The Causal Effects of Long-Term Exposure to Air Pollution: Evidence from Socialist East Germany

MORITZ LUBCZYK
ROCKWOOL Foundation Berlin

Maria Waldinger
ifo München, CESifo

Stockholm School of Economics

April 19, 2024

Why do we care about air pollution?

Pollution + Add to myFT

Almost all the world's population breathes harmful air, says WHO

New database reveals 99% of people on the planet are exposed to air pollution above safe limits

– *Financial Times*, 04 April 2022

Air pollution

Air pollution linked to almost a million stillbirths a year

– *The Guardian*, 29 November 2022

Why do we care about air pollution?

Pollution + Add to myFT

Almost all the world's population breathes harmful air, says WHO

New database reveals 99% of people on the planet are exposed to air pollution above safe limits

– *Financial Times*, 04 April 2022

Air pollution

Air pollution linked to almost a million stillbirths a year

– *The Guardian*, 29 November 2022

But we **know very little** about **long-term exposure** and **outcomes**.

This paper

Variation from natural experiment in Socialist East Germany: rapid substitution of oil with lignite in the 1980s, exogenously increasing air pollution

- East German authoritarian dictatorship provides novel setting for identifying long-term causal effects
  - **Address identification concerns**: spatial sorting, mitigation strategies, latent selection.
  - **Large historic shock**: spatial variation in shock exposure, individuals forcibly exposed for eight years.
  - **Individual-level data**: follow employment biographies in social security data.

Findings:

Lower wages, less employment, earlier retirement for affected individuals up to four decades after shock. Largest effects at both ends of the age-at-exposure distribution.

related literature

This paper

Variation from natural experiment in Socialist East Germany: rapid substitution of oil with lignite in the 1980s, exogenously increasing air pollution

- East German authoritarian dictatorship provides novel setting for identifying long-term causal effects
  - **Address identification concerns**: spatial sorting, mitigation strategies, latent selection.
  - **Large historic shock**: spatial variation in shock exposure, individuals forcibly exposed for eight years.
  - **Individual-level data**: follow employment biographies in social security data.

Findings:

Lower wages, less employment, earlier retirement for affected individuals up to four decades after shock. Largest effects at both ends of the age-at-exposure distribution.

related literature

This paper

Variation from natural experiment in Socialist East Germany: rapid substitution of oil with lignite in the 1980s, exogenously increasing air pollution

- East German authoritarian dictatorship provides novel setting for identifying long-term causal effects
  - **Address identification concerns**: spatial sorting, mitigation strategies, latent selection.
  - **Large historic shock**: spatial variation in shock exposure, individuals forcibly exposed for eight years.
  - **Individual-level data**: follow employment biographies in social security data.

Findings:

Lower wages, less employment, earlier retirement for affected individuals up to four decades after shock. Largest effects at both ends of the age-at-exposure distribution.

▸ related literature

This paper

Variation from natural experiment in Socialist East Germany: rapid substitution of oil with lignite in the 1980s, exogenously increasing air pollution

- East German authoritarian dictatorship provides novel setting for identifying long-term causal effects
  - **Address identification concerns**: spatial sorting, mitigation strategies, latent selection.
  - **Large historic shock**: spatial variation in shock exposure, individuals forcibly exposed for eight years.
  - **Individual-level data**: follow employment biographies in social security data.

Findings:

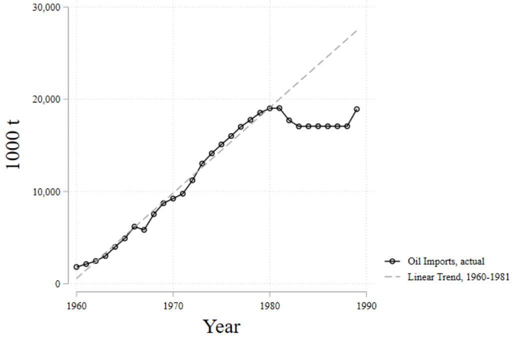
Lower wages, less employment, earlier retirement for affected individuals up to four decades after shock. Largest effects at both ends of the age-at-exposure distribution.

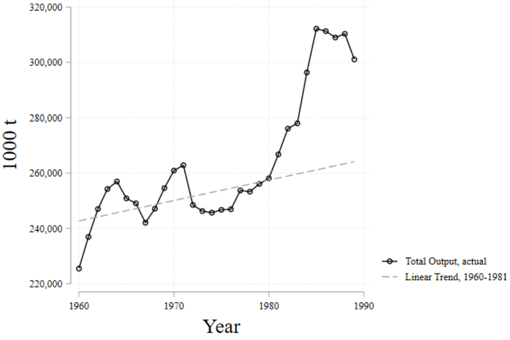
related literature

Socialism and environmental policy

- Autocratic environmental policy
  - Command economy
  - No freedom of movement, limited occupational choice
  - No pollution pricing
- East Germany and the environment
  - Environmental policy as a path to diplomatic legitimacy (Ohlenforst, 2019)
  - Early development of measuring stations, readings declared state secrets after 1982
  - Trade advantages within the Eastern Bloc: moving-average pricing, direct offset between exports and imports, subsidies

Identification from overnight trade shock: Lignite coal as oil substitute





policy context

Lignite mining in Socialist East Germany

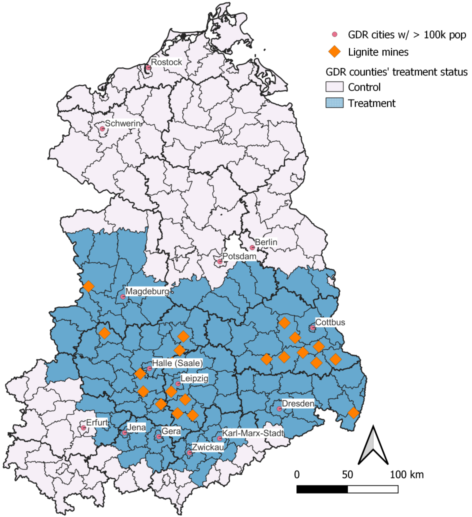
Open pit lignite mine in Welzow-Senftenberg.

Lignite is highly polluting and difficult to transport

- Soft sedimentary rock formed by naturally compressed peat
- "Lowest rank of coal" (EIA, 2021)
- Low net-carbon value, highly polluting
- Inefficient to transport and usually not traded
  - Low energy density
  - Low structural integrity (crumbling)
  - High inherent moisture content (up to 45 %)
  - Rapidly degrades when exposed to air ("slackening")

additional detail

Spatial variation in treatment: Base map



▸ accounting for wind directions ▸ power plants map ▸ falsification tests

Validation approach: Spatial variation

Natural experiment allows identification of *long-term causal effects* because individuals were not able to adjust to exposure during East German dictatorship.

