Mines ☛ Rivers ☛ Yields

Downstream Mining Impacts on Agriculture in Africs

Lukas Vashold (lvashold@wu.ac.at)

Gustav Pirich (gustav@pirich.at)

Max Heinze (mheinze@wu.ac.at)

Nikolas Kuschnig (nkuschnig@wu.ac.at)

Department of Economics, Vienna University of Economics and Business (WU Vienna)

May 16, 2025

Snapshot

Background

Methods & Data

Results

Mines – Curse or Blessing?

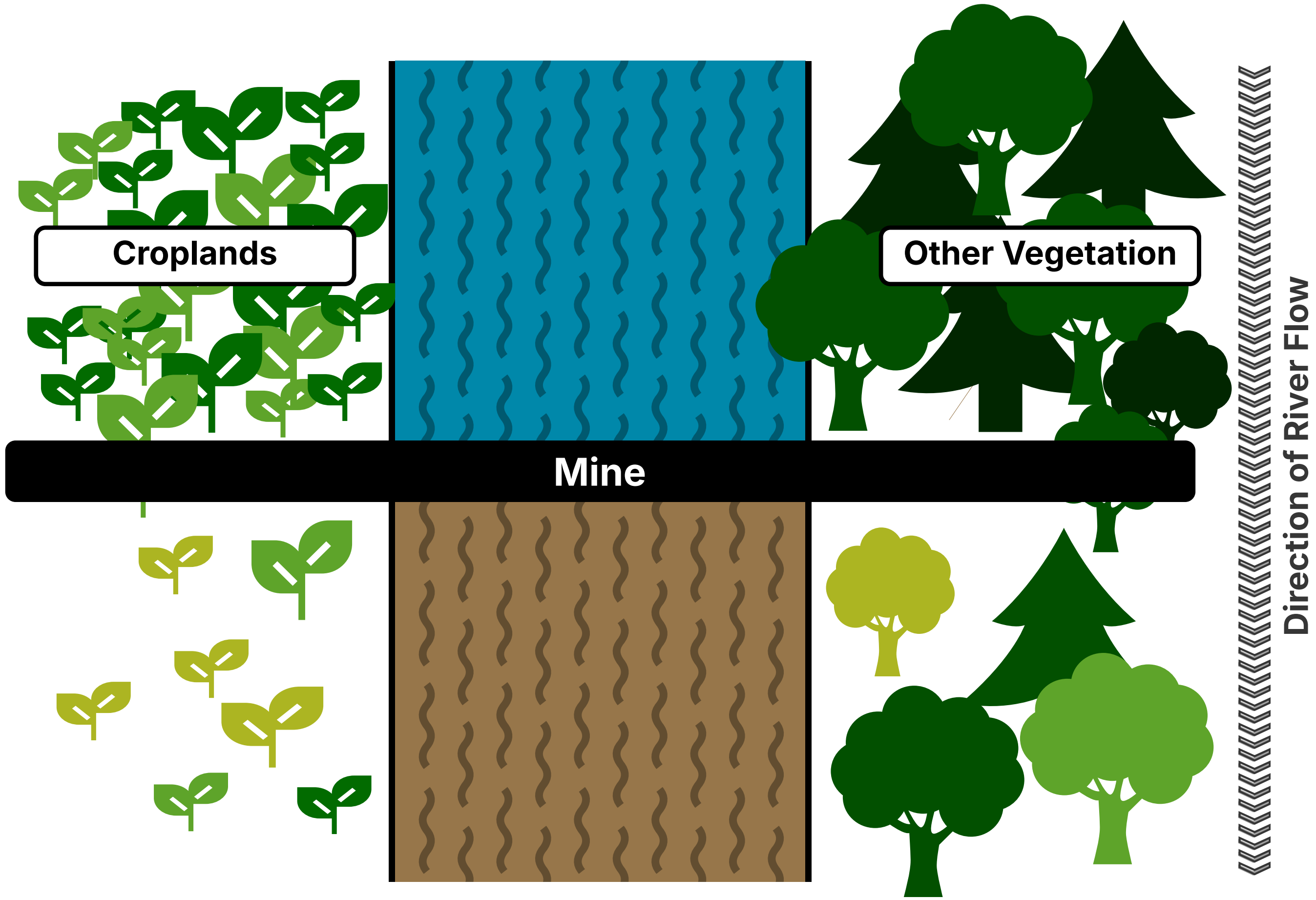
A Blessing?

- Demand for relevant **minerals** is projected to increase **fourfold** until 2050 (Hund et al., 2023).
- **Extraction Benefits** include:
    - enabling the **green transition**,
    - increasing local **incomes** (Bazillier & Girard, 2020),
    - and improving **wealth** and **asset ownership** (Goltz & Barnwal, 2019).

A Curse?

- Resource extraction causes **negative externalities**.
- **Ecological effects** include the following:
    - Mines **use water** and produce **sediments and tailings** (Moura et al., 2022).
    - Pollutants include **mercury** and **lead** (Schwarzenbach et al., 2010).
    - Industrial pollution **harms plant growth** (Ruppen et al., 2023).

How Pollution Travels



If **water pollution** from mines affects vegetation, we should observe reduced vegetation health **downstream** of a mine.

Using a **remotely-sensed vegetation index**, we find evidence for less healthy vegetation **downstream**.

Snapshot

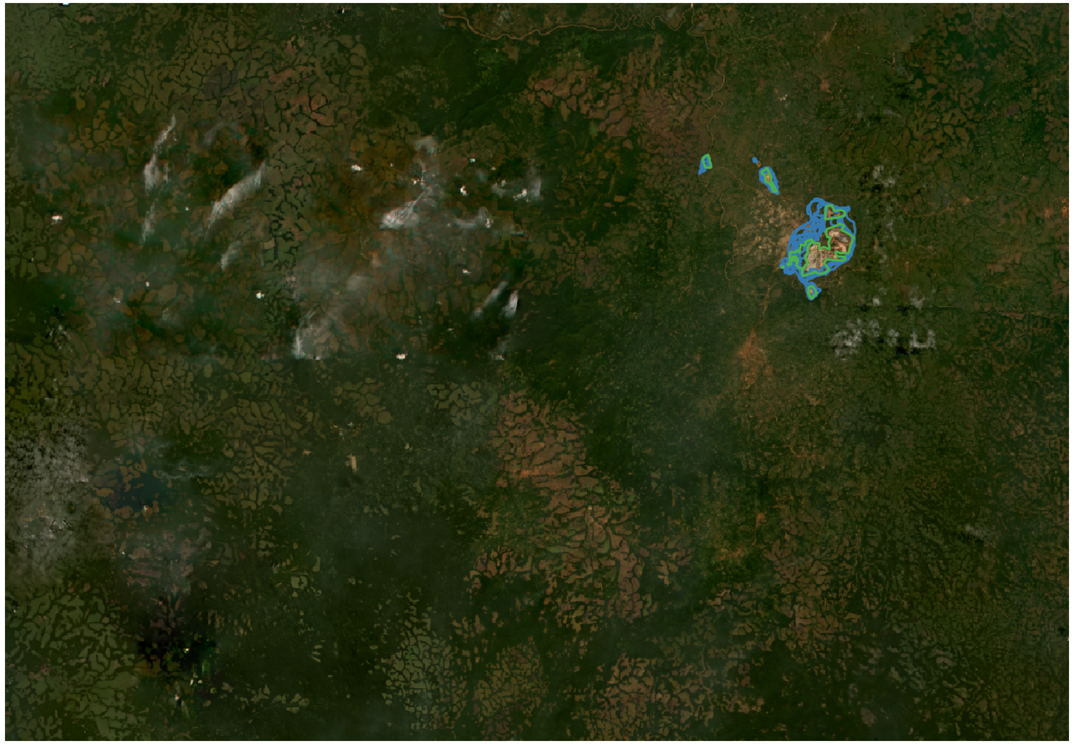
Background

Methods & Data

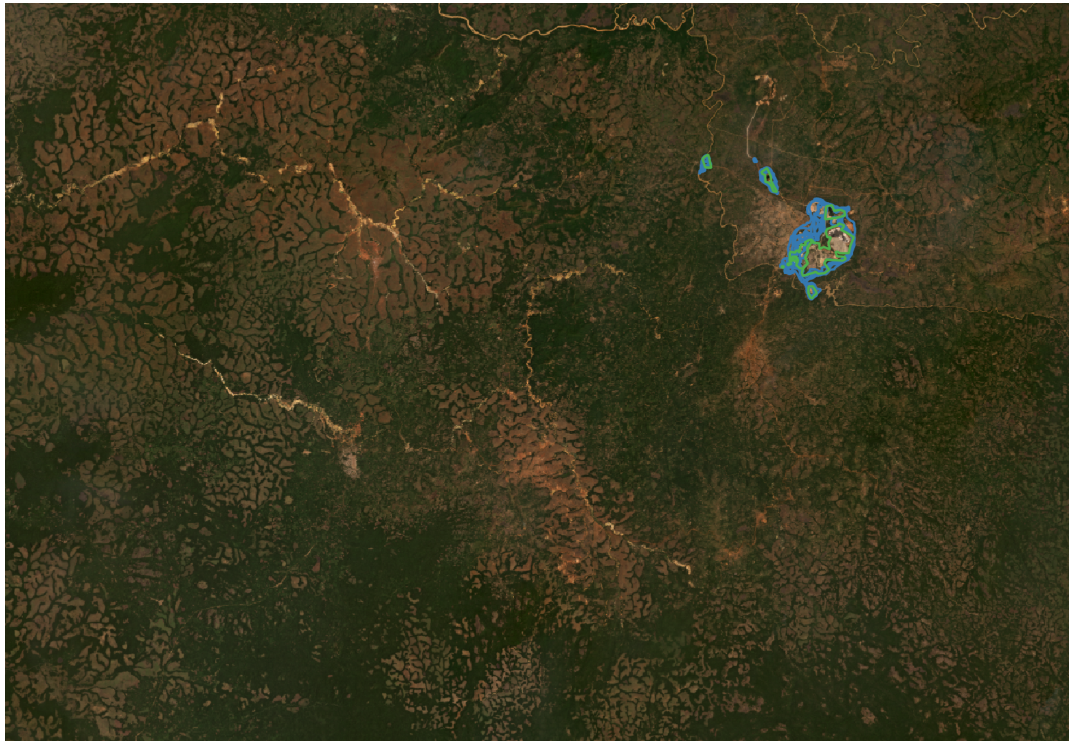
Results

Appendix

Mining and Water Pollution



2015



2025

Research Question

What is the causal effect of water pollution from mining on agricultural productivity in Africa?

- **Africa** is a particularly interesting focus because
  - it has a **booming mining industry** (International Council on Mining and Metals, 2022),
  - with many **artisanal and small-scale mines** (ASM Inventory, 2022; Girard et al., 2022),
  - and a **lack of containment facilities** (Kossoff et al., 2014; Macklin et al., 2023).
- Negative effects are **more locally concentrated** than benefits.
- Informing this discussion enables **improved environmental governance**.

