Mines → Rivers → Yields

Downstream Mining Impacts on Agriculture in Africa

Lukas Vashold (lukas.vashold@wu.ac.at)
Gustav Pirich (gustav.pirich@wu.ac.at
Max Heinze (maximilian.heinze@wu.ac.at)
Nikolas Kuschnig (nikolas.kuschnig@wu.ac.at)

Vienna Development Workshop

June 24, 2025

Background

Methods and 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** (von der Goltz & Barnwal, 2019).

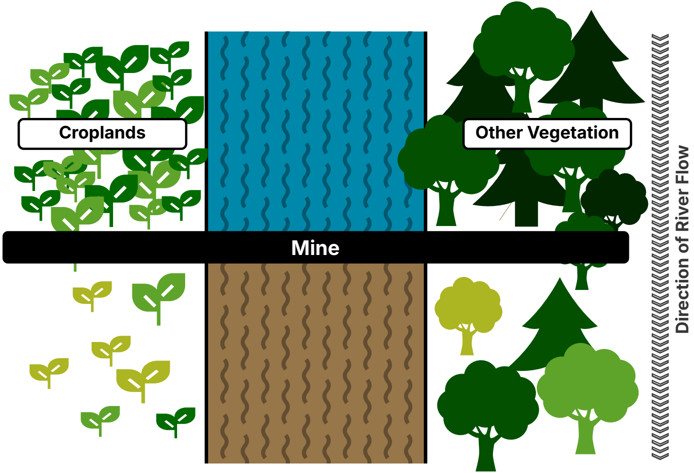
A Curse?

- Resource extraction causes **negative externalities.**
- **Ecological effects** include:
  - Mines **use water** and produce **sediments and tailings** (Moura et al., 2022).
  - Pollutants include **mercury**, **sodium cyanide** 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 this.



Research Question

What is the causal effect of mining-induced water pollution on vegetation and agricultural productivity in Africa?

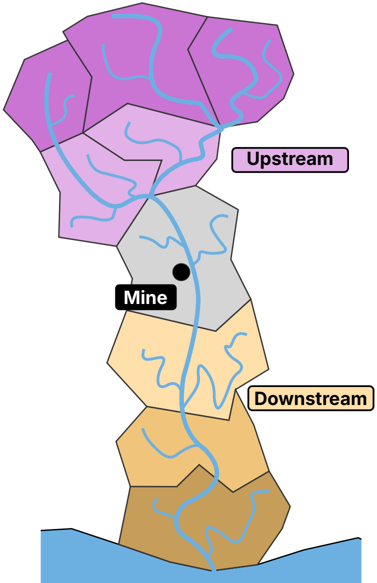
- We study this question in the context of Africa because
  - it has a **booming mining industry** (International Council on Mining and Metals [ICMM], 2022)
  - with many **artisanal and small-scale mines** (ASM Inventory, 2022; Girard et al., 2022)
  - and a lack of oversight and **containment facilities** (Kossoff et al., 2014; Macklin et al., 2023).

Background

Methods and Data

Results

How to Find Affected Areas



Our unit of observation is the **river basin**. Lehner and 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 basins **upstream** to a mine to those **downstream**.

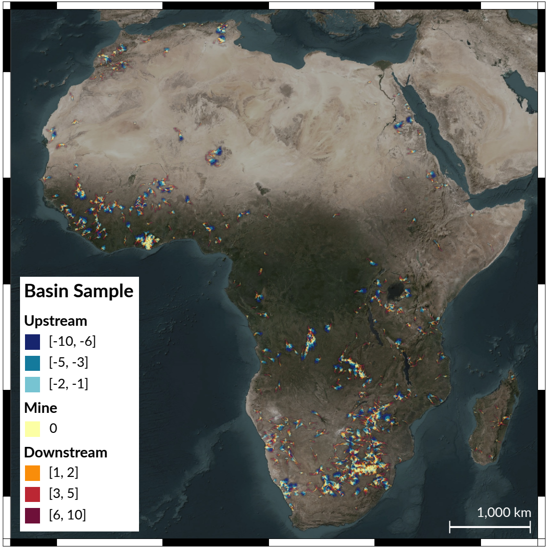
show more on basins

Mines

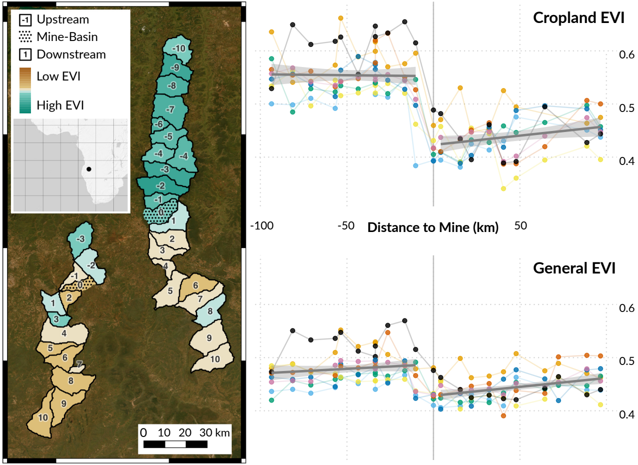
We use **mine locations** from Maus et al.'s (2022) dataset, which includes both industrial and artisanal mining sites.

We designate **mine basins** and up to ten levels of **upstream** and **downstream** basins.

more maps



Intuition



Example of two Angolan mine basin systems with EVI measurements for croplands and general vegetation over years 2016, 2017, 2018, 2019, 2020, 2021, 2022, and 2023.