Validation exercises:

1. Shock strongly affected local levels of air pollution ▸ pollution
2. Shock strongly affected individual (health) outcomes: infant mortality and birth weights ▸ health
3. Shock did not significantly affect net migration across counties in East Germany ▸ migration

▸ empirical approach: validation exercises

Lignite power plant in Espenhain.

Validation approach: Spatial variation

Natural experiment allows identification of *long-term causal effects* because individuals were not able to adjust to exposure during East German dictatorship.

Validation exercises:

1. Shock strongly affected local levels of air pollution ▸ pollution
2. Shock strongly affected individual (health) outcomes: infant mortality and birth weights ▸ health
3. Shock did not significantly affect net migration across counties in East Germany ▸ migration

▸ empirical approach: validation exercises

Lignite power plant in Espenhain.

Validation approach: Spatial variation

Natural experiment allows identification of *long-term causal effects* because individuals were not able to adjust to exposure during East German dictatorship.

Validation exercises:

1. Shock strongly affected local levels of air pollution ▸ pollution
2. Shock strongly affected individual (health) outcomes: infant mortality and birth weights ▸ health
3. Shock did not significantly affect net migration across counties in East Germany ▸ migration

▸ empirical approach: validation exercises

Lignite power plant in Espenhain.

Validation approach: Spatial variation

Natural experiment allows identification of *long-term causal effects* because individuals were not able to adjust to exposure during East German dictatorship.

Validation exercises:

1. Shock strongly affected local levels of air pollution ▸ pollution
2. Shock strongly affected individual (health) outcomes: infant mortality and birth weights ▸ health
3. Shock did not significantly affect net migration across counties in East Germany ▸ migration

▸ empirical approach: validation exercises

Lignite power plant in Espenhain.

Validation approach: Spatial variation

Natural experiment allows identification of *long-term causal effects* because individuals were not able to adjust to exposure during East German dictatorship.

Validation exercises:

1. Shock strongly affected local levels of air pollution ▸ pollution
2. Shock strongly affected individual (health) outcomes: infant mortality and birth weights ▸ health
3. Shock did not significantly affect net migration across counties in East Germany ▸ migration

▸ empirical approach: validation exercises

Lignite power plant in Espenhain.

Mechanism: Air pollution and long-term effects

Different health effects of airborne pollution (Aguilar-Gomez et al., 2022):

- Lung and heart functionality
  ⇒ Inflammation of respiratory and cardiovascular systems.
- Brain functionality
  ⇒ Neuro-inflammation and oxidative stress.
- Epigenetic effects
  ⇒ Latent effects on gene expression; long-term cognitive impairments.

Jointly, these effects may impact individuals' labor market performance.

VEB Gaskombinat "Schwarze Pumpe".

Mechanism: Air pollution and long-term effects

Different health effects of airborne pollution (Aguilar-Gomez et al., 2022):

- Lung and heart functionality
  ⇒ Inflammation of respiratory and cardiovascular systems.
- Brain functionality
  ⇒ Neuro-inflammation and oxidative stress.
- Epigenetic effects
  ⇒ Latent effects on gene expression; long-term cognitive impairments.

Jointly, these effects may impact individuals' labor market performance.

VEB Gaskombinat "Schwarze Pumpe".

Mechanism: Air pollution and long-term effects

Different health effects of airborne pollution (Aguilar-Gomez et al., 2022):

- Lung and heart functionality
  $\Rightarrow$ Inflammation of respiratory and cardiovascular systems.
- Brain functionality
  $\Rightarrow$ Neuro-inflammation and oxidative stress.
- Epigenetic effects
  $\Rightarrow$ Latent effects on gene expression; long-term cognitive impairments.

Jointly, these effects may impact individuals' labor market performance.

VEB Gaskombinat "Schwarze Pumpe".

Mechanism: Air pollution and long-term effects

Different health effects of airborne pollution (Aguilar-Gomez et al., 2022):

- Lung and heart functionality
  $\Rightarrow$ Inflammation of respiratory and cardiovascular systems.
- Brain functionality
  $\Rightarrow$ Neuro-inflammation and oxidative stress.
- Epigenetic effects
  $\Rightarrow$ Latent effects on gene expression; long-term cognitive impairments.

Jointly, these effects may impact individuals' labor market performance.

VEB Gaskombinat "Schwarze Pumpe".

Mechanism: Air pollution and long-term effects

Different health effects of airborne pollution (Aguilar-Gomez et al., 2022):

- Lung and heart functionality
  $\Rightarrow$ Inflammation of respiratory and cardiovascular systems.
- Brain functionality
  $\Rightarrow$ Neuro-inflammation and oxidative stress.
- Epigenetic effects
  $\Rightarrow$ Latent effects on gene expression; long-term cognitive impairments.

Jointly, these effects may impact individuals' labor market performance.

VEB Gaskombinat "Schwarze Pumpe".

Long-term effects: Measurement

- **What are the long-term causal effects of air pollution exposure on individual labor market outcomes?**
- Use social security data on universe of workers in East Germany from 1991/92 onwards (6.2m individuals)
  - Assume that individuals' first recorded location in social security data corresponds to their location under socialism
  - For each individual: measure baseline characteristics when they first appear in social security data in 91/92; follow their labor market outcomes until 2020.
    → how many years do they spend in employment/unemployment?
    → what is their average daily wage?
    → at what age do they retire?

Long-term effects: Measurement

- **What are the long-term causal effects of air pollution exposure on individual labor market outcomes?**
- Use social security data on universe of workers in East Germany from 1991/92 onwards (6.2m individuals)
  - Assume that individuals' first recorded location in social security data corresponds to their location under socialism
  - For each individual: measure baseline characteristics when they first appear in social security data in 91/92; follow their labor market outcomes until 2020.
    $\rightarrow$ how many years do they spend in employment/unemployment?
    $\rightarrow$ what is their average daily wage?
    $\rightarrow$ at what age do they retire?

Long-term effects: Measurement

- **What are the long-term causal effects of air pollution exposure on individual labor market outcomes?**
- Use social security data on universe of workers in East Germany from 1991/92 onwards (6.2m individuals)
  - Assume that individuals' first recorded location in social security data corresponds to their location under socialism
  - For each individual: measure baseline characteristics when they first appear in social security data in 91/92; follow their labor market outcomes until 2020.
    → how many years do they spend in employment/unemployment?
    → what is their average daily wage?
    → at what age do they retire?