Snapshot

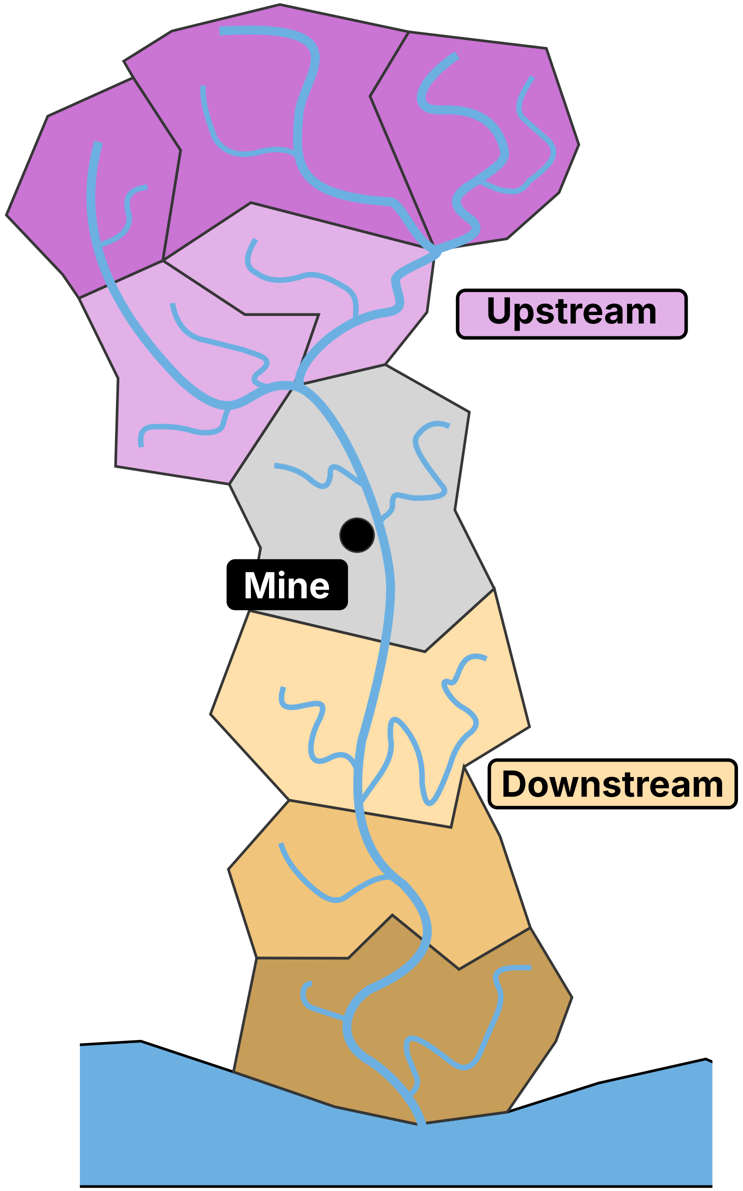
Background

Methods & Data

Results

Appendix

How to Find Affected Areas



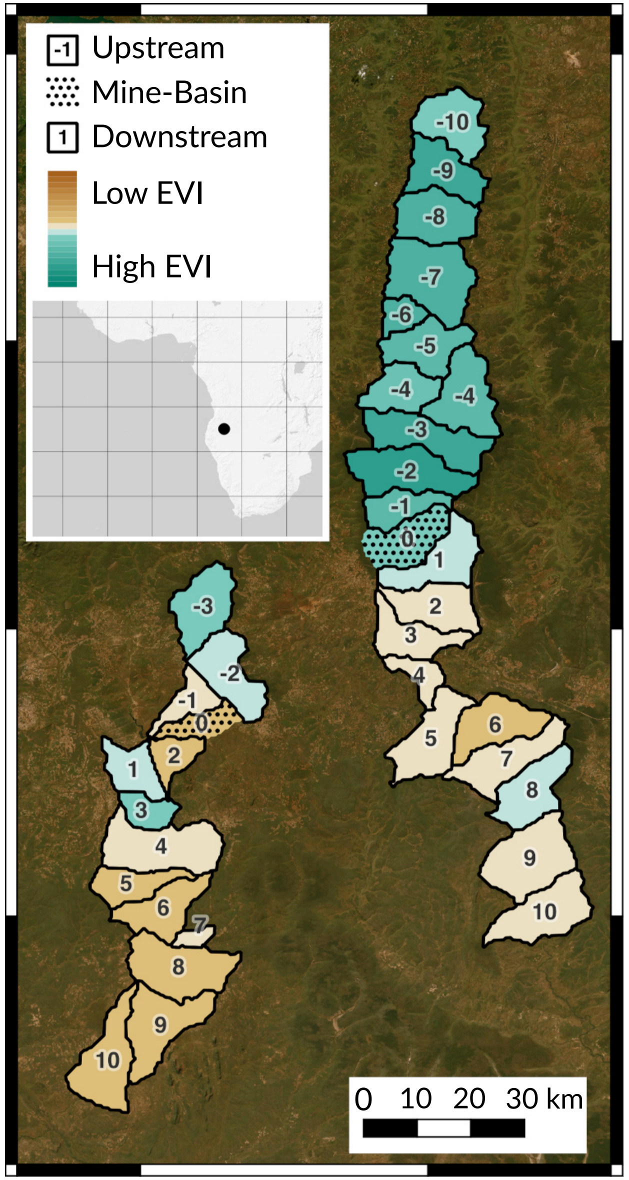
Our **unit of observation** is the **river basin**. Lehner & Grill (2013) provide a nested basin collection, of which we use the **most granular level**.

If we spill a cup of water anywhere in a basin, it always ends up in the next basin **downstream**.

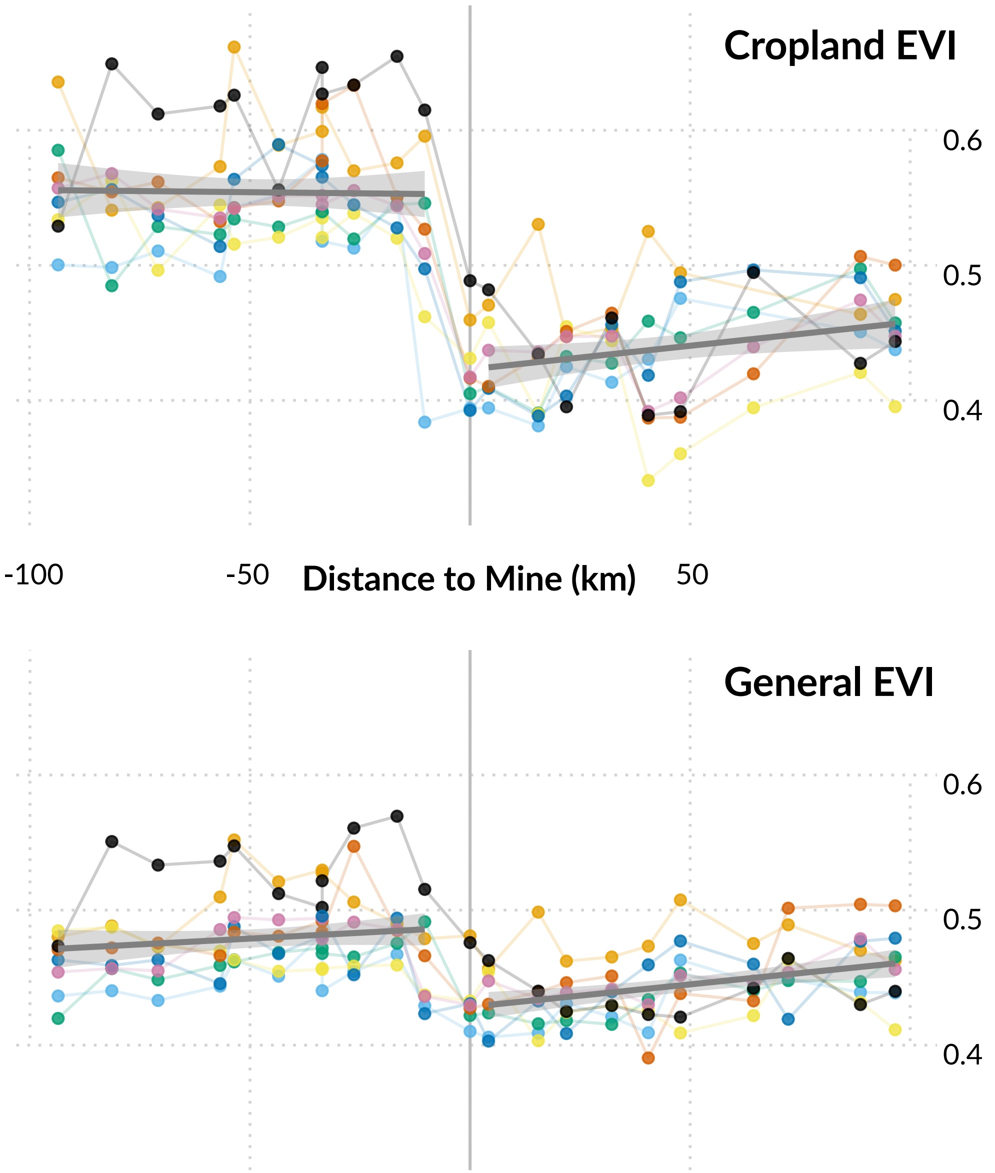
Water moves from **upstream** to **downstream** of a mine.

• show actual basins

Intuition



-1 Upstream
Mine-Basin
1 Downstream
Low EVI
High EVI
0 10 20 30 km

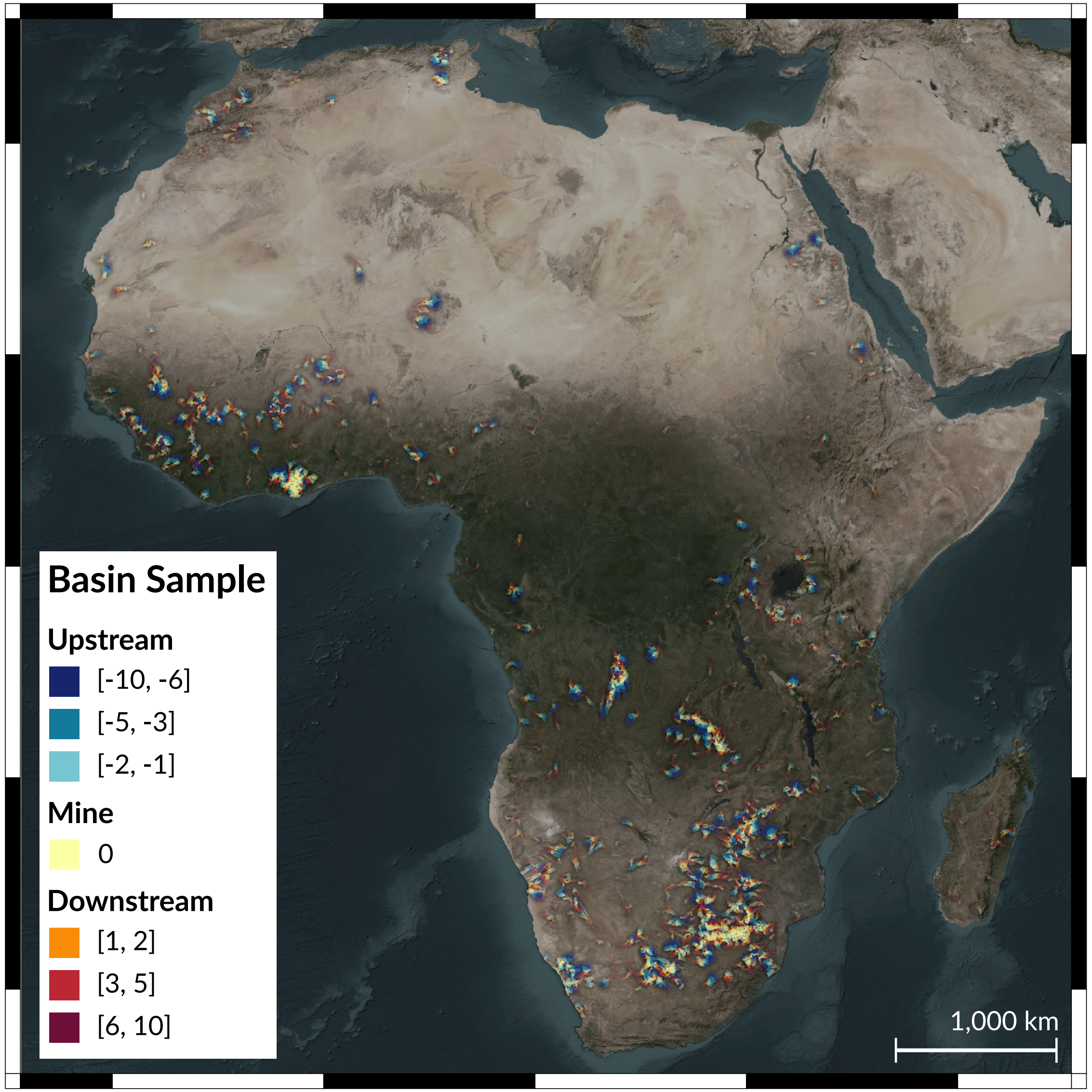


Cropland EVI
0.6
0.5
0.4
-100
-50
Distance to Mine (km)
50
General EVI
0.6
0.5
0.4

Mines

We use **mine locations** from Maus et al. (2022), which includes some ASM sites.