Variables and Observations

Outcome

- We use the **Enhanced Vegetation Index (EVI)**, which
  - is **remotely sensed**, and
  - ranges **between −1** (water) **and 1** (dense vegetation). Details

annual max. on all areas covered by vegetation	annual max. on areas covered by croplands
↓ **Max. EVI**	↓ **Max. Cropl. EVI**

Observations and Covariates

6,307 upstream basins	**1,900 mine** basins	**6,127 downstream** basins

- We observe the basins for $T = 8$ years. show order × status
- We collate **covariates** on topography (Amatulli et al., 2018), soil type (Hengl et al., 2017), climate (Abatzoglou et al., 2018), and socioeconomic characteristics. (Weiss et al., 2018)

Empirical Strategy (Spatial RDD)

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

where we let $F(\cdot)$ return indicators:

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

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

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

Background

Methods and Data

Results

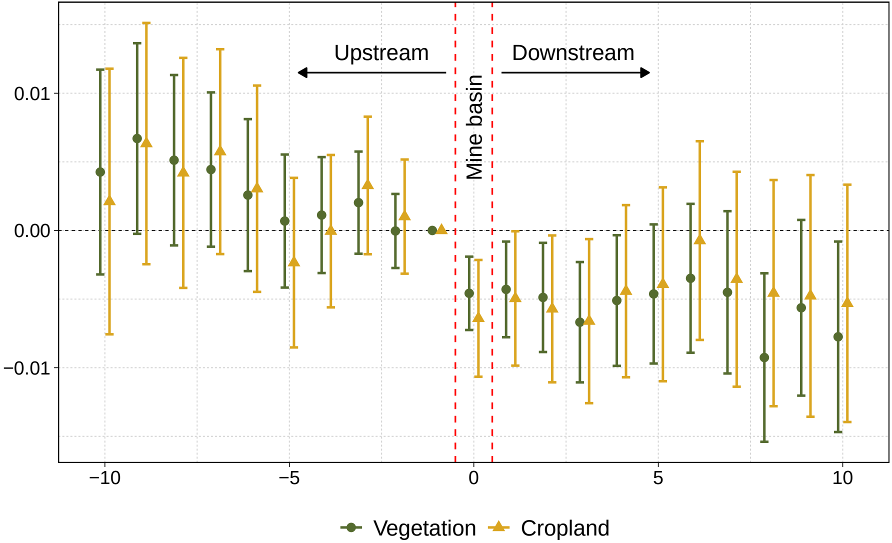
Results Overview

We find a **significant reduction** of vegetation health **downstream** of mines.

Impacts are particularly strong in **fertile regions** and where **gold mining** predominates.

These results are **robust** to varying the outcome measure, the sample, and the level of fixed effects.

Results: Figure



Results: Table

Peak vegetation is reduced by **1.28–1.35%** relative to the mean on a total affected area of **255,000 km²**.

Peak cropland vegetation is reduced by **1.38–1.47%** on a total affected area of **74,000 km²**.

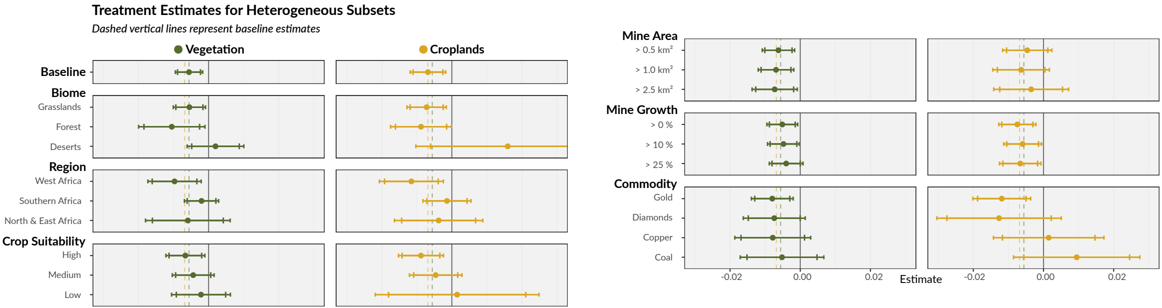
Outcome	*Peak Vegetation*		*Peak Cropland Veg.*	
(Specification)	(Plain)	(Full)	(Plain)	(Full)
Individual Order				
Downstream (1st)	-0.0045***	-0.0043**	-0.0051**	-0.0050**
	(0.0017)	(0.0018)	(0.0025)	(0.0025)
Downstream (2nd)	-0.0049**	-0.0048**	-0.0058*	-0.0067**
	(0.0022)	(0.0024)	(0.0031)	(0.0032)
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)
Pooled Order				
Downstream (1st–3rd)	-0.0057***	-0.0056***	-0.0064**	-0.0068***
	(0.0018)	(0.0020)	(0.0025)	(0.0026)
Controls	No	Yes	No	Yes
Mine and year FEs	Yes	Yes	Yes	Yes
Observations	110,576	110,524	93,036	93,000

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 (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 **vaule** of **overall crop production.**
- This amounts to a reduction in **agricultural production** of about **91,000 metric tons of cereals**,
- comparable to **5.4%** of the 1.7 million tons thath the **World Food Program (WFP)** distributes annually.

