Affordable Clean Energy (ACE) Rule. In the dynamic difference-in-differences specification, the 2017 dummy is omitted as the reference year. Control variables, $\bar{X}_{i,t-1}$ and $\bar{Z}_{t-1}$, are specified in the corresponding table legend. α_i and τ_t denote firm fixed effect and year fixed effect, respectively.

Appendix C. The Discussion on the ACE Rule

The following paragraphs report the discussion of the impact of the ACE rule from the Congressional Research Service and the affected electricity firms' 10K reports and other disclosures.

Congressional Research Service Report

EPA's Affordable Clean Energy Rule and Related Issues: Frequently Asked Questions, August 13, 2020

The lack of a numeric performance standard introduces uncertainty about the level of stringency expected from implementation of HRI[1] measures under the ACE rule. EPA reported that "numerous" state and industry commenters requested a presumptive standard or additional clarity about the CO_2 performance standards. . .

EPA points out that its emission projections contain uncertainty. In addition to the factors identified by EPA that provide uncertainty for all the scenarios—electricity demand, natural gas supply and demand, and long-term planning by utilities—the agency describes "considerable uncertainty" that is particular to its ACE rule policy scenario. In its Regulatory Impact Analysis of the final rule, EPA states that "there is inadequate and incomplete information regarding how states might specifically implement this rule, and the estimated range of costs and impacts presented in this chapter is based on the assumptions."

DTE Energy Co.

Quotes from DTE Energy Co. 2019 and 2020 Disclosure Action

In June 2019, EPA finalized the Affordable Clean Energy (ACE) Rule to regulate emissions of carbon dioxide from existing coal-fired power plants. The ACE Rule will impact at least two DTE Energy coal plants that will be operating beyond 2022 when the State of Michigan is required to submit its implementation plan to EPA. DTE has assembled a team to assess the emission reductions and energy efficiency opportunities that may be achieved at the emission units subject to the ACE Rule.

Quotes from DTE Energy Co. 2019 10K

In August 2018, the EPA proposed revised emission guidelines for GHGs from existing electric utility generating units... Comments on the proposed ACE rule were due on October 31, 2018. It is not possible to determine the potential impact of the EPA's proposed ACE rule on existing sources at this time.

Pending or future legislation or other regulatory actions could have a material impact on DTE Electric's operations...

Quotes from DTE Energy Co. 2021 Disclosure Action

In June 2019, EPA finalized the Affordable Clean Energy (ACE) Rule to regulate emissions of carbon dioxide from existing coal-fired power plants. The ACE Rule was vacated and remanded back to EPA by the D.C. District Court in January 2021. Future actions by the EPA to develop a replacement for the ACE rule are uncertain and contribute to the risk of potential emerging regulations.

Ameren Corp.

[1] HRI refers to heat rate improvement.

Quotes from Ameren Corp. 2019 10K Report

In August 2018, the EPA proposed to repeal and replace the Clean Power Plan with a proposed new rule known as the Affordable Clean Energy Rule, which establishes emission guidelines for states to follow in developing plans to limit CO2 emissions from power plants... The EPA is expected to finalize the Affordable Clean Energy rule in the first half of 2019. We cannot predict the outcome of EPA's rulemaking or the outcome of legal challenges related to such rulemaking.

Quotes from Ameren Corp. 2020 10K Report

In July 2019, the EPA issued the Affordable Clean Energy Rule, which establishes emission guidelines for states to follow in developing plans to limit CO2 emissions from coal-fired electric generating units... The plan is expected to include a standard of performance for each affected generating unit. We are evaluating the impact of the adoption and implementation of the Affordable Clean Energy Rule and, along with other stakeholders, will be working with the state of Missouri to develop the compliance plan submitted to the EPA. At this time, we cannot predict the outcome of Missouri's compliance plan development process. As such, the impact on the results of operations, financial position, and liquidity of Ameren and Ameren Missouri is uncertain. We also cannot predict the outcome of any potential legal challenges to the rule.

Southern Co.

Quotes from Southern Co. 2019 10K Report

The EPA has published rules limiting CO2 emissions from new, modified, and reconstructed fossil fuel-fired electric generating units and guidelines for states to develop plans to meet EPA-mandated CO2 emission performance standards for existing units (known as the Clean Power Plan or CPP). On August 31, 2018, the EPA published a proposed rule known as the Affordable Clean Energy (ACE) Rule, which is intended to replace a regulation enacted in 2015 known as the Clean Power Plan (CPP), that would limit CO2 emissions from existing fossil fuel-fired electric generating units. The CPP has been stayed by the U.S. Supreme Court since 2016. The ACE Rule would require states to develop GHG unit-specific emission rate standards based on heat-rate efficiency improvements for existing fossil fuel-fired steam units... As of January 1, 2019, the Southern Company system has ownership interests in 40 fossil fuel-fired steam units to which the proposed ACE Rule is applicable. The ultimate impact of this rule to the Southern Company system is currently unknown and will depend on changes between the proposal and the final rule, subsequent state plan developments and requirements, and any associated legal challenges.