We then designate **mine basins** and determine 10 levels each of **upstream** and **downstream** basins.



Variables and Observations

Outcome

- We use the **Enhanced Vegetation Index (EVI)**, which
  - is **remotely sensed**, and
  - ranges **between −1** (water) **and 1** (dense vegetation).
- We extract the **annual maximum** on
  - all areas covered by **vegetation**, and
  - only on **croplands**.

Observations and Covariates

- We observe **6,307 upstream** basins, **1,900 mine** basins, and **6,127 downstream** basins over $T = 8$ years. • show order × basins
- We collate **covariates** on:
  - topography (Amatulli et al., 2018),
  - soil type (Hengl et al., 2017),
  - climate (Abatzoglou et al., 2018), and
  - soioeconomic characteristics (Weiss et al., 2018).

Empirical Strategy (Spatial RDD)

We estimate:

$$y_{imt} = \boldsymbol{\beta}' \, F(x_i) + \boldsymbol{\theta}' W_{it} + \mu_m + \psi_t + \varepsilon_{imt},$$

where we let $F(x_i)$ return indicators:

$$f(x_i)_j = \mathbb{I}(x = j) \qquad \text{for } j \in \{-10, \ldots, -2, 0, 1, 2, \ldots, 10\}.$$

- y_{imt}: **outcome** of basin i near mine m in year t,
- μ_m and ψ_t: mine and year **fixed effects**,
- $\boldsymbol{W}_{it}$: basin-specific **covariates**.

- Parameter $\boldsymbol{\beta}'$ is identified under the assumption that there are **no other discontinuous changes** at the mine basin.

Identification Assumption

- To validate our results, we conduct a battery of robustness checks to address potential threats to identification.

  (1) Extensive set of controls: **geophysical features** (elevation, slope, distance to coast, soil composition), **meteorological conditions** (temperature, precipitation), and **socioeconomic indicators** (population, accessibility, conflict)

  (2) **placebo outcomes** • show more

  (3) **matching procedure** • show more

  (4) **permutation test** • show more

Snapshot

Background

Methods & Data

Results

Appendix

Results Overview

- We find a **significant reduction** in vegetation health **downstream** of mines.
- Impacts are particularly strong in **fertile regions** and where **gold mining** predominates.
- These results are **robust** to varying the sample, the outcome measurement, and the level of fixed effects.

Results

- Reduction in peak biomass by **1.28–1.35%** relative to the mean on an affected area of **255,000 km²**.
- For croplands, the estimates imply a **1.38–1.47% reduction** across **74,000 km²**.

Outcome	*Peak Vegetation*		*Peak Cropland Veg.*	
	(Plain)	**(Full)**	**(Plain)**	**(Full)**
Pooled Order				
Downstream (1st–3rd)	-0.0057***	-0.0056***	-0.0064**	-0.0068***
	(0.0018)	(0.0020)	(0.0025)	(0.0026)
Fit statistics				
Observations	110,576	110,524	93,036	93,000
R^2	0.903	0.908	0.816	0.822
Controls				
Geophysical	No	Yes	No	Yes
Meteorological	No	Yes	No	Yes
Socioeconomic	No	Yes	No	Yes
Fixed-effects				
Year (2016–2023)	Yes	Yes	Yes	Yes
Mine	Yes	Yes	Yes	Yes

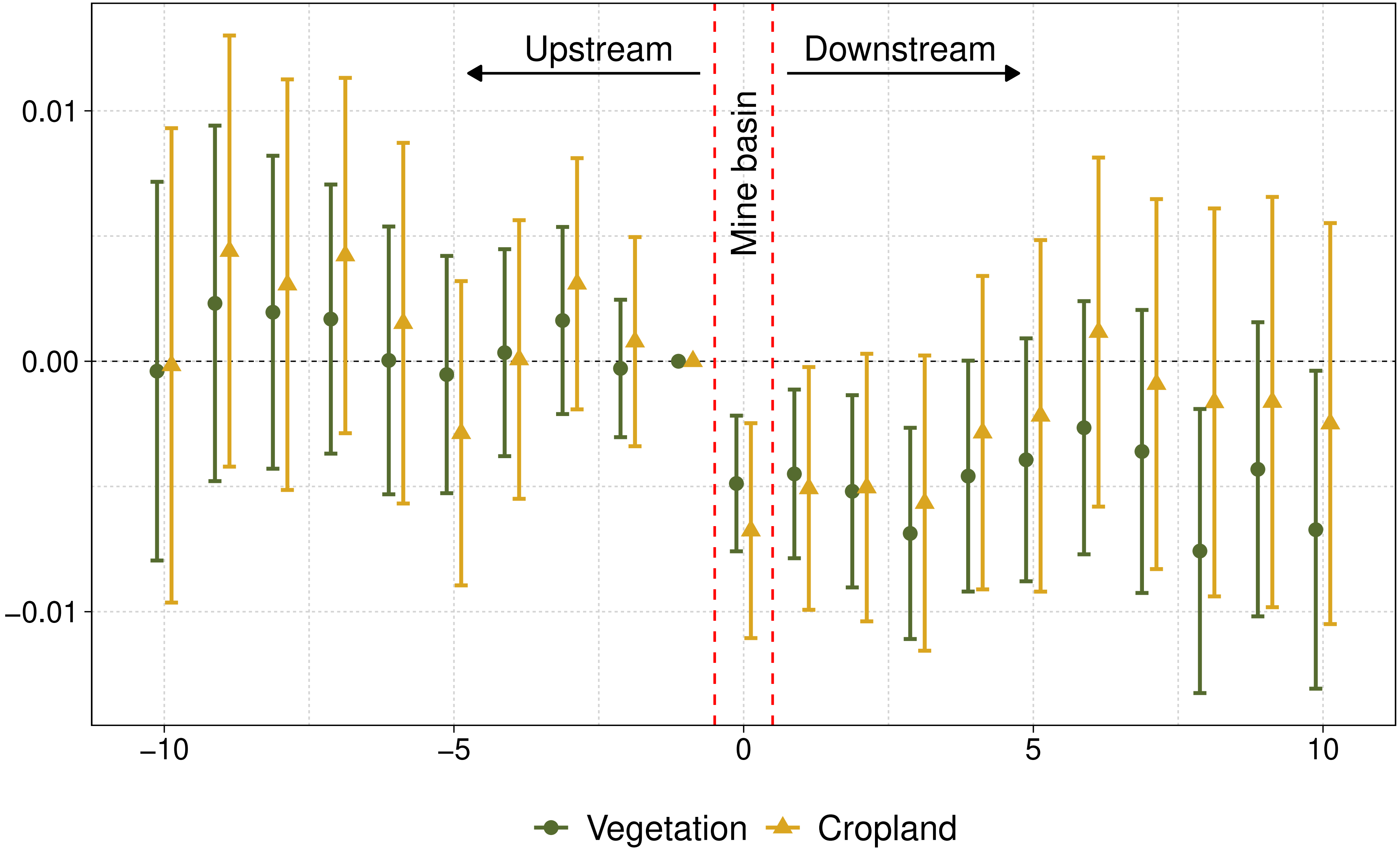
Clustered (by mine-basin) standard errors in parentheses.

*Significance levels: ***: 0.01, **: 0.05, *: 0.1.*

Results – What Do These Impacts Mean?

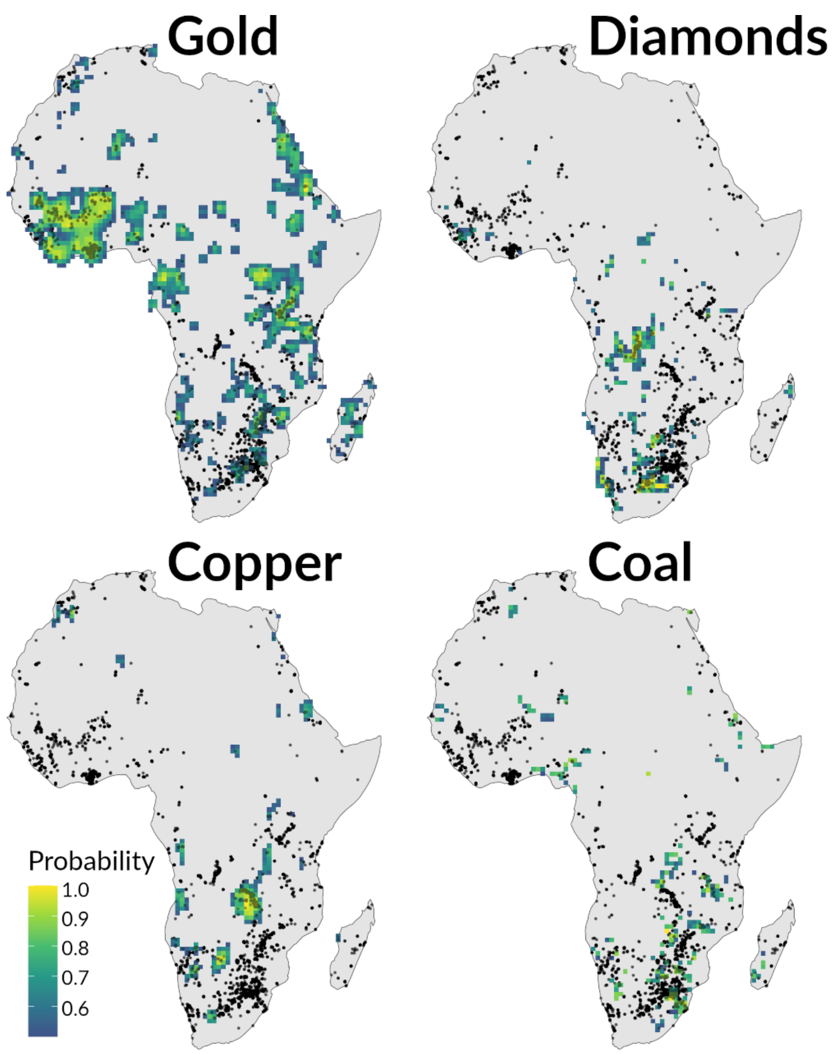
- We use **survey data** from farmers in Africa from IFPRI (2020) to estimate the crop yield–EVI elasticity.
- Our measure is **highly predictive of crop yields.** • show more
- We estimate a **2.16–2.31% decrease** in the **value** of **overall crop production**.
- This amounts to a rerduction in **agricultural production** of about **91,000 metric tons of cereals**,
- comparable to **5.4%** of the 1.7 million tons that the **World Food Program WFP** distributes annually.