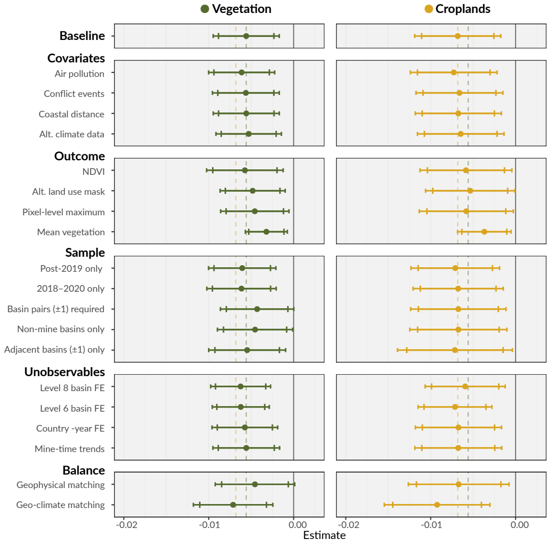
Heterogeneity – Which areas are affected?



Robustness

- Battery of **robustness checks** along several dimensions:
  - Additional covariates
  - Varying the outcome variable
  - Varying the sample specification
  - Varying fixed effect levels
  - Using continuous distance
- Additional **validity checks** of identifying assumptions:

permutation | covariates as placebo outcomes | matching on covariates | continuous distance



Discussion & Limitations

- **Water pollution** is a convincing mediator, but we have **no direct evidence**.
  - Water quality data is sparse, especially in Africa (Jones et al., 2024).
  - We collect data on water pollution (United Nations Environment Programme, 2025), only for South Africa.
  - We find elevated pollutant levels in mine and downstream basins. show more
- Noise from multiple data sources likely leads to **attenuation of estimates**.
  - EVI captures crop yields only indirectly, and
  - we cannot reliably estimate how far effects reach using this research design.
- We cannot disentangle **artisanal and industrial mining**.
- We do not directly observe **farmers' adaptive responses** (e.g., migrating upstream).

Conclusion

We identified the causal effects of mining

- on vegetation and agriculture,
- mediated by water pollution.

Our results show **negative impact** on vegetation and crop 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 (1)

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 1958–2015. *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

References (2)

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, November). *A practical introduction to regression discontinuity designs: Foundations*. Cambridge University Press. https://doi.org/10.1017/9781108684606

Digital Earth Africa. (2022). *Cropland extent maps for Africa* (tech. rep.). (Dataset). DE Africa Services. https://docs.digitalearthafrica.org/en/latest/data_specs/Cropland_extent_specs.html

Girard, V., Molina-Millán, T., Vic, G., et al. (2022). Artisanal mining in Africa. In *Working Paper Series No. 2201*. Novafrica.

References (3)

Hengl, T., de Jesus, J. M., Heuvelink, G. B. M., Gonzalez, M. R., Kilibarda, M., Blagotić, A., Shangguan, W., Wright, M. N., Geng, X., Bauer-Marschallinger, B., Guevara, M. A., Vargas, R., MacMillan, R. A., Batjes, N. H., Leenaars, J. G. B., Ribeiro, E., Wheeler, I., Mantel, S., & Kempen, B. (2017).SoilGrids250m: Global gridded soil information based on machine learning. *PLOS ONE*, *12*(2), e0169748. https://doi.org/10.1371/journal.pone.0169748

Hund, K., La Porta, D., Fabregas, T. P., Laing, T., & Drexhage, J. (2023). *Minerals for climate action: The mineral intensity of the clean energy transition* (tech. rep.). World Bank Group. Washington DC, United States. https://pubdocs.worldbank.org/en/961711588875536384/Minerals-for-Climate-Action-The-Mineral-Intensity-of-the-Clean-Energy-Transition.pdf

IFPRI. (2020). *AReNA's DHS-GIS database*. https://doi.org/10.7910/DVN/OQIPRW

International Council on Mining and Metals. (2022). Mining Contribution Index (MCI) (6th Edition). https://www.icmm.com/website/publications/pdfs/social-performance/2022/research_mci-6-ed.pdf?cb=16134

References (4)

Johnson, D. M. (2016). A comprehensive assessment of the correlations between field crop yields and commonly used MODIS products. *International Journal of Applied Earth Observation and Geoinformation*, *52*, 65–81. https://doi.org/10.1016/j.jag.2016.05.010

Jones, E. R., Graham, D. J., van Griensven, A., Flörke, M., & van Vliet, M. T. H. (2024). Blind spots in global water quality monitoring. *Environmental Research Letters*, *19*(9), 091001. https://doi.org/10.1088/1748-9326/ad6919

Kolesár, M., & Rothe, C. (2018). Inference in regression discontinuity designs with a discrete running variable. *American Economic Review*, *108*(8), 2277–2304. https://doi.org/10.1257/aer.20160945

Kossoff, D., Dubbin, W. E., Alfredsson, M., Edwards, S. J., Macklin, M. G., & Hudson-Edwards, K. A. (2014). Mine tailings dams: Characteristics, failure, environmental impacts, and remediation. *Applied Geochemistry*, *51*, 229–245. https://doi.org/10.1016/j.apgeochem.2014.09.010

References (5)

Lehner, B., & Grill, G. (2013).Global river hydrography and network routing: Baseline data and new approaches to study the world's large river systems. *Hydrological Processes*, *27*(15), 2171–2186. https://doi.org/10.1002/hyp.9740

Macklin, M. G., Thomas, C. J., Mudbhatkal, A., Brewer, P. A., Hudson-Edwards, K. A., Lewin, J., Scussolini, P., Eilander, D., Lechner, A., Owen, J., Bird, G., Kemp, D., & Mangalaa, K. R. (2023).Impacts of metal mining on river systems: A global assessment. *Science*, *381*(6664), 1345–1350. https://doi.org/10.1126/science.adg6704

Maus, V., Giljum, S., da Silva, D. M., Gutschlhofer, J., da Rosa, R. P., Luckeneder, S., Gass, S. L., Lieber, M., & McCallum, I. (2022).An update on global mining land use. *Scientific data*, *9*(1), 1–11.