OLS results and identification concerns

Y =	Employment	Unemployment	Retirement Age	Daily Wages
D_r	-0.241***	0.178***	-0.046**	0.111
	(0.0679)	(0.053)	(0.020)	(0.488)
Unit	Years	Years	Years	EUR
Mean Y	15.849	3.865	62.272	63.844
N	6,108,781	6,108,781	6,108,781	6,108,781
R^2	0.440	0.133	0.063	0.350

OLS estimates including year of birth FE, level of education FE, sex FE, occupation FE, NC3 FE.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

- Freedom of movement begins after German reunification
- Reunification shakes up local labor markets
  → economic transformation
  → place-based subsidies
- Solution: isolate effect of historic shock

OLS results and identification concerns

Y =	Employment	Unemployment	Retirement Age	Daily Wages
D_r	-0.241***	0.178***	-0.046**	0.111
	(0.0679)	(0.053)	(0.020)	(0.488)
Unit	Years	Years	Years	EUR
Mean Y	15.849	3.865	62.272	63.844
N	6,108,781	6,108,781	6,108,781	6,108,781
R^2	0.440	0.133	0.063	0.350

OLS estimates including year of birth FE, level of education FE, sex FE, occupation FE, NC3 FE.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

- Freedom of movement begins after German reunification
- Reunification shakes up local labor markets
  → economic transformation
  → place-based subsidies
- Solution: isolate effect of historic shock

OLS results and identification concerns

Y =	Employment	Unemployment	Retirement Age	Daily Wages
D_r	-0.241***	0.178***	-0.046**	0.111
	(0.0679)	(0.053)	(0.020)	(0.488)
Unit	Years	Years	Years	EUR
Mean Y	15.849	3.865	62.272	63.844
N	6,108,781	6,108,781	6,108,781	6,108,781
R^2	0.440	0.133	0.063	0.350

OLS estimates including year of birth FE, level of education FE, sex FE, occupation FE, NC3 FE.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

- Freedom of movement begins after German reunification
- Reunification shakes up local labor markets
  → economic transformation
  → place-based subsidies
- Solution: isolate effect of historic shock

OLS results and identification concerns

Y =	Employment	Unemployment	Retirement Age	Daily Wages
D_r	-0.241***	0.178***	-0.046**	0.111
	(0.0679)	(0.053)	(0.020)	(0.488)
Unit	Years	Years	Years	EUR
Mean Y	15.849	3.865	62.272	63.844
N	6,108,781	6,108,781	6,108,781	6,108,781
R^2	0.440	0.133	0.063	0.350

OLS estimates including year of birth FE, level of education FE, sex FE, occupation FE, NC3 FE.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

- Freedom of movement begins after German reunification
- Reunification shakes up local labor markets
  → economic transformation
  → place-based subsidies
- Solution: isolate effect of historic shock

Inverse movers design – isolating variation from lignite exposure

How can we identify the effect of increased air pollution exposure in light of post-reunification sorting?

Idea: Look at individuals who move *right* after reunification.

1. Pairs of individuals: Both move to the same destination C, but one of them moves from a treatment region A and one from a control region B.
2. Condition on observables: age, education, sex, occupation, industry.
3. Individuals who move exactly once during their post-reunification life.
4. Single move happens right after reunification, in our data: 91/92 $\rightarrow$ 92/93

Inverse movers design – isolating variation from lignite exposure

How can we identify the effect of increased air pollution exposure in light of post-reunification sorting?

Idea: Look at individuals who move *right* after reunification.

1. Pairs of individuals: Both move to the same destination C, but one of them moves from a treatment region A and one from a control region B.
2. Condition on observables: age, education, sex, occupation, industry.
3. Individuals who move exactly once during their post-reunification life.
4. Single move happens right after reunification, in our data: 91/92 $\rightarrow$ 92/93

Inverse movers design – isolating variation from lignite exposure

How can we identify the effect of increased air pollution exposure in light of post-reunification sorting?

Idea: Look at individuals who move *right* after reunification.

1. Pairs of individuals: Both move to the same destination C, but one of them moves from a treatment region A and one from a control region B.
2. Condition on observables: age, education, sex, occupation, industry.
3. Individuals who move exactly once during their post-reunification life.
4. Single move happens right after reunification, in our data: 91/92 $\rightarrow$ 92/93

Inverse movers design – isolating variation from lignite exposure

How can we identify the effect of increased air pollution exposure in light of post-reunification sorting?

Idea: Look at individuals who move *right* after reunification.

1. Pairs of individuals: Both move to the same destination C, but one of them moves from a treatment region A and one from a control region B.
2. Condition on observables: age, education, sex, occupation, industry.
3. Individuals who move exactly once during their post-reunification life.
4. Single move happens right after reunification, in our data: 91/92 $\rightarrow$ 92/93

Inverse movers design – isolating variation from lignite exposure

How can we identify the effect of increased air pollution exposure in light of post-reunification sorting?

Idea: Look at individuals who move *right* after reunification.

1. Pairs of individuals: Both move to the same destination C, but one of them moves from a treatment region A and one from a control region B.
2. Condition on observables: age, education, sex, occupation, industry.
3. Individuals who move exactly once during their post-reunification life.
4. Single move happens right after reunification, in our data: 91/92 $\rightarrow$ 92/93

Inverse movers design – isolating variation from lignite exposure

How can we identify the effect of increased air pollution exposure in light of post-reunification sorting?

Idea: Look at individuals who move *right* after reunification.

1. Pairs of individuals: Both move to the same destination C, but one of them moves from a treatment region A and one from a control region B.
2. Condition on observables: age, education, sex, occupation, industry.
3. Individuals who move exactly once during their post-reunification life.
4. Single move happens right after reunification, in our data: $91/92 \rightarrow 92/93$

Inverse movers design – isolating variation from lignite exposure

How can we identify the effect of increased air pollution exposure in light of post-reunification sorting?

Idea: Look at individuals who move *right* after reunification.

1. Pairs of individuals: Both move to the same destination C, but one of them moves from a treatment region A and one from a control region B.
2. Condition on observables: age, education, sex, occupation, industry.
3. Individuals who move exactly once during their post-reunification life.
4. Single move happens right after reunification, in our data: 91/92 $\rightarrow$ 92/93

$$Y_i = \alpha_i + \overbrace{\beta \times D_r}^{treatment} + \underbrace{\mathbf{Xb}}_{\text{baseline characteristics}} + \underbrace{destinationFE}_{\text{post-reunification effects}} + \underbrace{\epsilon_{it}}_{\text{unobservables}}$$

Inverse movers design – interpretation

$$Y_i = \alpha_i + \overbrace{\beta \times D_r}^{treatment} + \mathbf{Xb} + destinationFE + \epsilon_{it}$$

- $\mathbf{Xb}$: baseline characteristics
- $destinationFE$: post-reunification effects
- ϵ_{it}: unobservables

- Analogous to movers designs for place-based effects in health economics (Baum et al., 2020; Deryugina and Molitor, 2020, 2021; Finkelstein et al., 2021)
- Addresses (joint) selection into migration **and** selection into destination concerns
- Key identifying assumption: unobservable differences between origin regions are not correlated with later-life labor market outcomes
- Interpretation β: Effect on individuals moving to destination C from treatment region A relative to individuals moving to destination C from control region B, conditional on observables

Results: Long-term labor market outcomes

Y =	Years of Employment	Years of Unemployment	Retirement Age	Daily Wages
D_r	-0.372***	0.088	-0.163**	-2.054**
	(0.129)	(0.059)	(0.068)	(0.807)
Unit	Years	Years	Years	EUR
Mean Y	14.185	3.497	61.490	67.996
Observations	144,338	144,338	144,338	144,338
R-squared	0.369	0.140	0.086	0.373

Movers comparison including year of birth FE, destination municipality FE, level of education FE, sex FE, occupation FE, NC3 FE. Estimated using OLS. Errors clustered on the level of the origin county.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Robustness: Movers design

Selection into migration.