How Far Downstream Do These Effects Persist?



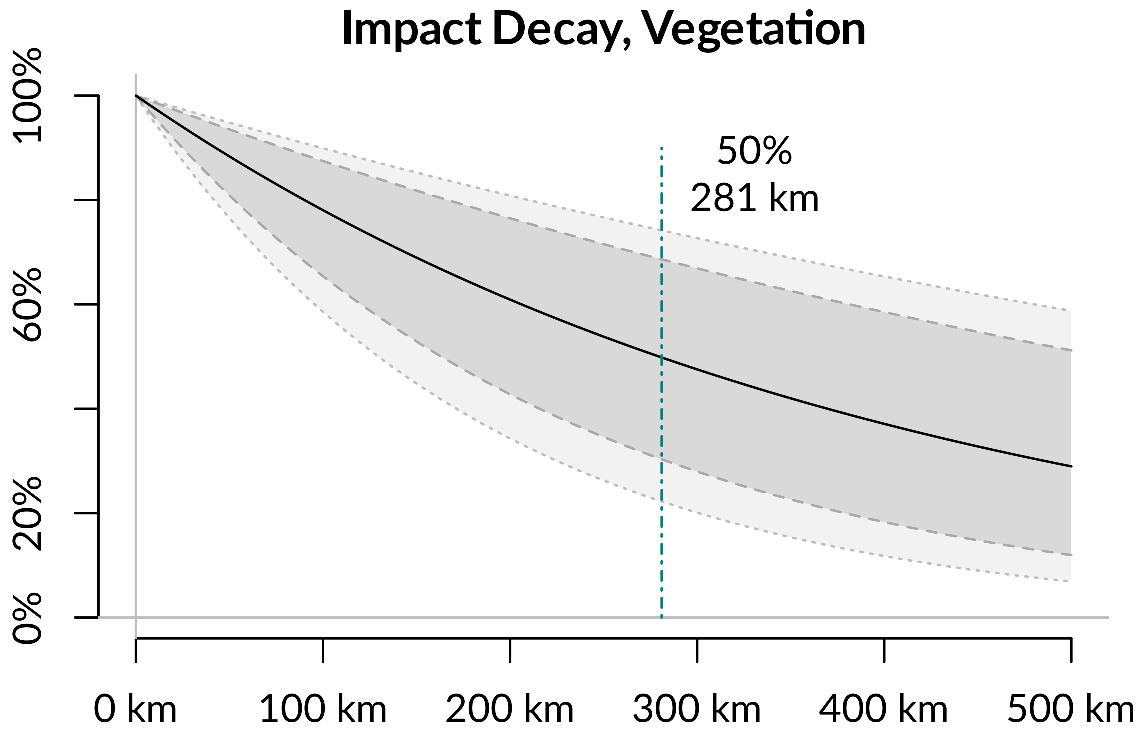
Heterogeneity – Which Areas Are Affected?

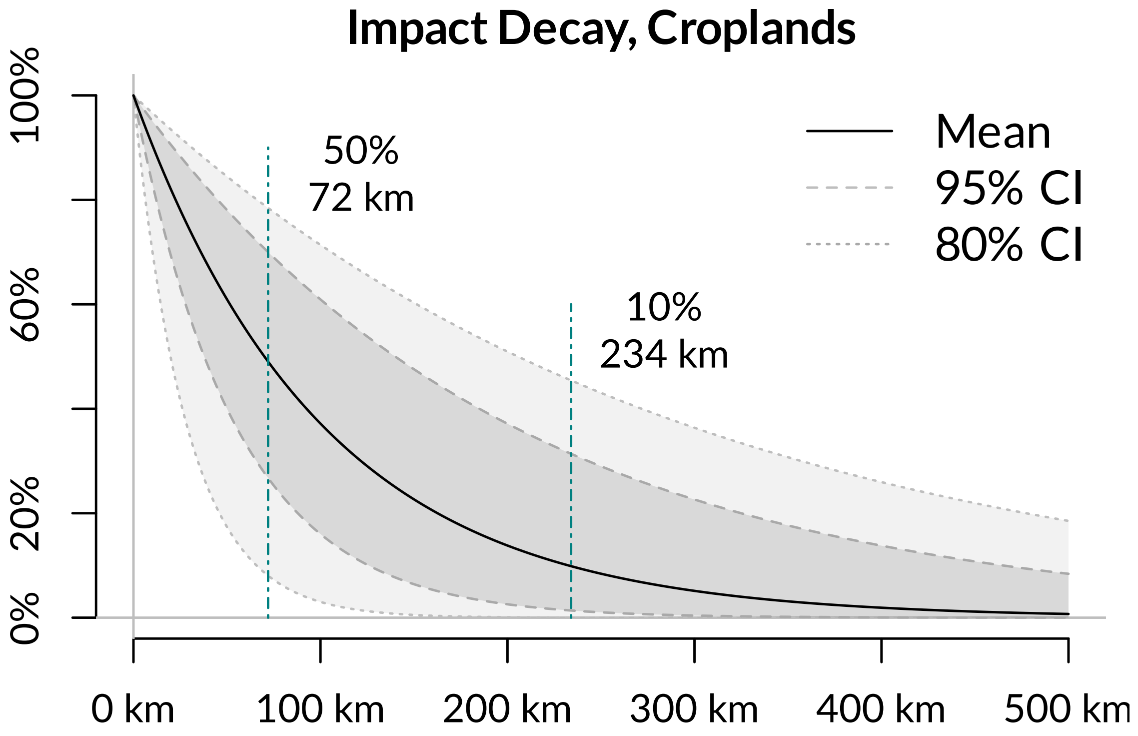
- **Commodity Types** (Padilla et al., 2021):
  - Gold mining regions are particularly affected
- **Region**:
  - greater effect in West Africa
- **Characteristics of the mine** (Jasansky et al., 2023):
  - greater effect for larger mines
- **Biome** (Food and Agriculture Organization of the United Nations et al., 2023)
  - larger effect for mines located on highly fertile croplands
  - larger effect in grasslands and forests



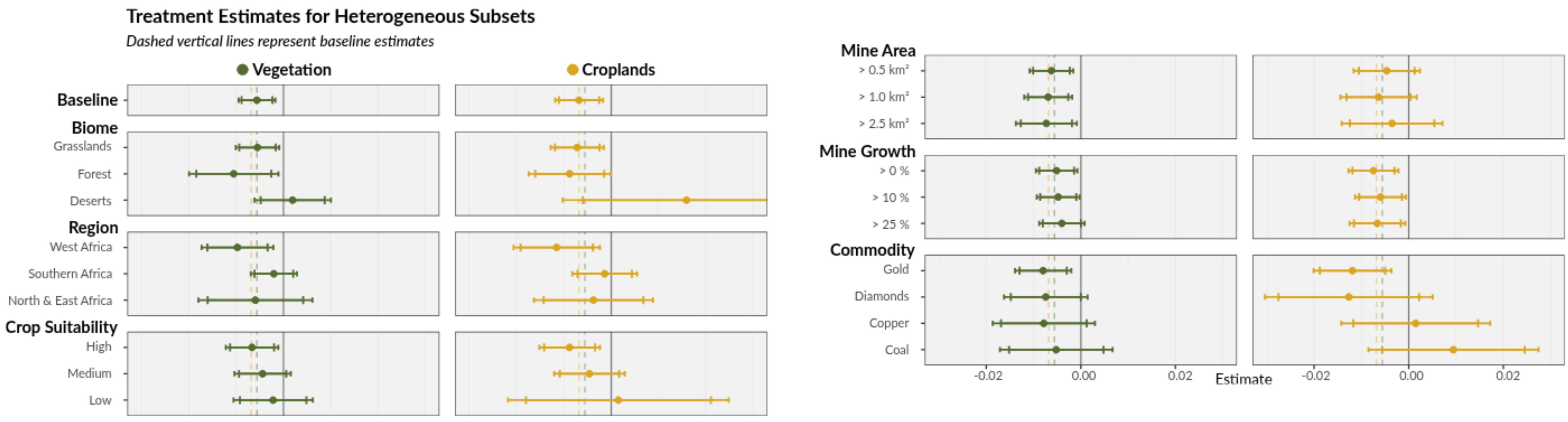
Impact Decay Assessment

- We re-estimate the main specification using an **exponential distance decay** function, $\exp(-\delta d_{ij})$, where d_{ij} is the distance along the river from a mine.





Heterogeneity – Which Areas Are Affected?

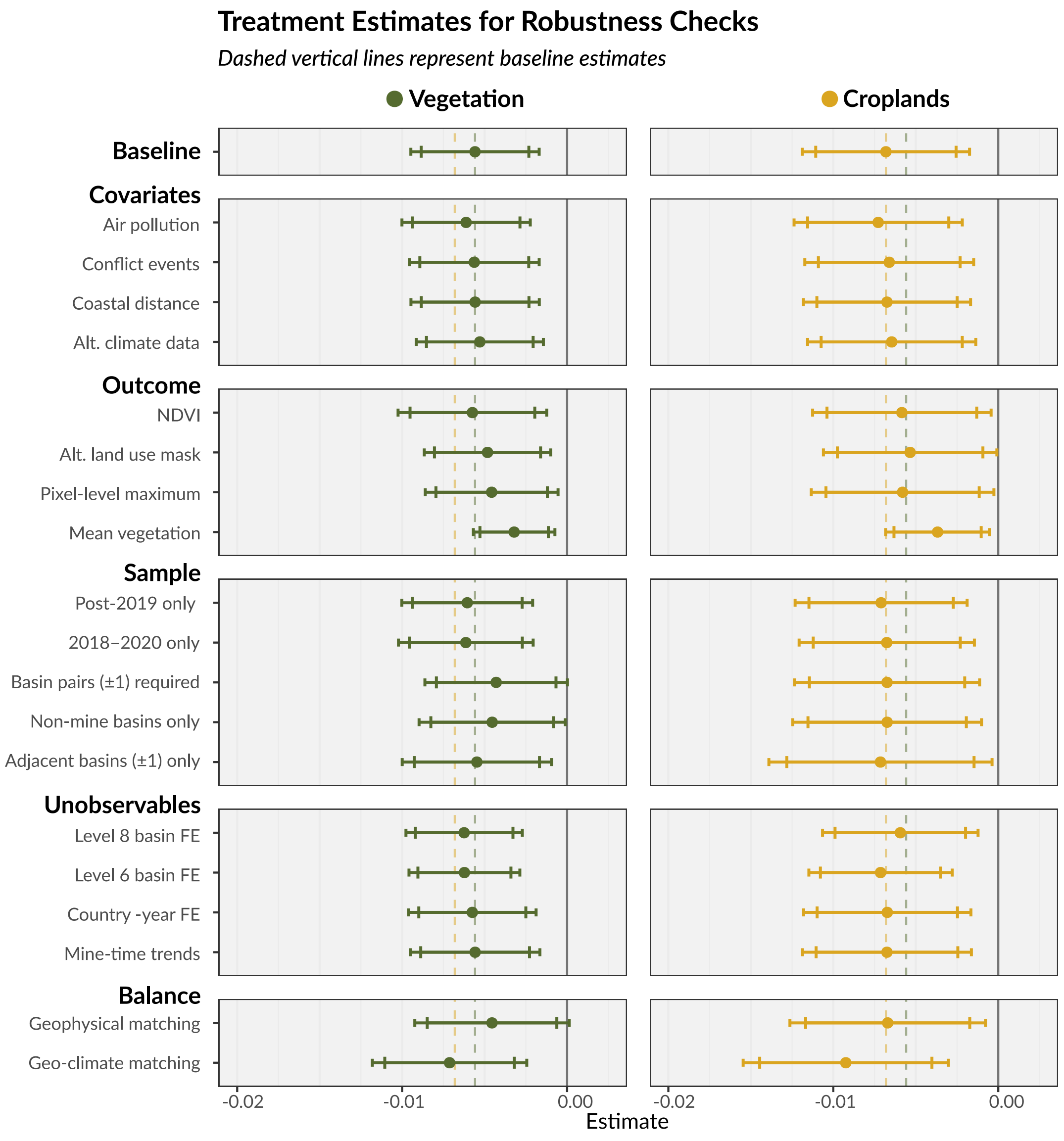


- Stronger effects in **West Africa** and on **highly fertile croplands**.
- **Gold mining** regions show **75% larger impacts** on croplands.