Moura, A., Lutter, S., Siefert, C. A. C., Netto, N. D., Nascimento, J. A. S., & Castro, F. (2022).Estimating water input in the mining industry in Brazil: A methodological proposal in a data-scarce context. *Extractive Industries and Society*, *9*, 101015. https://doi.org/10.1016/j.exis.2021.101015

References (6)

Ruppen, D., Runnalls, J., Tshimanga, R. M., Wehrli, B., & Odermatt, D. (2023). Optical remote sensing of large-scale water pollution in Angola and DR Congo caused by the Catoca mine tailings spill. *International Journal of Applied Earth Observation and Geoinformation*, *118*, 103237. https://doi.org/10.1016/j.jag.2023.103237

Schwarzenbach, R. P., Egli, T., Hofstetter, T. B., von Gunten, U., & Wehrli, B. (2010). Global water pollution and human health. *Annual Review of Environment and Resources*, (Volume 35, 2010), 109–136. https://doi.org/10.1146/annurev-environ-100809-125342

United Nations Environment Programme. (2025). GEMStat database of the global environment monitoring system for freshwater (GEMS/Water) Programme [Accessed 21 February 2025. Available upon request from GEMS/Water Data Centre: `gemstat.org`].

von der Goltz, J., & Barnwal, P. (2019). Mines: The local wealth and health effects of mineral mining in developing countries. *Journal of Development Economics*, *139*, 1–16. https://doi.org/10.1016/j.jdeveco.2018.05.005

References (7)

Weiss, D. J., Nelson, A., Gibson, H. S., Temperley, W., Peedell, S., Lieber, A., Hancher, M., Poyart, E., Belchior, S., Fullman, N., Mappin, B., Dalrymple, U., Rozier, J., Lucas, T. C. D., Howes, R. E., Tusting, L. S., Kang, S. Y., Cameron, E., Bisanzio, D., ... Gething, P. W. (2018). A global map of travel time to cities to assess inequalities in accessibility in 2015. *Nature*, *553*, 333–336. https://doi.org/10.1038/nature25181

Appendix

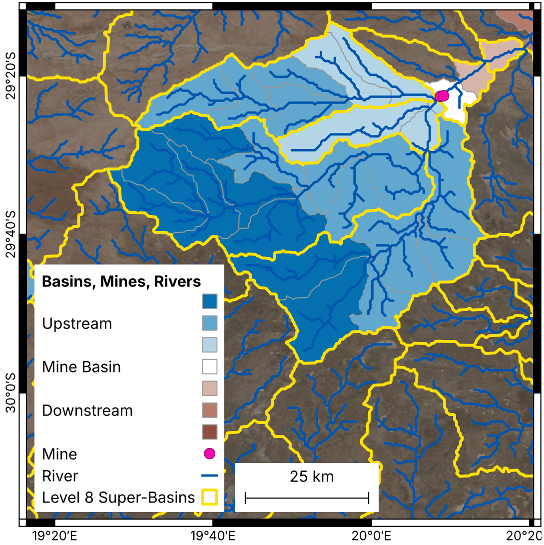
Our **unit of observation** is the **river basin**.

Lehner and 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**.

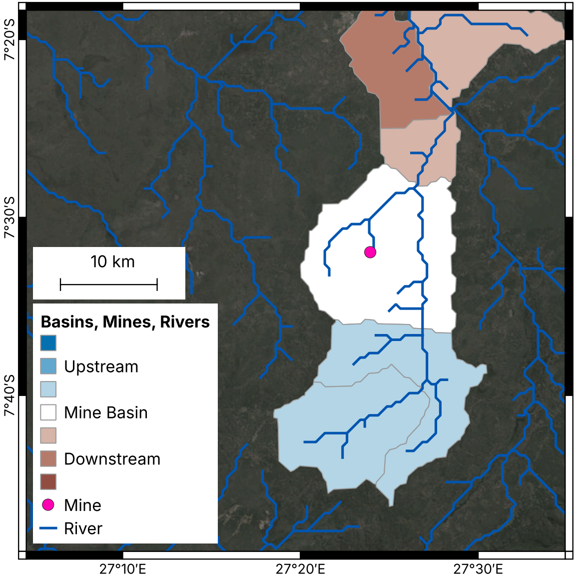
Illustration from Lehner and Grill (2013)

Appendix Forked Upstream Basins

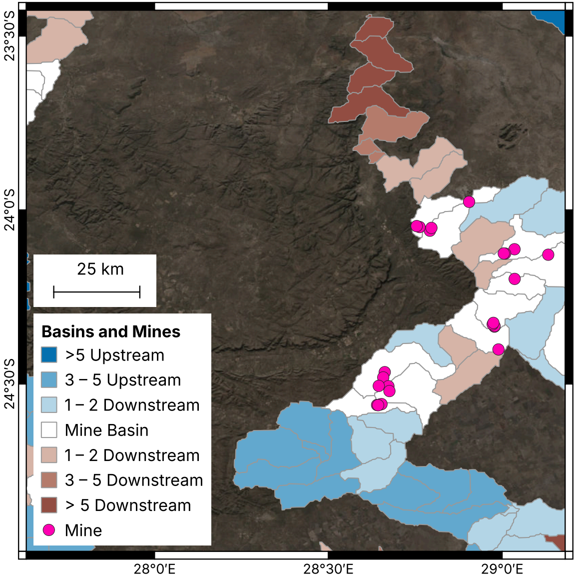


Multiple forked upstream (Level 12) basins join into a single mine basin further downstream. The superimposed yellow lines indicate Level 8 basins; these contain varying numbers of sub-basins (due to a level-skipping mechanism) and clearly divide tributary and main basins. The blue lines, which represent river streams, provide additional intuition for the basin topology.