- Differential pull: destination FE
- Differential push:
  - Predict migration status with treatment selection into migration
  - Use age-at-exposure variation to estimate destination-by-origin FE

Time invariant regional characteristics.

- State level FE
- Balance test treatment and control counties origin county balance table
- Exclude energy sector workers

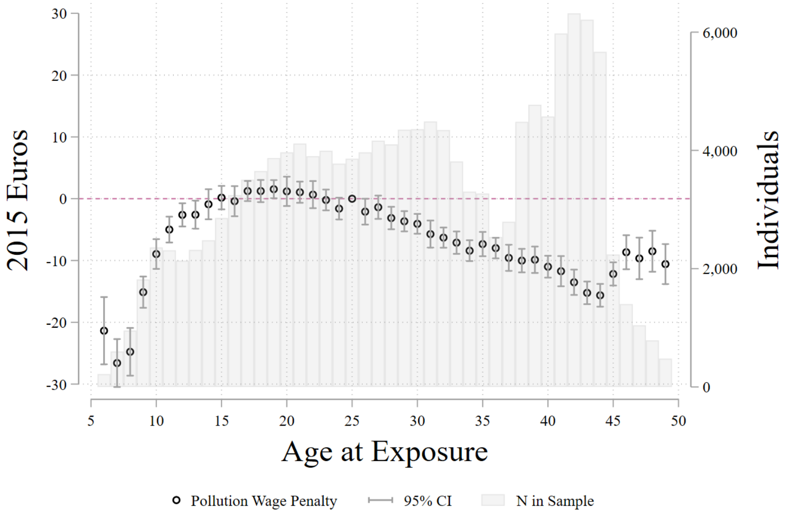
Falsification exercise: Potash mines.

- Placebo-in-place potash placebo test
- Place-based effects of mining culture and exploitative industry

Mechanism: Who is affected and why?

- Which individuals are most impacted by long-term exposure to air pollution?
- Interact treatment indicator with individual's year of birth
- Relative size of effect of pollution on wages:
  - Largest effect for the youngest **and** the oldest individuals in the sample
  - Possible mechanism: effect of pollution exposure depends on individuals' underlying health capital

Effect by Age Cohorts



Conclusion

- This paper investigates the causal impacts of air pollution in a setting with minimal scope for endogenous mitigation.
- Large adverse effects on local infant mortality and birth weights.
- Movers design: long-term detrimental impacts of air pollution on employment, retirement, wages – up to four decades after the initial pollution shock.
- Cohort effects: largest impact on youngest and oldest individuals

Thanks!

Get in touch:
ml@rfberlin.com
@MoritzLubczyk

Overnight trade shock: Fossil fuel imports drop sharply after 1982

1982: Soviet Union suddenly and unexpectedly ends East Germany's preferential access to fossil fuel imports (Schürer, 1999; Pfaff, 2006)

- Limited supply: Soviet Union limits availability of oil, capping it at 17 mio. t. (Steiner, 2004) $\rightarrow$ domestic economic issues.
- New requirement: East Germany has to pay imports from Soviet Union by sending "exportable" goods in return.
- East German energy mix depends on imports $\rightarrow$ domestic lignite as only alternative.

back to lignite mining

Overnight trade shock: Fossil fuel imports drop sharply after 1982

1982: Soviet Union suddenly and unexpectedly ends East Germany's preferential access to fossil fuel imports (Schürer, 1999; Pfaff, 2006)

- Limited supply: Soviet Union limits availability of oil, capping it at 17 mio. t. (Steiner, 2004) → domestic economic issues.
- New requirement: East Germany has to pay imports from Soviet Union by sending "exportable" goods in return.
- East German energy mix depends on imports → domestic lignite as only alternative.

back to lignite mining

Overnight trade shock: Fossil fuel imports drop sharply after 1982

1982: Soviet Union suddenly and unexpectedly ends East Germany's preferential access to fossil fuel imports (Schürer, 1999; Pfaff, 2006)

- Limited supply: Soviet Union limits availability of oil, capping it at 17 mio. t. (Steiner, 2004) → domestic economic issues.
- New requirement: East Germany has to pay imports from Soviet Union by sending "exportable" goods in return.
- East German energy mix depends on imports → domestic lignite as only alternative.

back to lignite mining

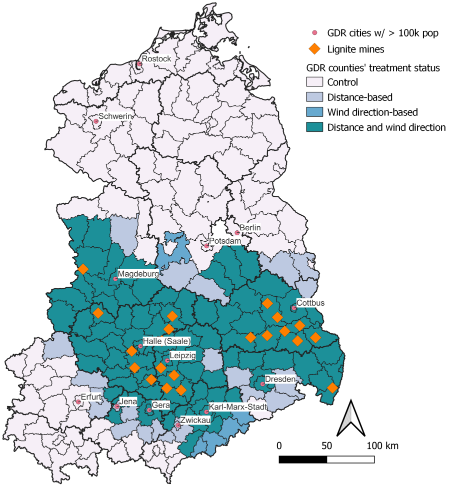
Overnight trade shock: Fossil fuel imports drop sharply after 1982

1982: Soviet Union suddenly and unexpectedly ends East Germany's preferential access to fossil fuel imports (Schürer, 1999; Pfaff, 2006)

- Limited supply: Soviet Union limits availability of oil, capping it at 17 mio. t. (Steiner, 2004) → domestic economic issues.
- New requirement: East Germany has to pay imports from Soviet Union by sending "exportable" goods in return.
- East German energy mix depends on imports → domestic lignite as only alternative.