Robustness

- We conduct a battery of robustness checks along several dimensions.

- We check the validity of our assumptions by conducting additional exercises.
  - permutation test
  - placebo outcomes
  - matching on covariates



More on Water Pollution

- We identify the impact of water pollution from mining on agricultural productivity.
- Water quality data is sparse, and limited to South Africa.
- We provide suggestive evidence of levels of **water pollution** as a mediating mechanism (United Nations Environment Programme, 2025)
- We find **elevated levels of key pollutant measures** for mine and downstream basins. • show more

Limitations

- Remotely sensed vegetation indices capture **crop yields** only indirectly.
- Differentiating between **artisanal and industrial mining** remains challenging due to data scarcity and co-location of mines.
- No direct evidence on **farmers' adaptive responses** (e.g., migration, altered crop choices, or barrier to irrigation).
- Noise from multiple data sources likely leads to **attenuation of true estimates**.

Conclusion

We identified the causal effects of mining

- on agricultural productivity,
- mediated by water pollution.

Our results showed a **negative impact** on vegetation health.

Effects were particularly **strong** for larger mines, gold mining regions, and in regions with highly fertile croplands.

Results were **robust** to changes of treatment, outcome, sample, methods, and estimation procedures.

References

*This list is **scrollable**.*

Abatzoglou, J. T., Dobrowski, S. Z., Parks, S. A., & Hegewisch, K. C. (2018). TerraClimate, a high-resolution global dataset of monthly climate and climatic water balance from 19582015. *Scientific Data*, *5*(170191), 1–12. https://doi.org/10.1038/sdata.2017.191

Amatulli, G., Domisch, S., Tuanmu, M.-N., Parmentier, B., Ranipeta, A., Malczyk, J., & Jetz, W. (2018). A suite of global, cross-scale topographic variables for environmental and biodiversity modeling. *Scientific Data*, *5*(180040), 1–15. https://doi.org/10.1038/sdata.2018.40

ASM Inventory. (2022). *World maps of artisanal and small-scale mining*. http://artisanalmining.org/Inventory/

Azzari, G., Jain, M., & Lobell, D. B. (2017). Towards fine resolution global maps of crop yields: Testing multiple methods and satellites in three countries. *Remote Sensing of Environment*, *202*, 129--141. https://doi.org/10.1016/j.rse.2017.04.014

Bazillier, R., & Girard, V. (2020). The gold digger and the machine. Evidence on the distributive effect of the artisanal and industrial gold rushes in Burkina Faso. *Journal of Development Economics*, *143*, 102411. https://doi.org/10.1016/j.jdeveco.2019.102411

Becker-Reshef, I., Vermote, E., Lindeman, M., & Justice, C. (2010). A generalized regression-based model for forecasting winter wheat yields in Kansas and Ukraine using MODIS data. *Remote Sensing of Environment*, *114*(6), 1312--1323. https://doi.org/10.1016/j.rse.2010.01.010

Bolton, D. K., & Friedl, M. A. (2013). Forecasting crop yield using remotely sensed vegetation indices and crop phenology metrics. *Agricultural and Forest Meteorology*, *173*, 74--84. https://doi.org/10.1016/j.agrformet.2013.01.007

Cattaneo, M. D., Idrobo, N., & Titiunik, R. (2019). *A practical introduction to regression discontinuity designs: foundations*. Cambridge University Press. https://doi.org/10.1017/9781108684606

Food and Agriculture Organization of the United Nations, International Fund of Agricultural Developmenmt, United Nations International Children's Emergency Fund, United Nations World Food Programme, & World Health Organization. (2023).

Background

Methods & Data

Results

Appendix

Illustration from Lehner & Grill (2013)

• go back

Appendix · A Proxy for Agricultural Activity

We get a **proxy** for **agricultural productivity** like this:

(1) Filter out **cloud cover**.

(2) Aggregate the **mean EVI** per basin.

(3) Take the **annual maximum** per basin per year. **→ Max. EVI**

(4) Apply a **cropland mask**. **→ Max. Cropland EVI**

This **peak vegetation index** has been shown to proxy well for crop yields (Azzari et al., 2017; Becker-Reshef et al., 2010; Bolton & Friedl, 2013; Johnson, 2016).

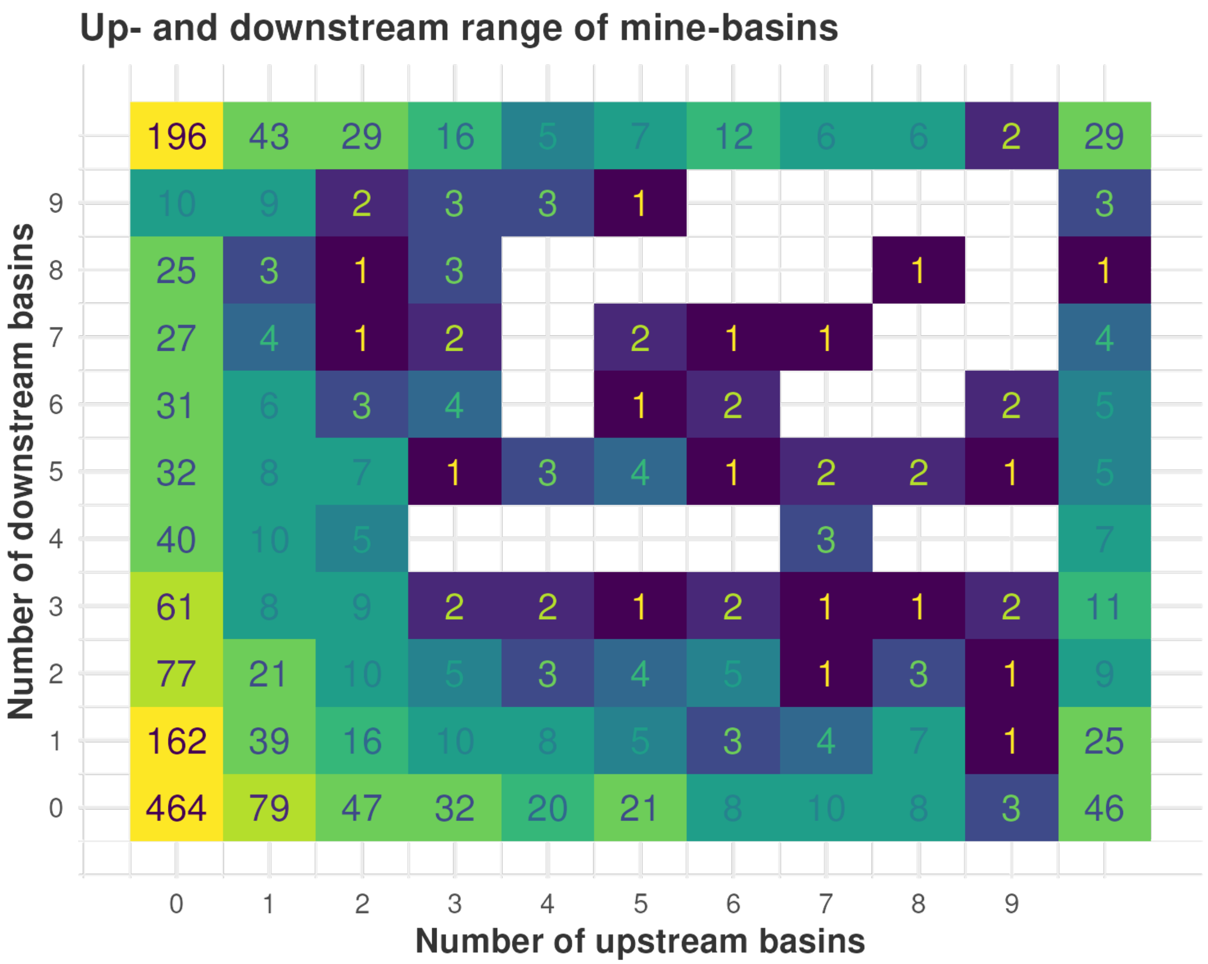
Appendix · Summary Statistics

Variable	Unit	NT	Mean	St. Dev.	Min.	Max.
Peak Vegetation	Index [−1, 1]	110,576	0.428	0.154	0.016	0.993
Mean Vegetation	Index [−1, 1]	110,576	0.279	0.112	−0.021	0.578
Peak Cropland Veg.	Index [−1, 1]	93,036	0.464	0.133	−0.068	0.978
Mean Cropland Veg.	Index [−1, 1]	93,036	0.298	0.101	−0.104	0.601
Elevation	Meters	110,568	820.4	481.1	−118.3	3,059.7
Slope	Degrees	110,568	2.23	2.34	0.0	20.9
Max. Temperature	Degree Celsius	110,572	34.3	3.9	15.6	48.8
Precipitation	Millimeters	110,576	901.8	595.2	0.64	4,456.7
Population	Capita	110,576	8,471	37,716	0.0	1,396,921
Accessibility	Minutes	110,528	164.3	179.1	1.0	2,659.9

Appendix · Impact Decay Assessment

- We re-estimate our main specification using an **exponential decay** function: $\exp(-\delta d_{ij})$.
- **Hydrological studies** on dispersion patterns suggest using an exponential decay function.
- Since the **decay parameter** is unknown, we conduct a grid search for $\delta \in [0.001, 2]$.
- We then use a **Bayesian Model Averaging** approach with BIC as marginal likelihood approximation.
- Finally, we compute the **mean effect decay** at increasing distances.