Appendix D. Variable definitions

Variables	Definitions	Sources
lnCPU	The natural logarithm of the climate policy uncertainty index.	Gavriilidis (2021)
Polarization	The average absolute difference between the percentage of Yea and Nay votes on bills of environment and climate issues in the U.S. House of Representatives.	The League of Conservation Voters (LCV)
$lnCPU_{residualized}$	The natural logarithm of the residual from the regression of the CPU index on a time trend.	Gavriilidis (2021)
$GreenPat\_adj$	The number of green patents filed by a firm in a year that are successfully granted scaled by the total green patents that applied in the same year.	BFMP & KPSS
$GreenPatCitations\_adj$	The number of citations received by green patents filed by a firm scaled by the average number of citations for all green patents filed in the same year.	BFMP & KPSS WRDS Patents
$GreenPat\_adj_{avg(1,2)}$	The two-year forward average of adjusted green patent counts.	BFMP & KPSS
$GreenPat\ citation\_adj_{avg(1,2)}$	The two-year forward average of adjusted citations of green patents.	BFMP & KPSS WRDS Patents
$GreenPat\ value_{avg(1,2)}$	The average economic value of two-year forward green patents.	KPSS
$OtherPat\_adj_{avg(1,2)}$	The two-year forward average of adjusted non-green patent counts.	BFMP &KPSS
$OtherPat\ citation\_adj_{avg(1,2)}$	The two-year forward average of adjusted citations of non-green patents.	BFMP &KPSS WRDS Patents
$OtherPat\ value_{avg(1,2)}$	The average economic value of two-year forward non-green patents.	KPSS
Post-ACE	A dummy variable, which takes the value one in 2019 and 2020, and zero from 2014 to 2017.	
Treated	A dummy variable, which takes the value one for fossil fuel-fired power plants (facility-level) or firms with fossil fuel-fired power plants (firm level) that are affected by the ACE rule and zero for other facilities (firms) reporting to the EPA GHG program.	EPA GHG Reporting Program
Sentence frequency	The fraction of sentences that contain at least one "uncertainty" keywords and at least one "regulation" keywords in Item 1A. Risk Factors of 10-K filings.	SEC Edgar

Continued on the next page

Variables	Definition	
$ln(emission)_{avg(1,2)}$	The average of the natural logarithm of the two-year forward total pounds of toxic chemical release, aggregated at the firm or facility level.	EPA Toxic Release Inventory Program
$ln(emission\_intensity)_{avg(1,2)}$	The average of the natural logarithm of two-year forward emission intensity. Emission intensity is measured as toxic chemical release scaled by total assets.	As above
ln(GHG)	The natural logarithm of GHG emissions at the firm level.	EPA GHGRP
ln(GHG_intensity)	The natural logarithm of the GHG emissions scaled by total assets.	As above
Raw material	A dummy variable that equals one if there is at least one raw material-related abatement activity in a firm-year and zero, otherwise.	EPA Pollution Prevention
Equipment	A dummy variable that equals one if there is at least one equipment-related abatement activity in a firm-year and zero, otherwise.	As above
Operating	A dummy variable that equals one if there is at least one operating-related abatement activity in a firm-year and zero, otherwise.	As above
Process	A dummy variable that equals one if there is at least one process-related abatement activity in a firm-year and zero, otherwise.	As above
Size	The natural logarithm of a firm's market value of equity.	Compustat
Firm age	The natural logarithm of a firm's age.	As above
MTB	Market value of total assets (AT+PRCC_F*CSHO-TXDB-CEQ) to book value of assets.	As above
ROA	Net income scaled by total assets.	As above
Capex	Capital expenditures (CAPEX) scaled by total assets.	As above
R&D	R&D expenditures scaled by total assets.	As above
Tangibility	Property, plant and equipment (PPENT) scaled by total assets.	As above
Leverage	Total debt scaled by total assets.	As above
Cash flow	(IB+DP) scaled by total assets.	As above
Past R&D high	A dummy variable equals one if a firm's total R&D expenditures in the past five years (t-5, t-1) are above the sample median and zero, otherwise.	As above
Past GreenPat high	A dummy variable equals one if a firm's total green patents ("fixed-effects" adjusted) in the past five years (t-5, t-1) are above the sample median and zero, otherwise.	BFMP&KPSS
Capex decrease	A dummy variable equals one if capital expenditure reduced in a year and	Compustat