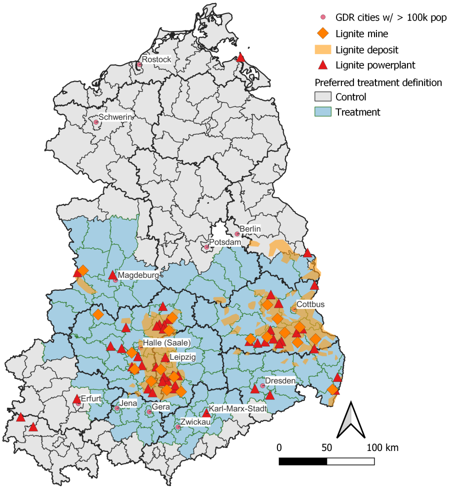
back to lignite mining

Spatial variation in treatment: Alternative map accounting for wind directions ▸ base map



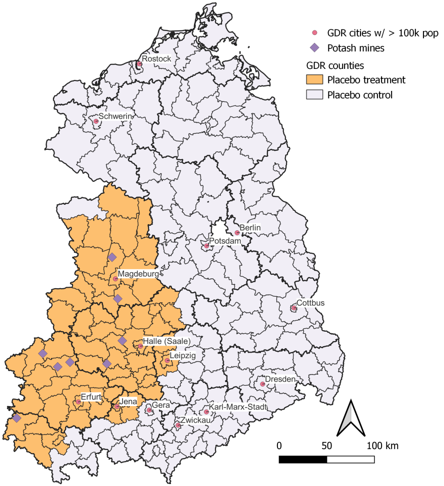
Spatial variation in treatment: Power plant and deposits map

▸ base map



Spatial variation in treatment: Placebo map based on distance to potash mines

▸ base map

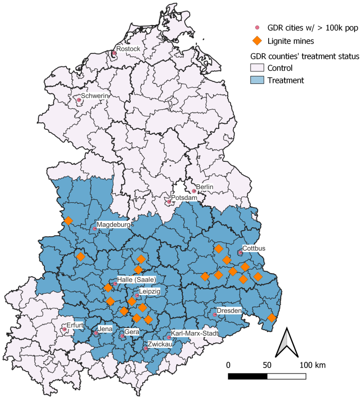


Difference-in-Differences: Approach

Empirical strategy: Use the rapid policy shift towards lignite in a difference-in-differences setting, compare counties close to/far away from lignite deposits before/after 1982. ▸ back

- Uniform timing, binary treatment indicator ▸ treatment assignment
- Reduced form
- "design based" DiD (Rambachan & Roth, 2020)
  - TWFE
  - Errors clustered at county level
  - ATT interpretation

$$Y_{it} = \alpha_i + \phi_t + (1[t > 1982] \times D_i)\beta + \epsilon_{it}$$



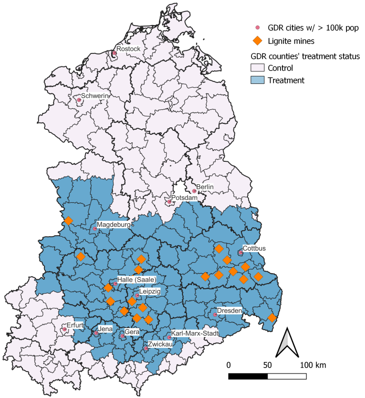
▸ back

Difference-in-Differences: Approach

Empirical strategy: Use the rapid policy shift towards lignite in a difference-in-differences setting, compare counties close to/far away from lignite deposits before/after 1982. back

- Uniform timing, binary treatment indicator treatment assignment
- Reduced form
- "design based" DiD (Rambachan & Roth, 2020)
  - TWFE
  - Errors clustered at county level
  - ATT interpretation

$$Y_{it} = \alpha_i + \phi_t + (1[t > 1982] \times D_i)\beta + \epsilon_{it}$$

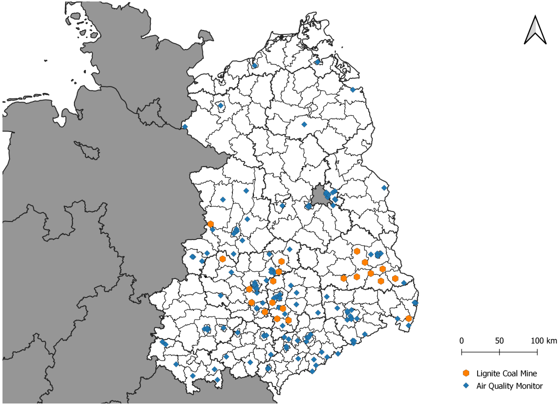


back

Did natural experiment affect local levels of air pollution? Measurement.

- Retrieve previously confidential East German air pollution data
- Annual mean SO_2 readings from >100 official measurement stations
- Considerable heterogeneity in coverage across time and space
- Inverse distance weighted interpolation of air pollution levels
- Unit: county centroids between 1979 and 1989
  → robust to limiting max distance at 50km/100km

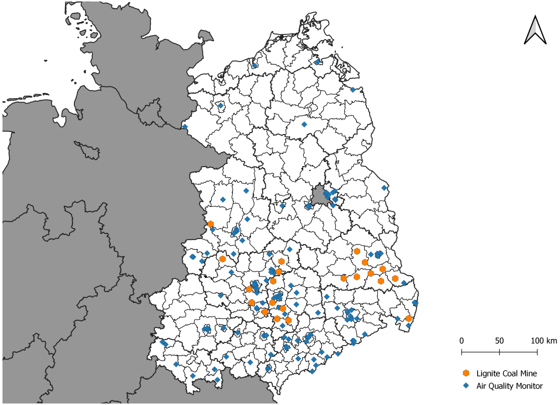
▸ validation overview



Did natural experiment affect local levels of air pollution? Measurement.

- Retrieve previously confidential East German air pollution data
- Annual mean SO_2 readings from >100 official measurement stations
- Considerable heterogeneity in coverage across time and space
- Inverse distance weighted interpolation of air pollution levels
- Unit: county centroids between 1979 and 1989
  $\rightarrow$ robust to limiting max distance at 50km/100km

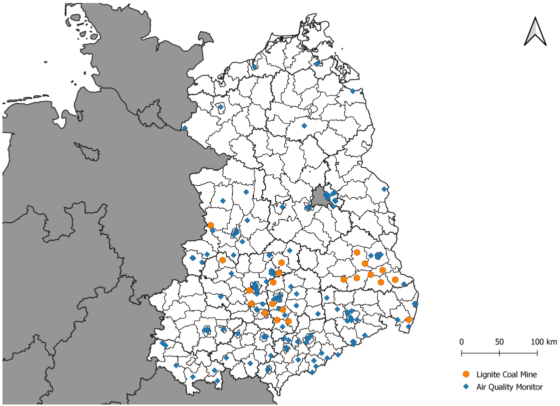
validation overview



Did natural experiment affect local levels of air pollution? Measurement.

- Retrieve previously confidential East German air pollution data
- Annual mean SO_2 readings from $>$100 official measurement stations
- Considerable heterogeneity in coverage across time and space
- Inverse distance weighted interpolation of air pollution levels
- Unit: county centroids between 1979 and 1989
  $\rightarrow$ robust to limiting max distance at 50km/100km

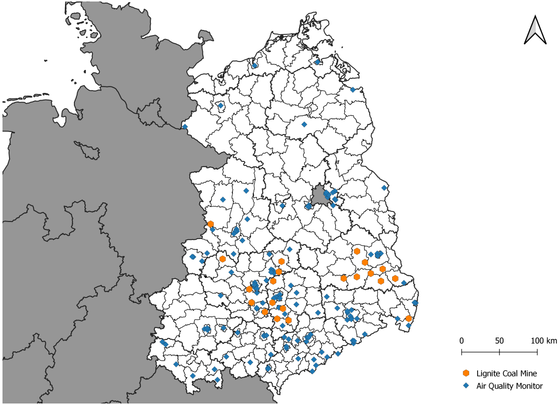
▸ validation overview



Did natural experiment affect local levels of air pollution? Measurement.