Appendix · Basin Numbers



The number of mine-basins with Y upstream and X downstream basins in the dataset. • go back

Appendix · Basins by Order

Order	Upstream		Downstream	
	N	Distance	N	Distance
0	(1900)	(0.0)	(1900)	(0.0)
1	847	13.9	1162	11.1
2	781	24.5	882	22.0
3	722	35.0	743	32.7
4	698	44.9	643	43.3
5	653	55.3	578	54.0
6	576	66.3	512	64.3
7	562	75.8	458	74.1
8	522	86.5	416	84.4
9	494	95.8	382	95.0
10	452	104.2	351	104.7

Appendix · Full Order Specifiaction Results

Outcome	*Peak Vegetation*		*Peak Cropland Veg.*	
(Specification)	**(Plain)**	**(Full)**	**(Plain)**	**(Full)**
Individual Order				
Downstream (1st)	-0.0045***	-0.0043**	-0.0051**	-0.0050**
	(0.0017)	(0.0018)	(0.0025)	(0.0025)
Downstream (2nd)	-0.0049**	-0.0048**	-0.0058*	-0.0067**
	(0.0022)	(0.0024)	(0.0031)	(0.0032)
Downstream (3rd)	-0.0085***	-0.0087***	-0.0088**	-0.0099***
	(0.0028)	(0.0029)	(0.0037)	(0.0038)
Downstream (4th)	-0.0049*	-0.0062*	−0.0029	−0.0044
	(0.0030)	(0.0033)	(0.0038)	(0.0040)
Downstream (5th)	−0.0034	−0.0053	0.0007	−0.0016
	(0.0033)	(0.0037)	(0.0042)	(0.0045)
Fit statistics				
Observations	110,576	110,524	93,036	93,000
R^2	0.903	0.908	0.816	0.822
Pooled Order				
Downstream (1st–3rd)	-0.0057***	-0.0056***	-0.0064**	-0.0068***
	(0.0018)	(0.0020)	(0.0025)	(0.0026)
Fit statistics				
Observations	110,576	110,524	93,036	93,000
R^2	0.903	0.908	0.816	0.822
Controls				
Geophysical	No	Yes	No	Yes
Meteorological	No	Yes	No	Yes
Socioeconomic	No	Yes	No	Yes
Fixed-effects				
Year (2016–2023)	Yes	Yes	Yes	Yes
Mine	Yes	Yes	Yes	Yes

Clustered (Mine) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Appendix · Distance Specification Results

Dependent Variables:	Maximum Vegetation EVI				Maximum Croplands EVI			
Model:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Linear distance								
Downstream	-0.0050^{**}	-0.0045^{**}	-0.0033	-0.0034	-0.0050^{*}	-0.0049^{*}	-0.0041	-0.0042
	(0.0023)	(0.0022)	(0.0022)	(0.0022)	(0.0029)	(0.0029)	(0.0029)	(0.0029)
Downstream × Distance	-7.57×10^{-6}	-3.59×10^{-5}	-8.32×10^{-5}	-8.47×10^{-5}	1.47×10^{-5}	-4.19×10^{-6}	-5.85×10^{-5}	-5.96×10^{-5}
	(4.69×10^{-5})	(5.36×10^{-5})	(5.38×10^{-5})	(5.32×10^{-5})	(5.85×10^{-5})	(6.91×10^{-5})	(6.96×10^{-5})	(6.94×10^{-5})
Distance	7.75×10^{-6}	3.26×10^{-5}	5.61×10^{-5}	6.18×10^{-5}	2.75×10^{-5}	4.08×10^{-5}	6.32×10^{-5}	5.66×10^{-5}
	(3.91×10^{-5})	(4.13×10^{-5})	(4.12×10^{-5})	(4.04×10^{-5})	(4.97×10^{-5})	(5.45×10^{-5})	(5.45×10^{-5})	(5.3×10^{-5})
Fit statistics								
Observations	110,576	110,568	110,564	110,524	93,036	93,036	93,032	93,000
R^2	0.90282	0.90452	0.90762	0.90783	0.81609	0.81748	0.82138	0.82165
Linear–quadratic distance								
Downstream	-0.0056^{**}	-0.0055^{**}	-0.0050^{**}	-0.0052^{**}	-0.0077^{**}	-0.0076^{**}	-0.0072^{**}	-0.0073^{**}
	(0.0027)	(0.0026)	(0.0025)	(0.0025)	(0.0035)	(0.0036)	(0.0035)	(0.0035)
Downstream × Distance	2.64×10^{-5}	2.01×10^{-5}	5.75×10^{-6}	5.45×10^{-6}	0.0002	0.0001	0.0001	0.0001
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0002)	(0.0002)	(0.0002)	(0.0002)
Downstream × Distance^2	-3.04×10^{-7}	-4.7×10^{-7}	-7.27×10^{-7}	-7.35×10^{-7}	-1.2×10^{-6}	-1.17×10^{-6}	-1.38×10^{-6}	-1.36×10^{-6}
	(8.52×10^{-7})	(8×10^{-7})	(8.09×10^{-7})	(7.99×10^{-7})	(1.2×10^{-6})	(1.2×10^{-6})	(1.18×10^{-6})	(1.18×10^{-6})
Distance	3.97×10^{-5}	3.93×10^{-5}	3.33×10^{-5}	3.64×10^{-5}	-4.23×10^{-6}	1.24×10^{-6}	1.17×10^{-5}	-1.21×10^{-6}
	(9.08×10^{-5})	(8.61×10^{-5})	(8.93×10^{-5})	(8.63×10^{-5})	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Distance^2	-2.43×10^{-7}	-5.04×10^{-8}	1.76×10^{-7}	1.97×10^{-7}	2.55×10^{-7}	3.18×10^{-7}	4.13×10^{-7}	4.64×10^{-7}
	(6.32×10^{-7})	(5.85×10^{-7})	(6.05×10^{-7})	(5.91×10^{-7})	(9.26×10^{-7})	(9.37×10^{-7})	(9.11×10^{-7})	(9.2×10^{-7})
Fit statistics								
Observations	110,576	110,568	110,564	110,524	93,036	93,036	93,032	93,000
R^2	0.90283	0.90453	0.90762	0.90784	0.81612	0.81751	0.82142	0.82168
Exponential decay								
	$\delta = 0.005$	$\delta = 0.006$	$\delta = 0.002$	$\delta = 0.002$	$\delta = 0.035$	$\delta = 0.035$	$\delta = 0.020$	$\delta = 0.010$
$\exp\{-\delta\times\text{Distance}\}$ × Downstream	-0.0062^{***}	-0.0062^{***}	-0.0060^{***}	-0.0062^{***}	-0.0093^{***}	-0.0091^{***}	-0.0074^{**}	-0.0068^{**}
	(0.0023)	(0.0023)	(0.0023)	(0.0023)	(0.0034)	(0.0033)	(0.0029)	(0.0029)
Fit statistics								
Observations	110,576	110,568	110,564	110,524	93,036	93,036	93,032	93,000
R^2	0.901147	0.902842	0.905958	0.906169	0.812592	0.813949	0.817862	0.818141
Fixed-effects								
Year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mine	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

*Clustered (mine basin) standard-errors in parentheses   Significance: ***: 0.01, **: 0.05, *: 0.1*

Appendix · Varying Sample Definition

Dependent Variables:	Maximum EVI					Maximum Cropland EVI				
Model:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Variables										
Downstream × Order = 0	-0.0059^{***}	-0.0076^{***}	-0.0062^{***}			-0.0095^{***}	-0.0082^{***}	-0.0094^{***}		
	(0.0013)	(0.0014)	(0.0012)			(0.0020)	(0.0024)	(0.0022)		
Downstream × Order = 1	-0.0057^{***}	-0.0053^{***}	-0.0053^{***}	-0.0049^{**}	-0.0051^{**}	-0.0061^{**}	-0.0049	-0.0051^{*}	-0.0061^{**}	-0.0069^{*}
	(0.0017)	(0.0020)	(0.0017)	(0.0020)	(0.0021)	(0.0026)	(0.0032)	(0.0030)	(0.0030)	(0.0039)
Downstream × Order = 2	-0.0066^{***}	-0.0054^{**}		-0.0056^{**}		-0.0062^{**}	-0.0057		-0.0062^{*}	
	(0.0021)	(0.0026)		(0.0023)		(0.0030)	(0.0037)		(0.0033)	
Fixed-effects										
Year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mine	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fit statistics										
Observations	114,496	61,712	32,360	99,320	9,168	94,604	50,914	27,589	81,278	7,623
R^2	0.92395	0.91566	0.93993	0.92392	0.93378	0.78597	0.76613	0.84032	0.78332	0.81766
Within R^2	0.05582	0.05702	0.05650	0.05511	0.07364	0.02531	0.02382	0.03446	0.02322	0.03884

Clustered (Mine) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Research Design

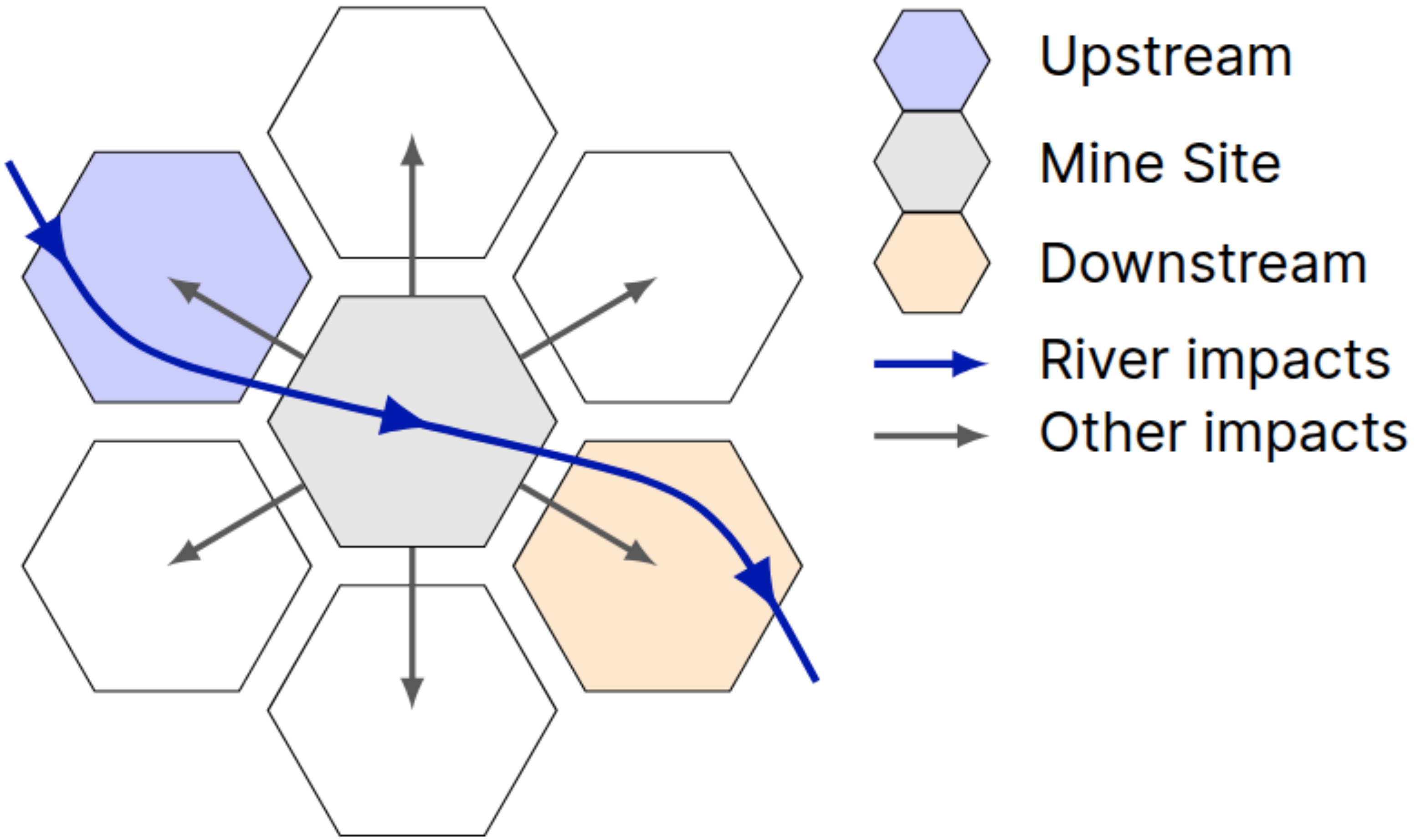


Illustration of the research design. The comparison of up- and downstream basins enables the identification of mine impacts that are mediated by the river.