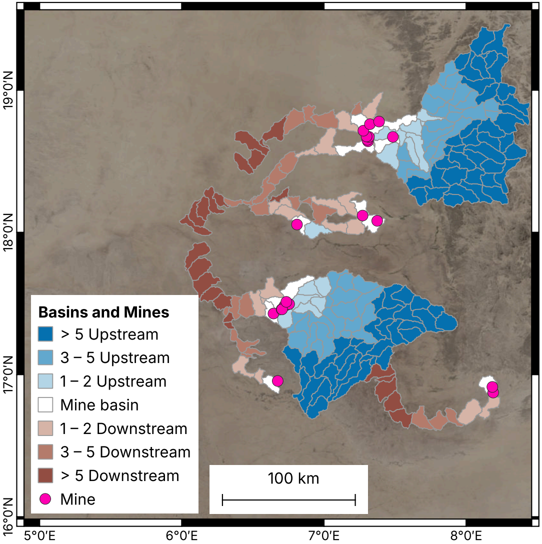
Appendix Mine in Basin



Appendix Mine Cluster



Appendix Multiple Basin Chains



go back

Research Design

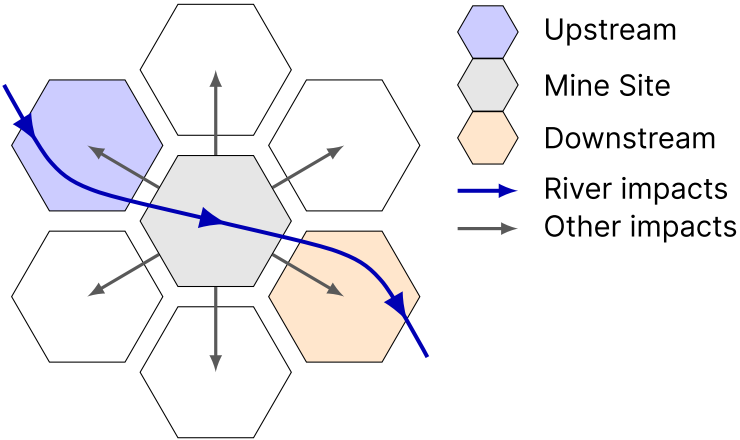


Illustration of the research design. The comparison of up- and downstream basins enables the identification of mine impacts that are mediated by the river.

A Proxy for Agricultural Activity

go back

- 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** (Digital Earth Africa, 2022). **→ 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).

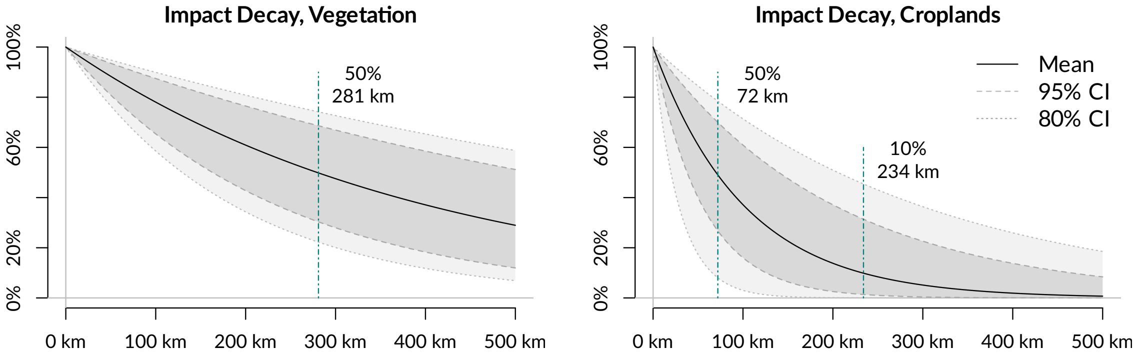
Appendix Summary Statistics go back

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

show balance

Impact Decay Assessment

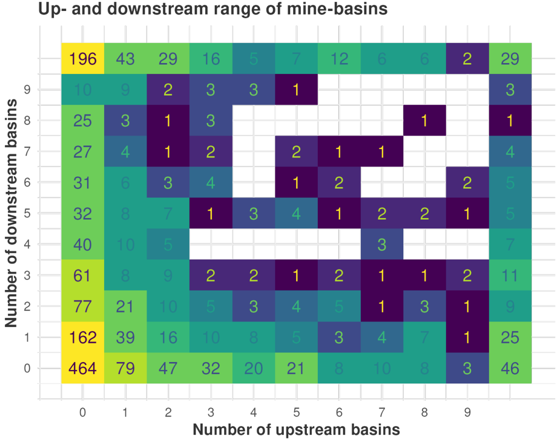
To gauge effect reach, we re-estimate using an **exponential distance decay** function, $exp(-\delta d_{ij})$, where d_{ij} is the distance from the mine along the river



Impact Decay Assessment

go back

- 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 not known, 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.



Number of mine-basins with *Y* upstream and *X* downstream basins in the dataset.