- Retrieve previously confidential East German air pollution data
- Annual mean SO_2 readings from $>$100 official measurement stations
- Considerable heterogeneity in coverage across time and space
- Inverse distance weighted interpolation of air pollution levels
- Unit: county centroids between 1979 and 1989
  $\rightarrow$ robust to limiting max distance at 50km/100km

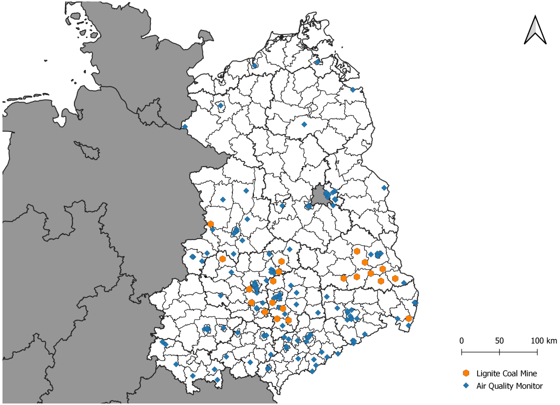
‣ validation overview



Did natural experiment affect local levels of air pollution? Measurement.

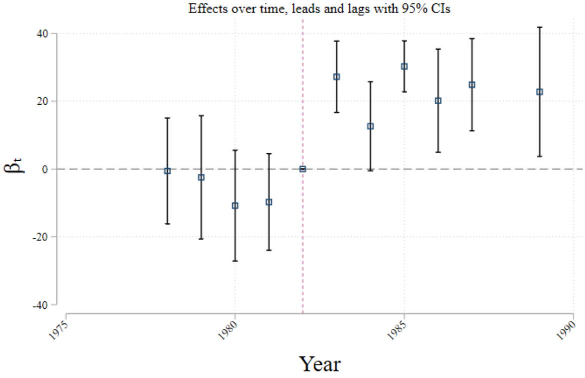
- Retrieve previously confidential East German air pollution data
- Annual mean SO_2 readings from $>$100 official measurement stations
- Considerable heterogeneity in coverage across time and space
- Inverse distance weighted interpolation of air pollution levels
- Unit: county centroids between 1979 and 1989
  $\rightarrow$ robust to limiting max distance at 50km/100km

validation overview



Did natural experiment affect local levels of air pollution? Yes.

	Y = Mean $\mu g/m^3$ of SO_2
	60km
$1[t>1982] \times D_i$	27.553***
	(3.178)
Mean Y	156.002
Observations	2,069
R-squared	0.850



validation overview

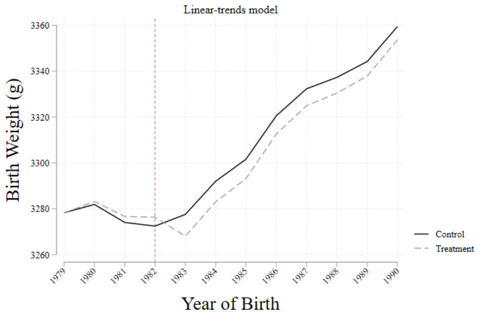
Did shock manifest in individual (health) outcomes? Measurement.

- Birth weight $\rightarrow$ East German birth register (1979-1990)
  - $N > 2$ mio. individuals
  - Additional controls on individual level
    - Municipality
    - Parental socioeconomic status (occupation, education)
    - Parents' age
    - Gender, exact date of birth
- Infant mortality per 1000 live births (1970-1989) $\rightarrow$ digitized from Statistical Yearbooks

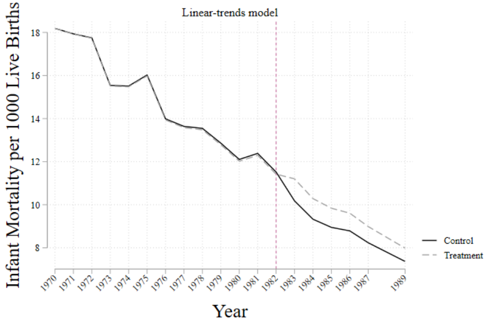
validation overview

Did shock manifest in individual (health) outcomes? Yes.

Birth weight in gram



Infant mortality rate per 1000 live births



▸ validation overview ▸ regression results ▸ DiD event plots ▸ DiD robustness

Difference-in-Differences: Treatment assignment

- Treatment: $D_i = \mathbb{1}$[mine within 60km]
- Robust to varying threshold distance between 40 km and 80 km
- Distance to mine: measure of substitutability
  $\rightarrow$ Lignite is not suitable for trade and transport
- Location of deposits exogenous
- Uses of lignite in East Germany:
  - Electricity generation (close to mines)
  - Other industrial applications
  - Domestic heating
- Corroborate effects with air quality monitor data as "catch-all"

back

DiD: Regression results

$$Y_{it} = \alpha_i + \phi_t + (1[t > 1982] \times D_i)\beta + \epsilon_{it}$$

Y = Birth weight in gram

	60km
1[t>1982] $\times D_i$	-8.328***
	(2.931)
Mean Y	3,282.4
Observations	2,210,149
R-squared	0.030

Y = Infant deaths per 1000 live births

	60km
1[t>1982] $\times D_i$	0.894***
	(0.322)
Mean Y	12.958
Observations	4,104
R-squared	0.386

TWFE specification with year and municipality/county fixed effects. Birth weight model additionally includes month-of-birth FE, gender FE, motherly employment status FE. Estimated using OLS, standard errors clustered at the county level.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

back DiD event plots

Robustness: Alternative specifications and tests

- **Rule out spurious correlation.** ▸ falsification tests
  ⇒ Use placebo in time.
  ⇒ Use potash mines as placebo in space.
- **Disentangle regional variation.** ▸ bins
  ⇒ Rerun specifications with non-parametric distance bins at 15km intervals.
- **Intensity of treatment regressions.** ▸ intensity
  ⇒ Continuous distance to mines. Number of mines within threshold distance.
- **Distributional effects on birth weight.** ▸ DiD weight distribution
  ⇒ Larger effects at lower end of birth weight distribution.
- **Rule out regional idiosyncrasies.**
  ⇒ Switch off and on East Germany's 15 states. Remove border regions.