Continued on the next page

Variables	Definition	
	zero, otherwise.	
High redeployability	A dummy variable equals one if a firm's asset redeployability is above the sample median and zero, otherwise.	Kim and Kung (2017)
ln(Employees)	The natural logarithm of the number of employees in a firm.	Compustat
ln(prod_waste)	The natural logarithm of the total production waste of a firm in a year.	EPA
Regulation sentiment	The natural logarithm of one plus 1000 multiply regulation sentiment measures from Sautner, Van Lent, Vilkov, and Zhang (2023).	
Penalty growth	The annual growth of total penalties of all firms in the sample.	EPA ECHO Program
Compliance growth	The annual growth of total compliance cost in all firms in the sample.	As above
Enforce state	The annual growth rate of the number of enforcements of polluting firms in a state.	As above
Penalty state	The annual growth rate of the penalty of the sample polluting firms in a state.	As above
Public attention	The Google Search Volume Index (SVI) of "climate change" search frequency divided by the total search frequency in a state-year. The data is available from 2004. The data between 2000 and 2003 were back-filled by the 2004 data.	Google Trends
lnEPU	The natural logarithm of the averaged monthly economic policy uncertainty index in a year.	Bloom, Baker, and Davis
GDP Growth	The growth rate of the U.S. GDP.	FRED St. Louis
CCAP	A binary variable, which takes the value one after a state has taken climate change adoption plans.	Georgetown Climate Center
RGGI	A binary variable, which takes the value one after a state has taken CO_2 Budget Trading programs.	Center for Climate Energy Solutions
ln(EPA and DOE budget)	The natural log of the sum of the EPA and DOE's annual budget.	U.S. Treasury
State R&D tax dummy	A dummy variable equals one after a state enacting Research Development Expenses and tax credit and zero, otherwise.	Wikipedia
REE tax cut dummy	A dummy variable of renewable energy and energy conservation tax act of 2007. It equals one for years after 2007 and zero from 2000 to 2007.	Congress.gov

Appendix E. Construction of the EPA variables

I utilize the granular pollution emission data from the EPA Toxic Release Inventory (TRI) database. I match the parent companies of the facilities in the EPA databases with the firms in Compustat. This novel matching of datasets provides detailed firm and facility-level information, such as pollutant emissions, abatement activities, production, the EPA enforcement, penalties, and compliance costs.

EPA Data

To obtain toxic emissions and related information, I utilize four databases from the EPA, which I merge using their administrative identifiers. The first data source for emissions is the Toxic Release Inventory (TRI). This database is constructed following the Section 313 reporting requirements of the 1986 Emergency Planning and Community Right-to-Know Act (EPCRA). It provides self-reported data on toxic emissions and tracks the releases of 770 chemicals that pose threats to human well-being and the environment. Additionally, it includes physical locations of plants, names of parent companies, and identifiers assigned by the EPA.

The TRI data are self-reported and concerns about their accuracy may arise. However, existing research suggests that firms do not commonly misreport emissions. First, the EPA periodically conducts audits to verify data integrity. Misreporting could result in civil and criminal penalties. According to EPA estimates, reported amounts and auditor estimates were within 3% of each other for most industries.

The second database is drawn from the EPA's Pollution Prevention (P2) database, which provides information on firms' production and abatement activities. The third database is the Air Stationary Source Data, from which I obtain data on greenhouse gas emissions under the EPA Greenhouse Gas Reporting Program (GHGRP). Emissions are presented as facility-level aggregates and organized by pollutant and EPA program. The fourth and final database is the EPA's Enforcement and Compliance History Online (ECHO) system. The ECHO system incorporates Federal Enforcement and Compliance (FE&C) from the Integrated Compliance Information System (ICIS). It has tracked civil, judicial, and administrative EPA enforcement cases since 1980. Enforcement cases may result in monetary penalties. Firms could incur compliance costs to resolve environmental violations.[2]

There is no existing linkage between the administrative identifiers of the EPA databases and Compustat. Therefore, I use an algorithm to connect the facilities in the EPA databases to the firms in Compustat. During the matching process, I remove common suffixes, e.g., "Corp.", "Inc.", "CO.", etc. To ensure the matching quality, I manually verify each matching pair and manually matched the observations that remain unmatched following the initial algorithm.

Environmental Variables Construction

I use the total weight of the on-site toxic chemical releases, including air emissions, water discharges, underground injection, etc., from the TRI database to measure emissions. I aggregate the data at both the facility and firm levels. Following the literature, I use the natural logarithm of these variables (Akey and Appel, 2021; Bellon, 2021; Chu and Zhao, 2019; Akey and Appel, 2019).

[2]The above-mentioned major databases from the EPA are widely used in the existing studies, e.g., Bellon (2025, 2021); Chu, Guo, Zhao, and Zheng (2021); Campa (2018); Cordis, Hsu, and Zhang (2022); Levine, Lin, Wang, and Xie (2019); Hamilton (1995).

I normalize toxic emissions by total assets as robustness checks. Figure A2 illustrates the trend of pollutant emissions over the sample period.

To measure abatement activities, I utilize environment variables from the EPA's Pollution Prevention (P2) database to access facilities' abatement activities, which reduce toxic releases at the source. Under the Pollution Prevention Act of 1990 (PPA), the EPA collects information to monitor industry progress in waste reduction and adoption of safer waste management alternatives. I construct source reduction activity dummies using the data from the P2 database. I categorize the source reduction abatement activities into four categories: raw material, equipment, operating, and process, as detailed in Table A8.