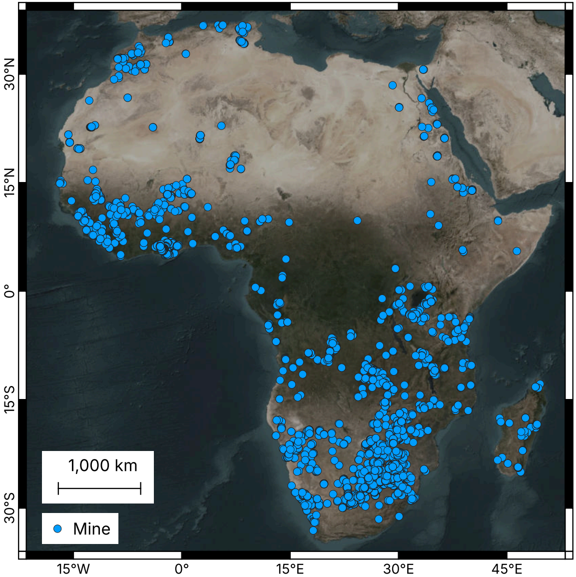
go back

Appendix Basins by Order

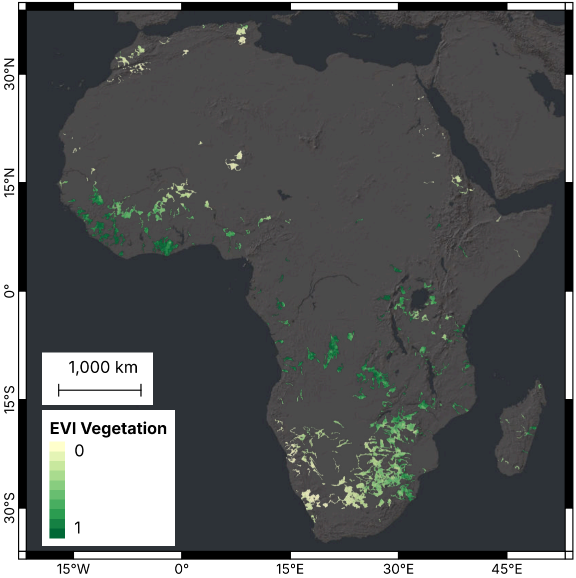
Order	Upstream		Downstream	
	N	Distance	*N*	Distance
0	(1900)	(0.0)	(1900)	(0.0)
1	847	13.9	1162	11.1
2	781	24.5	882	22.0
3	722	35.0	743	32.7
4	698	44.9	643	43.3
5	653	55.3	578	54.0
6	576	66.3	512	64.3
7	562	75.8	458	74.1
8	522	86.5	416	84.4
9	494	95.8	382	95.0
10	452	104.2	351	104.7