▸ back

DiD-IV estimation: Elasticities ▸ back

Y =	RF ln(Inf. Mor.)	FS ln(SO_2)	TSLS ln(Inf. Mor.)	TSLS ln(Inf. Mor.)	TSLS ln(Inf. Mor.)
ln(SO_2)			0.263**	0.501*	0.816*
			(0.118)	(0.269)	(0.434)
D_r	0.097**	0.366***			
	(0.044)	(0.034)			
Group trend	–	–	–	✓	–
Unit trend	–	–	–	–	✓
Observations	2,070	2,070	2,070	2,070	2,070
R-squared	0.090	0.146	0.893	0.887	0.966
K-P F-stat	–	118.35	118.35	35.72	26.67

Errors clustered on the county level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Additional facts on lignite

Lignite and pollution

- Lignite used in electricity generation, heating, chemical processing
- High pollutant to heat generation ratio and high transport cost to economic value ratio
- Sulfur and carbon dioxide
- Airborne pollutants can travel vast distances as they decay to particulate matter

back to lignite back to empirical setting

Alternative Outcome: Weight at Birth – Results

Outcome: Individuals' weight at birth in gram.

$$Y_{it} = \alpha_m + \phi_t + (1[t > 1982] \times D_m)\beta + \mathbf{Xb} + \epsilon_{it}$$

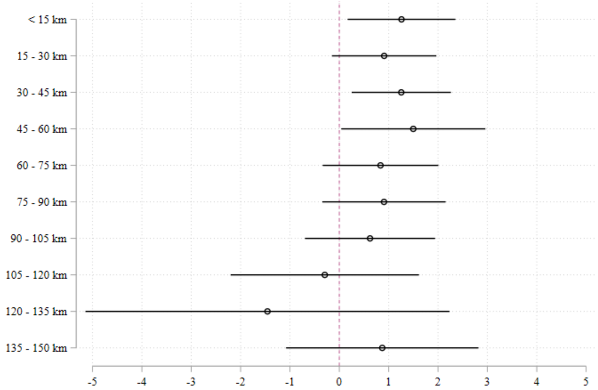
	60km	40km	50km	70km	80km
	(1)	(2)	(3)	(4)	(5)
$1[t>1982] \times D_i$	-8.328***	-7.512**	-8.713***	-5.346*	-2.112
	(2.931)	(3.293)	(3.166)	(2.987)	(2.943)
Mean Y	3,282.4	3,282.4	3,282.4	3,282.4	3,282.4
Observations	2,210,149	2,210,149	2,210,149	2,210,149	2,210,149
R-squared	0.030	0.030	0.030	0.030	0.030

TWFE specification with year and municipality fixed effects. Includes month-of-birth FE, gender FE, motherly employment status FE.
Errors clustered at county level. Estimated using OLS.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Where is infant mortality rising?

Indicator variables for nearest lignite mine in distance range. Base category: > 150 km

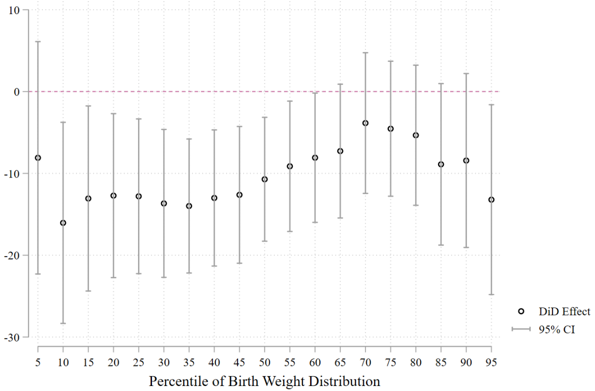
$$Y_{it} = \alpha_i + \phi_t + \Sigma_j (1[Distance_j] \times D_i)\beta_j + \epsilon_{it}$$



back

Air pollution effects larger at lower end of birth weight distribution

Municipality level panel: running separate difference-in-differences estimations across percentiles of within-municipality birth weight distribution.



Pctl.	DiD Effect	H0: Linear trends parallel		H0: No anticipation effect		Mean Y
	β	*F statistic*	*p value*	*F statistic*	*p value*	*gram*
10th	-16.737***	1.09	0.2970	0.46	0.7073	2,709.80
	(6.094)					
20th	-13.350***	0.00	0.9721	0.19	0.9056	2,925.84
	(4.651)					
30th	-12.552***	0.28	0.5983	0.71	0.5464	3,075.81
	(4.116)					
40th	-11.496***	0.21	0.6470	1.09	0.3506	3,204.73
	(3.674)					
50th	-9.848***	0.16	0.6852	0.52	0.6670	3,319.81
	(3.33)					
60th	-7.357**	0.12	0.7269	0.48	0.6947	3,434.15
	(3.476)					
70th	-5.306	1.51	0.2193	1.42	0.2353	3,555.89
	(3.716)					
80th	-4.225	0.75	0.3858	1.60	0.1862	3,691.27
	(3.889)					
90th	-6.216	0.01	0.9032	0.40	0.7532	3,867.30
	(4.593)					

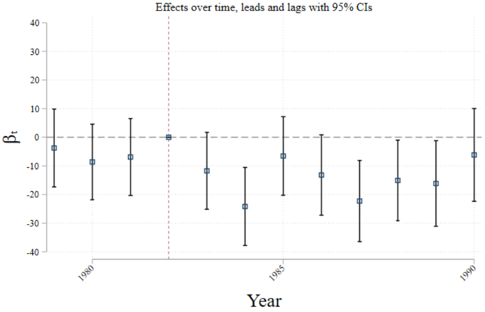
Separate TWFE-DiD specifications with year and municipality fixed effects at each percentile. Estimated using OLS.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

back

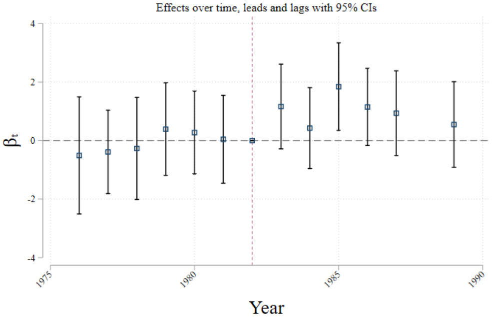
DiD: Event plots for short-term outcomes

Plots: Coefficients from DiD regression including lags and leads of treatment.

Birth weight in gram



Infant mortality rate per 1000 live births



back

Robustness: Placebo-Treatments – Effect on Infant Mortality rate

Aim: Rule out spurious correlation.