Appendix · Varying Outcome / Fixed Effects

Dependent Variables:	Maximum EVI			Mean EVI	Maximum Cropland EVI			Mean C EVI	ESA C EVI
Model:	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**
Variables									
Downstream × Order = 0	−0.0059***	−0.0065***	−0.0079***	−0.0048***	−0.0095***	−0.0104***	−0.0109***	−0.0073***	−0.0048*
	(0.0013)	(0.0013)	(0.0014)	(0.0009)	(0.0020)	(0.0020)	(0.0021)	(0.0013)	(0.0026)
Downstream × Order = 1	−0.0057***	−0.0060***	−0.0066***	−0.0035***	−0.0061**	−0.0062**	−0.0064***	−0.0043**	−0.0035
	(0.0017)	(0.0016)	(0.0017)	(0.0011)	(0.0026)	(0.0025)	(0.0025)	(0.0017)	(0.0032)
Downstream × Order = 2	−0.0066***	−0.0064***	−0.0067***	−0.0038***	−0.0062**	−0.0058**	−0.0064**	−0.0055***	−0.0015
	(0.0021)	(0.0020)	(0.0020)	(0.0013)	(0.0030)	(0.0029)	(0.0028)	(0.0019)	(0.0035)
Fixed-effects									
Year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mine	Yes			Yes	Yes			Yes	Yes
Pfaffstetter basin level 8		Yes				Yes			
Pfaffstetter basin level 6			Yes				Yes		
Fit statistics									
Observations	114,496	114,496	114,496	114,496	94,604	94,604	94,604	94,604	67,649
R^2	0.92395	0.91954	0.90419	0.95707	0.78597	0.77061	0.74193	0.88641	0.80154
Within R^2	0.05582	0.06500	0.08647	0.11783	0.02531	0.02957	0.04285	0.04478	0.02553

Clustered (Mine) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Appendix · Placebo Outcomes

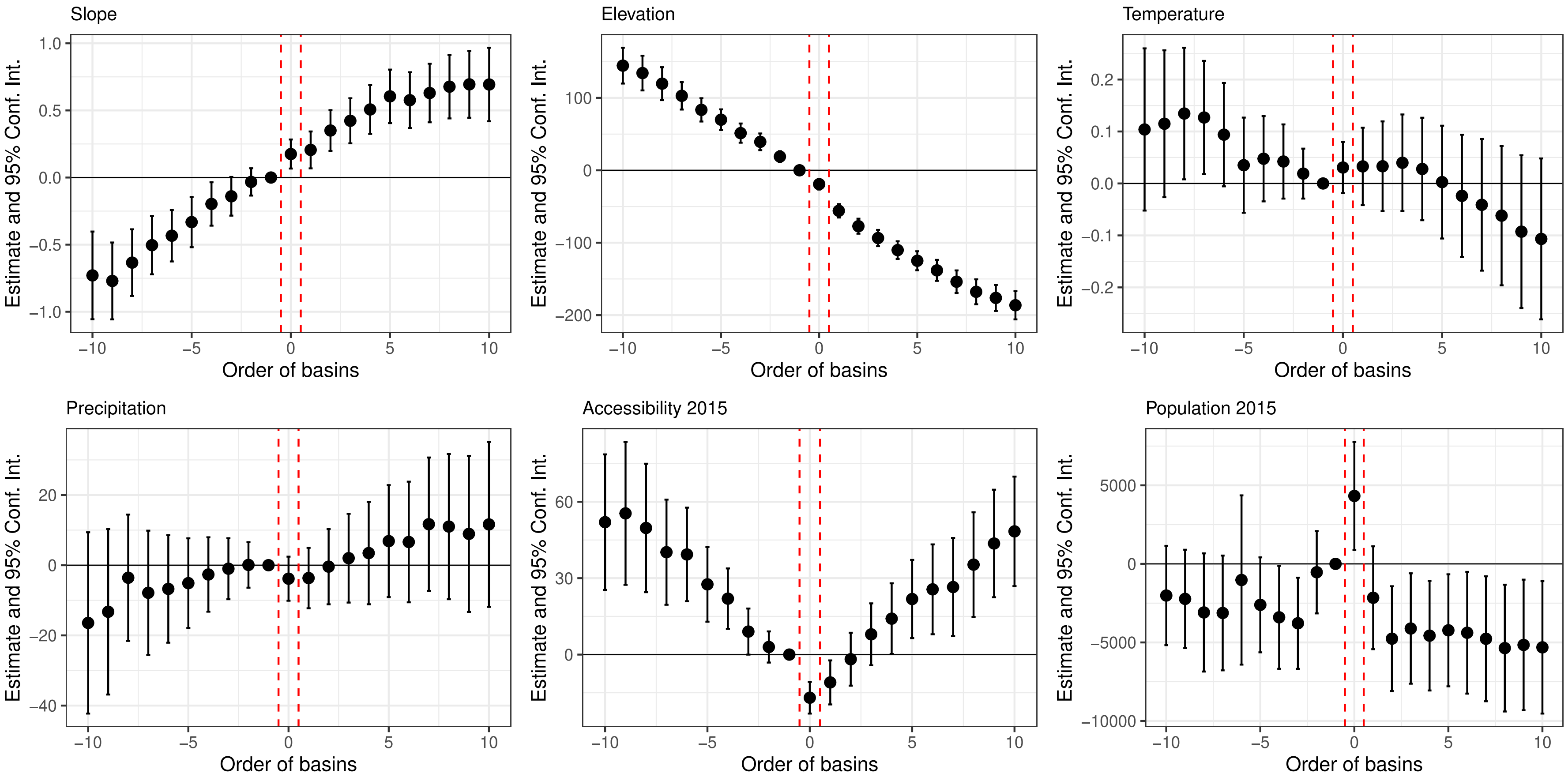
Dependent Variables:	Elevation	Slope	Max. Temp	Precipitation	Accessibility 2015	Population 2015
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Variables						
Downstream	−6.852	−0.0538	−0.0137	0.6025	−5.427	2,125.7
	(8.509)	(0.0912)	(0.0567)	(3.934)	(5.531)	(1,589.8)
Distance × Downstream	−5.008***	−0.0060	0.0135***	−0.1942	0.0839	−182.9***
	(0.4814)	(0.0044)	(0.0036)	(0.2860)	(0.3278)	(55.80)
Distance^2 × Downstream	0.0043	-8.25×10^{-6}	2.12×10^{-6}	0.0003	0.0004	1.081***
	(0.0039)	(4.01×10^{-5})	(3.36×10^{-5})	(0.0020)	(0.0028)	(0.3463)
Distance	2.326***	0.0025	−0.0067**	0.0879	0.7557***	−54.72
	(0.4215)	(0.0039)	(0.0032)	(0.2129)	(0.2587)	(45.17)
Distance^2	0.0005	1.12×10^{-6}	-5.34×10^{-6}	−0.0005	−0.0013	0.3439
	(0.0033)	(3.49×10^{-5})	(3.1×10^{-5})	(0.0015)	(0.0021)	(0.2724)
Fixed-effects						
Year	Yes	Yes	Yes	Yes	Yes	Yes
Mine	Yes	Yes	Yes	Yes	Yes	Yes
Fit statistics						
Observations	114,616	114,616	114,616	114,616	114,576	114,536
R^2	0.95627	0.70192	0.95579	0.96187	0.88768	0.59121
Within R^2	0.41042	0.01108	0.07605	0.00070	0.04659	0.00851

Clustered (Mine) standard-errors in parentheses

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

• go back to identification • go back to results

Appendix · Placebo Outcomes

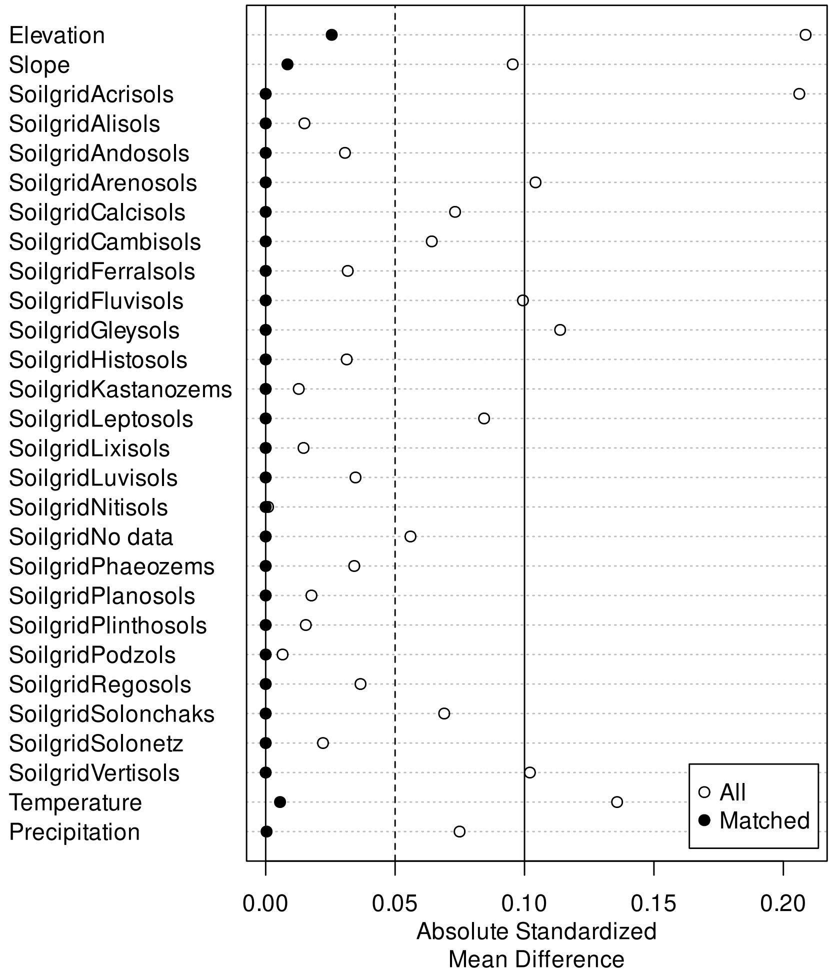


Order estimates when using elevation, slope, temperature, precipitation, accessibility to cities, and population as placebo outcomes.

• go back to identification

• go back to results

Appendix · Matching Exercise



Balance of elevation, slope, temperature, and precipitation before and after matching. (Soilgrids not pictured.)

• go back to identification
• go back to results

Appendix · Dist. Spec. w/ Aut. BW Selection (No Controls)

	Max EVI		Max C EVI	
	Conv.	Bias-Corr.	Conv.	Bias-Corr.
No Controls				
Conventional	-0.0050^{***}	-0.0056^{***}	-0.0112^{***}	-0.0116^{***}
	(0.0015)	(0.0015)	(0.0020)	(0.0025)
Observations	37,880	37,880	32,813	32,813
Bandwidth (Conv)	20.3	20.3	20.7	20.7
Bandwidth (Bias)	46.4	46.4	47.4	47.4

Note: Table shows results for the estimation of the main specification, where distance in kilometres along the river network is the running variable. Bandwidths are chosen automatically following Cattaneo et al. (2019), using a triangular kernel, the mean-squared-error criterion, and bias correction. The upper-panel models include no covariates; the lower-panel models include the full set of controls. Columns (1) and (2) use overall EVI as the outcome, while columns (3) and (4) use cropland-specific EVI. Columns (1) and (3) fit a linear polynomial in distance on each side of the cutoff; columns (2) and (4) fit a quadratic polynomial. All specifications include mine and year fixed effects. Standard errors are clustered at the mine-basin system level.