go back

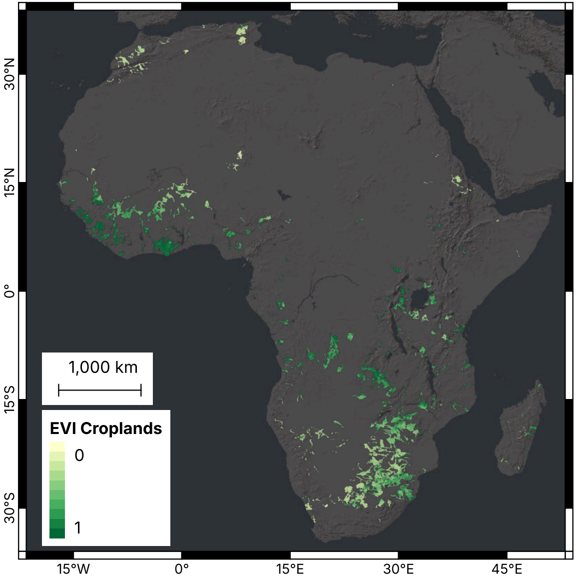
Appendix Distribution of Mines



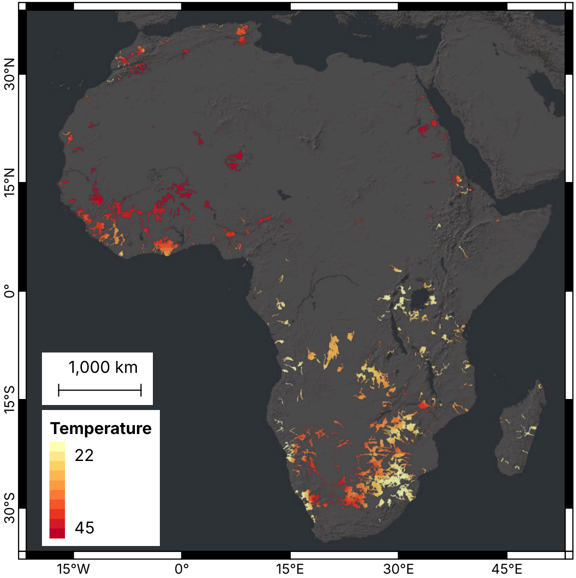
Appendix EVI Map



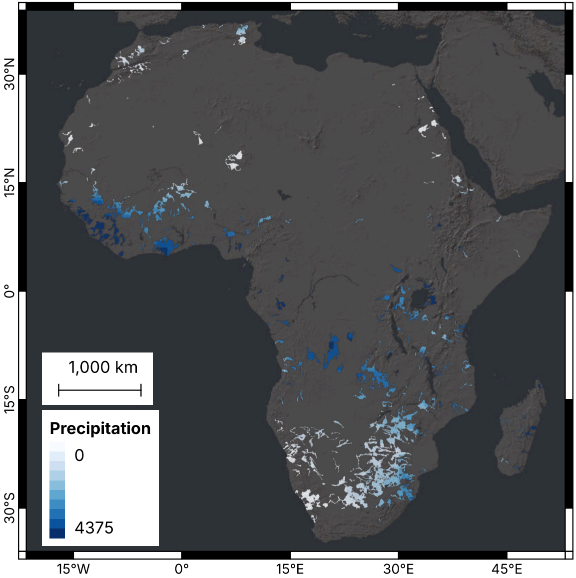
Appendix Croplands EVI Map



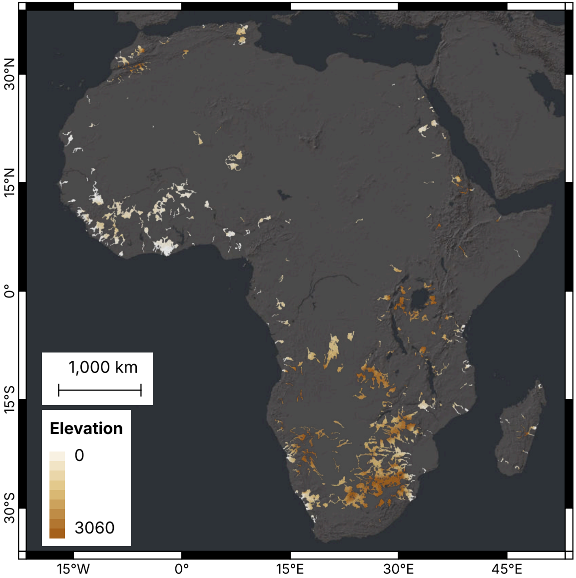
Appendix Temperature Map



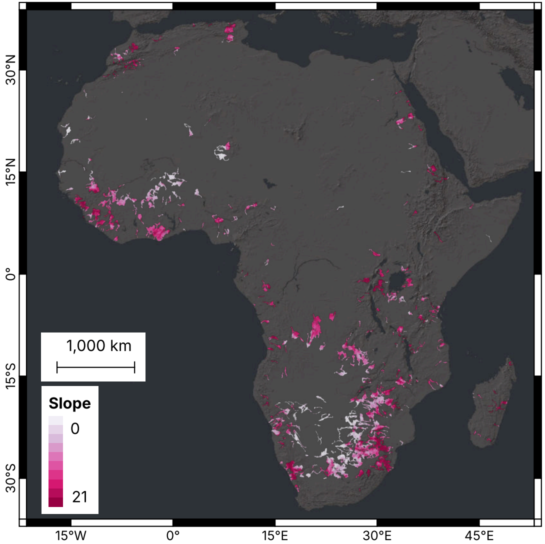
Appendix Precipitation Map



Appendix Elevation Map

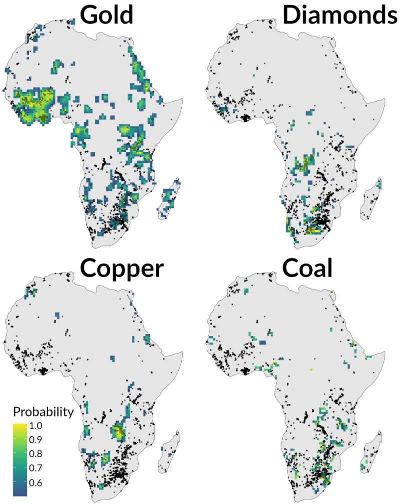


Appendix Slope Map



go back

Appendix Commodity Type Prediction



Appendix Full Order Specification Results

Outcome	*Peak Vegetation*		*Peak Cropland Veg.*	
(Specification)	(Plain)	(Full)	(Plain)	(Full)
Individual Order				
Downstream (1^{st})	-0.0045***	-0.0043**	-0.0051**	-0.0050**
	(0.0017)	(0.0018)	(0.0025)	(0.0025)
Downstream (2^{nd})	-0.0049**	-0.0048**	-0.0058*	-0.0067**
	(0.0022)	(0.0024)	(0.0031)	(0.0032)
Downstream (3^{rd})	-0.0085***	-0.0087***	-0.0088**	-0.0099***
	(0.0028)	(0.0029)	(0.0037)	(0.0038)
Downstream (4^{th})	-0.0049*	-0.0062*	-0.0029	-0.0044
	(0.0030)	(0.0033)	(0.0038)	(0.0040)
Downstream (5^{th})	-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 (1^{st}–3^{rd})	-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*

go back

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 $-\delta$ × 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*

go back

Appendix Varying Sample Definition

Dependent Variables:	Maximum EVI					Maximum Cropland EVI				
Model:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Variables										
Downstream x 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 x 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 x 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*

go back

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 x 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 x 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 x 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 in 2015	Population in 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