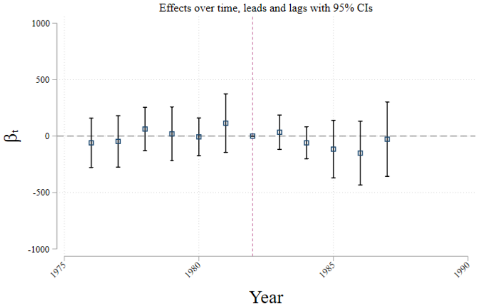
$$Y_{it} = \alpha_i + \phi_t + (1[Placebo] \times D_i)\beta + \epsilon_{it}$$

Placebo Treatment	60km	40km	50km	70km	80km	
	(1)	(2)	(3)	(4)	(5)	N
Panel A: Placebo-in-Time						
Brown Coal, 1973	-0.241	-0.208	0.036	-0.369	-0.527	2,808
	(0.492)	(0.493)	(0.479)	(0.512)	(0.528)	
Brown Coal, 1976	-0.394	-0.001	0.013	-0.501	-0.445	2,808
	(0.404)	(0.394)	(0.395)	(0.416)	(0.433)	
Brown Coal, 1979	0.086	0.299	0.334	0.010	0.186	2,808
	(0.399)	(0.389)	(0.387)	(0.415)	(0.432)	
Panel B: Placebo-in-Space						
Potash	-0.028	0.220	-0.361	-0.032	0.484	4,104
	(0.571)	(0.677)	(0.585)	(0.548)	(0.542)	

TWFE specifications with year and county fixed effects. Panel A: Placebo-treatments using data between 1970 and 1982, interacting distance to lignite mining with pseudo-treatment dates. Panel B: Placebo-treatments using full data set, interacting distance to potash mining with actual treatment data. Estimated using OLS.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

back

Robustness: Effect of Treatment on Population Change



- Key assumption: authoritarian nature of socialist regime limited individuals' freedom of movement (Lichter et al., 2021).
- To corroborate: generate measure of natural population movement. $NetMigration = Population_t - Population_{(t-1)} - Births + Deaths$
- Use resulting measure of population change as outcome in DiD.

did validation overview

Robustness: Intensity of Treatment – Effect on Infant Mortality rate

Aim: Leverage full extent of variation.

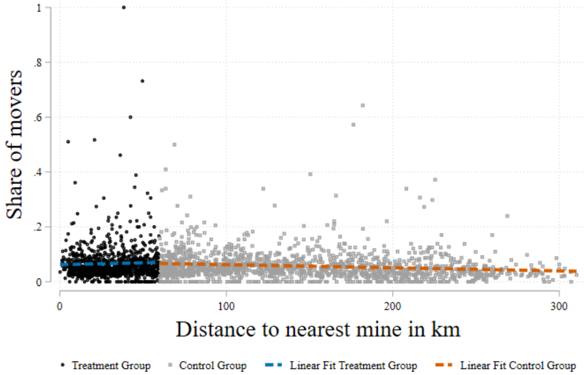
	Y = Infant Mortality Rate		
Binary Treatment 1[60km]	0.894***		
	(0.322)		
Distance to Nearest Mine		-0.007***	
		(0.003)	
# Mines within 60km			0.087*
			(0.049)
Mean Y	12.958	12.958	12.958
Observations	4,104	4,104	4,104
R-squared	0.386	0.386	0.386

TWFE specifications with year and county fixed effects. Estimated using OLS.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Selection into Migration

$Y = 1[\text{Move next year}]$	
	60km
D_i	0.005***
	(0.000)
Mean Y	0.0747
Observations	5,550,967
R-squared	0.0369

OLS estimate including year of birth FE, occupation FE, level of education FE, sex FE, NC3 FE. Restricted to individuals observed in 1991/92.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$



back

Balance test: Treatment vs. control counties (weighted)

back

	East German Counties	
	(1)	(2)
Log Mean Population, 1980-88	0.039	-0.077
	(0.386)	(0.174)
Share of Population $<$ 15y, 1989	-1.875***	-0.250
	(0.342)	(0.347)
Share of Population $>$ 64y, 1989	2.928***	0.661
	(0.442)	(0.509)
SED Party Membership	0.560	1.718
	(2.006)	(2.207)
March '90 Voter Turnout	0.222	-0.243
	(0.506)	(0.524)
December '90 Voter Turnout	0.340	-0.276
	(0.856)	(1.152)
Log Industrial Output, 1989	0.408	-0.132
	(0.429)	(0.298)
Patents Filed, 1989	-63.860	45.583
	(222.474)	(32.074)
I[University or College]	0.008	-0.006
	(0.156)	(0.189)
Number of Universities or Colleges	0.056	-0.235
	(1.091)	(0.461)
Observations	216	216
State FE		✓

	East German Counties	
	(1)	(2)
Employment Shares:		
Agriculture, 09/89	-1.558	1.061
	(2.267)	(3.155)
Energy, 09/89	2.929**	4.462**
	(1.291)	(2.230)
Chemicals, 09/89	2.324**	1.229
	(1.151)	(1.077)
Metal, 09/89	-0.084	-1.475
	(0.666)	(0.994)
Engineering, 09/89	4.029***	0.312
	(1.498)	(1.839)
Water and Sewage, 09/89	-0.054	-0.006
	(0.075)	(0.068)
Electronics, 09/89	-1.128	-1.904
	(1.660)	(2.000)
Mgmt, Planning and R&D, 09/89	0.308	0.070
	(0.545)	(0.604)
Cooperative Workers, 09/89	-0.725	0.497
	(1.709)	(2.350)
Observations	216	216
State FE		✓

Falsification test: Potash mines in East Germany

Y =	Employment	Unemployment	Retirement Age	Daily Wages
P_r	-0.147	0.075	-0.057	-0.502
	(0.139)	(0.067)	(0.064)	(0.661)
Unit	Years	Years	Years	EUR
Mean Y	14.187	3.497	61.490	68.002
N	144,352	144,352	144,352	144,352
R^2	0.370	0.140	0.086	0.373

OLS estimates including year of birth FE, level of education FE, sex FE, occupation FE, NC3 FE.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

- Potash in East Germany:
  $\rightarrow$ eight mines in southwest of country
  $\rightarrow$ potash: mineral salt mainly used in production of fertilizers
- Define distance to nearest potash mine as placebo test
- Rerun movers design

back

Literature: What are the causal impacts of air pollution?

- Mainly correlational evidence from epidemiological research
  → Many obstacles to causal interpretations
  → In particular: endogenous response to mitigate impact of pollution

- Causal inference using quasi-experimental variation shows large health effects.
  (e.g. Currie & Walker, 2011; Luechinger, 2014; Cesur et al., 2017; Beach & Hanlon, 2018; Deryugina et al., 2019; Anderson, 2020; Alexander & Schwandt, 2022)

- Mostly focused on short-run outcomes in localized applications
  (Notable exceptions: Chen et al., 2013; Ebenstein et al., 2017; Bishop et al., 2019; Barreca et al., 2021; Deryugina & Reif, 2022)

- Existing evidence on labor market effects limited to specific episodes of acute early-life exposure
  (Isen et al., 2017; Colmer et al., 2022)

back