Significance codes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$ · Clustered (Mine) standard errors in parentheses.

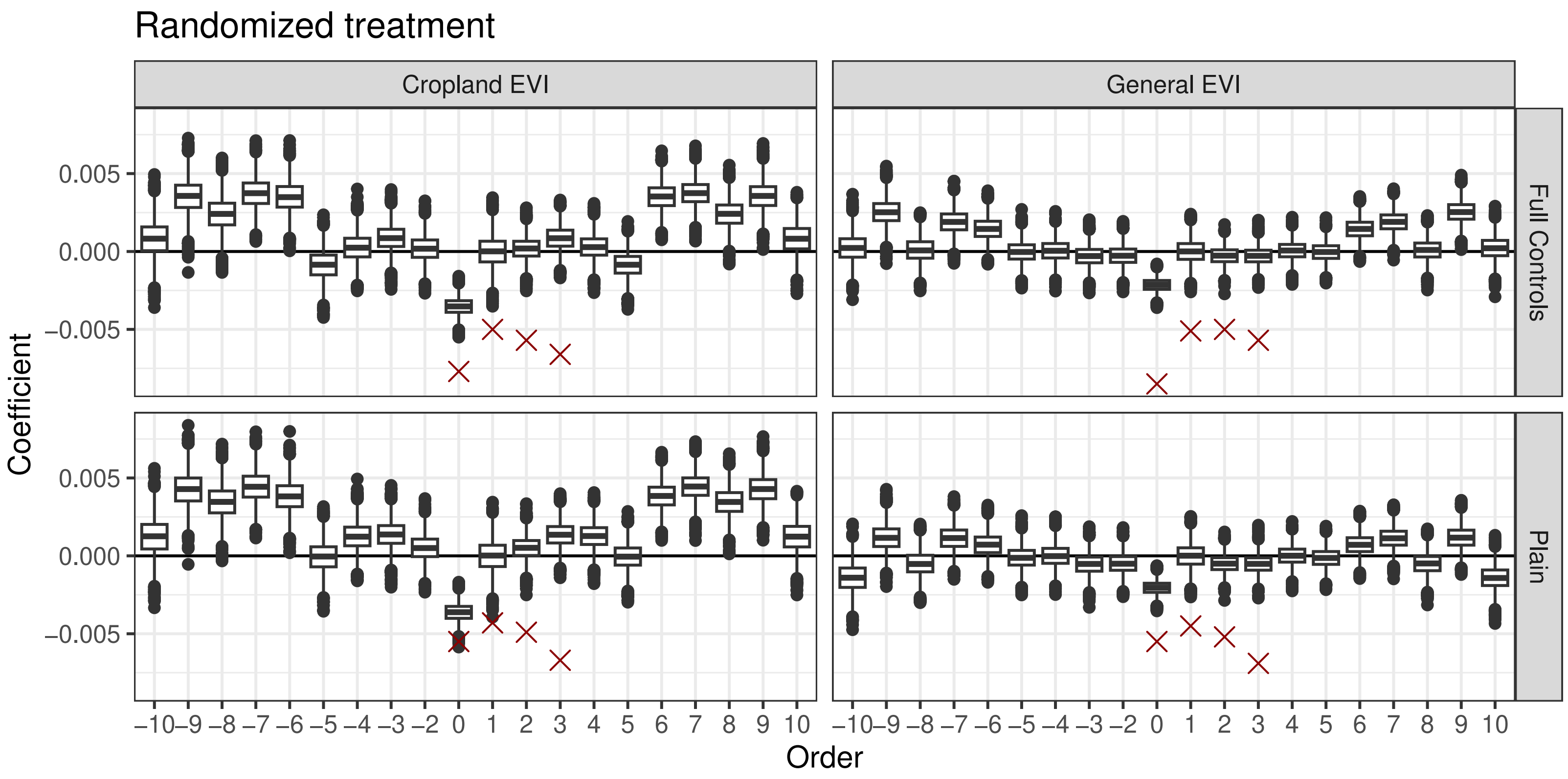
Appendix · Dist. Spec. w/ Aut. BW Selection (Full Controls)

	Max EVI		Max C EVI	
	Conv.	Bias-Corr.	Conv.	Bias-Corr.
With Full Controls				
Conventional	-0.0045^{***}	-0.0049^{***}	-0.0100^{***}	-0.0118^{***}
	(0.0015)	(0.0015)	(0.0020)	(0.0026)
Observations	38,200	38,200	32,629	32,629
Bandwidth (Conv)	20.6	20.6	20.5	20.5
Bandwidth (Bias)	43.4	43.4	45.4	45.4

Note: Table shows results for the estimation of the main specification, with distance (in kilometres along the river network) as the running variable. Bandwidths are selected automatically following Cattaneo et al. (2019), using a triangular kernel, the mean-squared-error criterion, and bias correction. Models in the upper panel include no covariates; models in the lower panel include the full set of controls. Columns (1) and (2) report results for overall EVI, while columns (3) and (4) report results for cropland-specific EVI. Columns (1) and (3) fit a linear polynomial in distance on each side of the cutoff; columns (2) and (4) fit a quadratic polynomial. All specifications include mine and year fixed effects. Standard errors are clustered at the mine-basin system level.

Significance codes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$ · Clustered (Mine) standard errors in parentheses.

Appendix · Permutation – Robustness



Estimation results when the treatment status (i.e., whether basins are down- or upstream) is randomized (5,000 runs, balance between statuses is kept). The red crosses indicate estimation results for the main specification.

• go back to identification

• go back to results

Appendix · Ord. Spec. w/ Aut. BW Selection (Full Controls)

	Max EVI		Max C EVI	
	No Cluster	**Cluster (Mine Basin)**	**No Cluster**	**Cluster (Mine Basin)**
No Controls				
I(order > 0)	−0.0048	−0.0048	-0.0090^{***}	-0.0090^{**}
	(0.0013)	(0.0019)	(0.0018)	(0.0030)
Observations	45,613	45,613	38,537	38,537
Bandwidth	2	2	2	2

Note: Table shows results for the estimation of the main specification, with distance defined by basin order relative to the mine basin as the running variable. Bandwidths are selected automatically following Kolesár & Rothe (2018), using a triangular kernel, the mean-squared-error criterion, and no bias correction. Models in the upper panel include no covariates; models in the lower panel include the full set of controls. Columns (1) and (2) report results for overall EVI, while columns (3) and (4) report results for cropland-specific EVI. Columns (1) and (3) use unclustered standard errors; columns (2) and (4) cluster standard errors at the mine-basin system level.
All specifications include mine and year fixed effects.

Significance codes: *** p < 0.01, ** p < 0.05, * p < 0.1

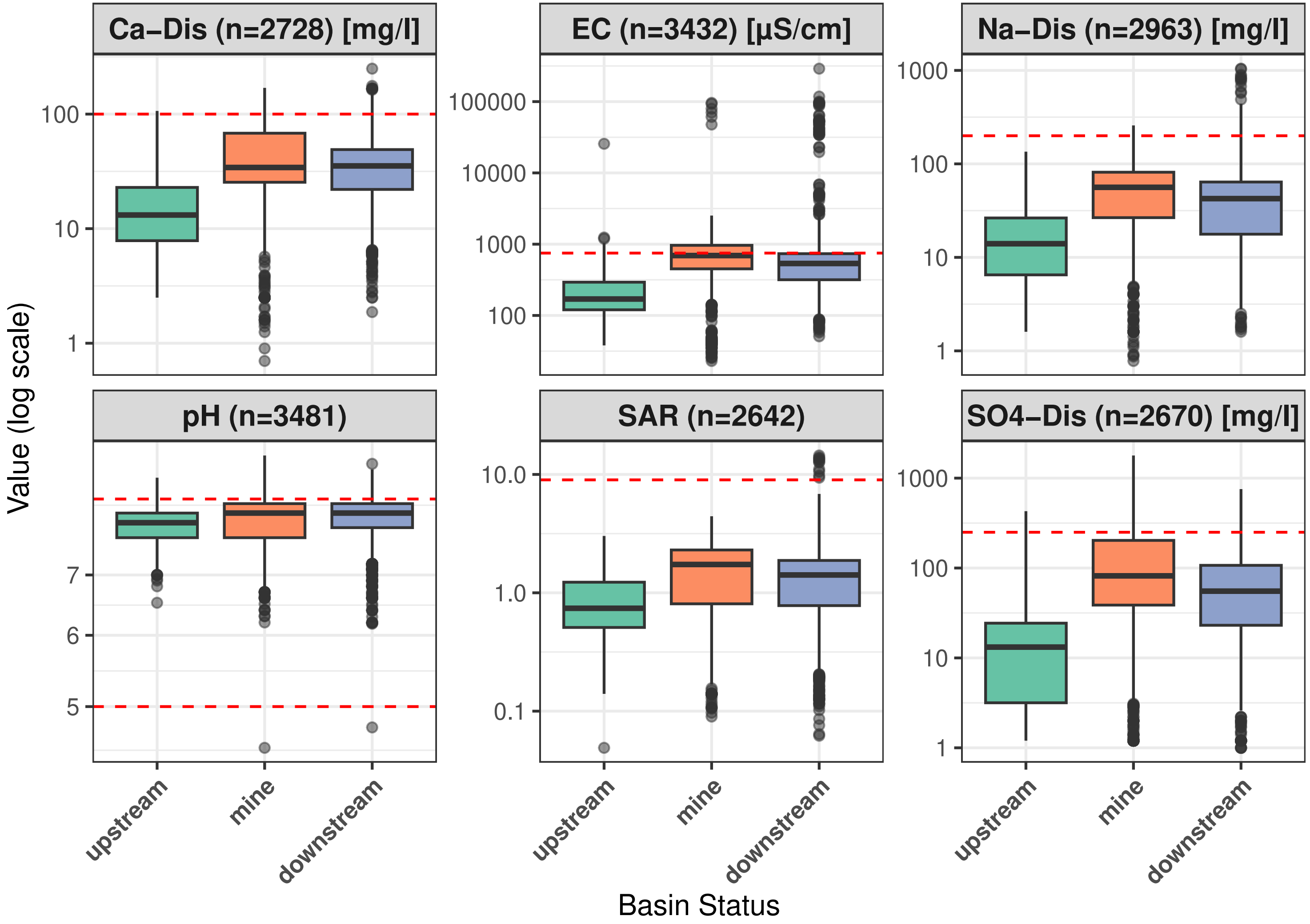
Appendix · Validation of Outcome IFPRI

Outcome:	ln(Crops, Value)	ln(Crops, FY)	ln(Cereals, Value)	ln(Cereals, Yield)	ln(Cereals, FY)
Model:	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**
Variables					
Max. Cropland EVI	3.398***	0.9519***	2.489***	0.8995***	0.5589**
	(0.4230)	(0.1828)	(0.9150)	(0.1586)	(0.2704)
Fixed effects					
Wave	Yes	Yes	Yes	Yes	Yes
Fit statistics					
Observations	44,682	44,380	44,682	44,682	44,171
R^2	0.65336	0.35656	0.50120	0.60944	0.32956
Within R^2	0.08225	0.00717	0.02177	0.02195	0.00153

*Clustered (wave) standard-errors in parentheses Significance: ***: 0.01, **: 0.05, *: 0.1*

• go back

Appendix · Water Pollution Data for South Africa



• go back