Appendix Placebo Outcomes

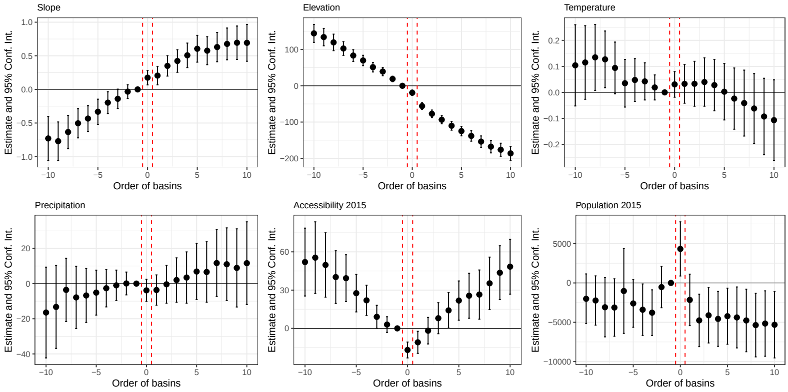
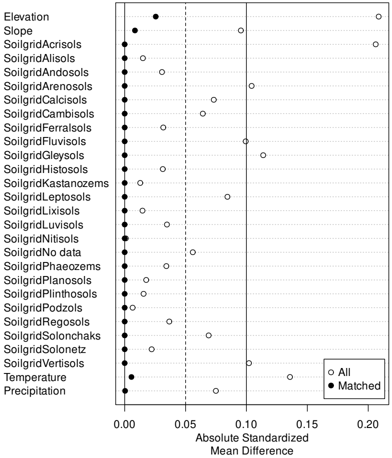


Figure: Order estimates when using elevation, slope, temperature, precipitation, accessibility to cities, and population as placebo outcomes.

Appendix Matching Exercise



Balance of elevation, slope, temperature, and precipitation before and after matching. go back

Appendix Dist. Spec. w/ Aut. Bandwith 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 estimation of 11, with distance as measured in kilometer along the river network used as the running variable, using practices suggested in Cattaneo et al., 2019 for automatic bandwidth selection using a triangular Kernel and the mean squared error distance as selection criterion, and bias correction. Models in the upper panel include no covariates, models in the lower panel include the full set of controls. Models in columns (1) and (2) report results using the overall EVI as outcome, models in columns (3) and (4) for the cropland-specific EVI. Models (1) and (3) fit a linear polynomial of the distance measure at each side of the cutoff, models in columns (2) and (4) 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.

Dist. Spec. w/ Aut. Bandwith 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 estimation of 11, with distance as measured in kilometer along the river network used as the running variable, using practices suggested in Cattaneo et al., 2019 for automatic bandwidth selection using a triangular Kernel and the mean squared error distance as selection criterion, and bias correction. Models in the upper panel include no covariates, models in the lower panel include the full set of controls. Models in columns (1) and (2) report results using the overall EVI as outcome, models in columns (3) and (4) for the cropland-specific EVI. Models (1) and (3) fit a linear polynomial of the distance measure at each side of the cutoff, models in columns (2) and (4) 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

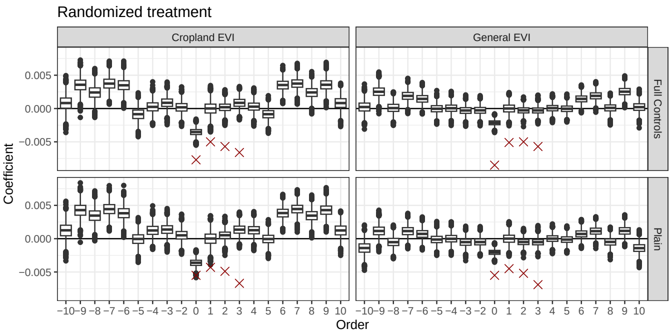


Figure: Estimation results when the treatment status (i.e., whether basins are down- or upstream) is randomized (5,000 runs, balance between statuses is kept). The red crosses indicate estimation results for the main specification.

go back

Appendix Ord. Spec. w/ Aut. Bandwith 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 (0.0013)	-0.0048 (0.0019)	-0.0090*** (0.0018)	-0.0090** (0.0030)
Observations	45,613	45,613	38,537	38,537
Bandwidth	2	2	2	2

Note: Table shows results for estimation of 11, with distance as measured by the ordering of basins with respect to the mine basin as the running variable, using practices suggested in Kolesár and Rothe, 2018 for automatic bandwidth selection using a triangular Kernel and the mean squared error distance as selection criterion. Models in the upper panel include no covariates, models in the lower panel include the full set of controls. Models in columns (1) and (2) report results using the overall EVI as outcome, models in columns (3) and (4) for the cropland-specific EVI. Models (1) and (3) do no cluster standard errors, models in columns (2) and (4) cluster standard errors are 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

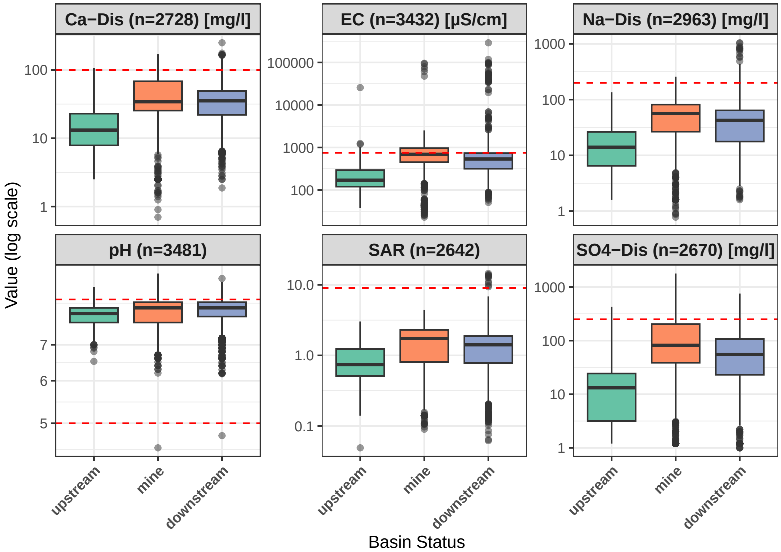
Appendix Validation of Outcome IFPRI

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