

# Re-Evaluating the Effect of Jesuit Missions on Present-Day Human Capital

Max Heinze ([maximilian.heinze@wu.ac.at](mailto:maximilian.heinze@wu.ac.at))

Topics in Macroeconomics · Project Presentation

January 17, 2025

1/31

## Recap

Using BMA to Evaluate the Model Setup

Spatial Scope

Spatial Methods

What Lies Ahead

2/31

## What was this about again?

In his paper, Valencia Caicedo (2019) finds **persistent effects on present-day human capital** in localities where **Jesuits had established a mission** 300 years earlier.

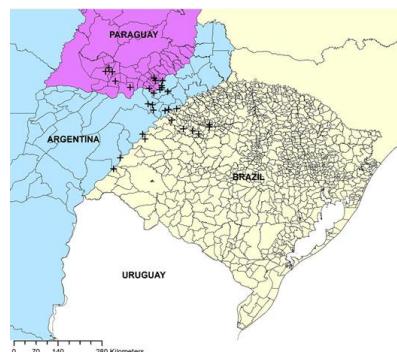


Figure I from Valencia Caicedo (2019)

3/31

### (A Third of) Table II from Valencia Caicedo (2019)

	Illiteracy	
	(1)	(2)
Mission distance	0.0105*** (0.004) {0.004}	0.0112** (0.005) {0.005}
Geo controls	No	Yes
State fixed effects	Yes	Yes
Observations	547	548
Within $R^2$	0.037	0.068
$R^2$	0.042	0.073

Note: Robust standard errors in parentheses and Conley standard errors in curly brackets.

\*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ .

4/31

### What Do I Want to Investigate?

Variables	Space (geographic, not model)
<ul style="list-style-type: none"> <li>The <b>mission distance</b> coefficient is not consistently positive.</li> <li>FVC conducts some robustness checks, but other small changes sometimes lead to the effect changing or vanishing.</li> </ul>	<ul style="list-style-type: none"> <li>Distance is accounted for <b>linearly</b>.</li> <li>The mission distance coefficient <b>becomes insignificant</b> when changing that.</li> <li>It might also make sense to consider, e.g., spatial autocorrelation.</li> </ul>
	5/31

### Varying the Distance Measure: Original Model

	Illiteracy			
	(1)	(2)	(3)	(4)
dismiss	0.011** (0.005)			
dismiss_log		0.078 (0.289)		
dismiss_exp			-3.293 (2.082)	
dismiss_gauss				-1.631 (1.246)
Geo Controls	Yes	Yes	Yes	Yes
Observations	548	548	548	548
$R^2$	0.073	0.063	0.067	0.066
Adjusted $R^2$	0.049	0.038	0.043	0.041
Residual Std. Error (df = 533)	3.912	3.934	3.924	3.927

Note:

\* $p < 0.1$ ; \*\* $p < 0.05$ ; \*\*\* $p < 0.01$

6/31

## Using BMA to Evaluate the Model Setup

Spatial Scope

Spatial Methods

What Lies Ahead

7/31

## How to Deal with Distance

### Distance Measures

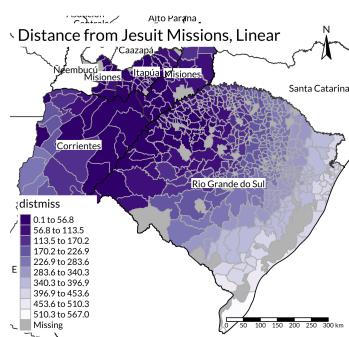
- **Linear distance**, just as in the original paper
- **Exponential distance**, with  $\lambda = \frac{1}{200}$
- **Logarithmic distance**
- **Gaussian distance**, with  $\exp\left(-\frac{d^2}{2 \cdot 100^2}\right)$

### Fixed Inclusion?

- **Naively averaging** over all possible models in the model space yields a **low PIP** for *any* distance measure.
- I fix the **distance measure** and average over that model subspace. Even a non-robust determinant may be causally related to the outcome.

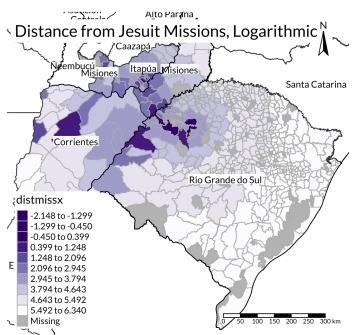
8/31

## Linear Distance



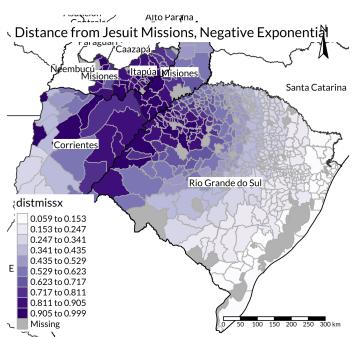
9/31

## Logarithmic Distance



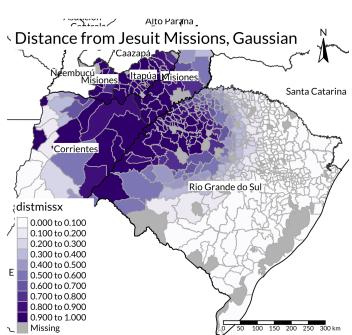
10/31

## Exponential Distance



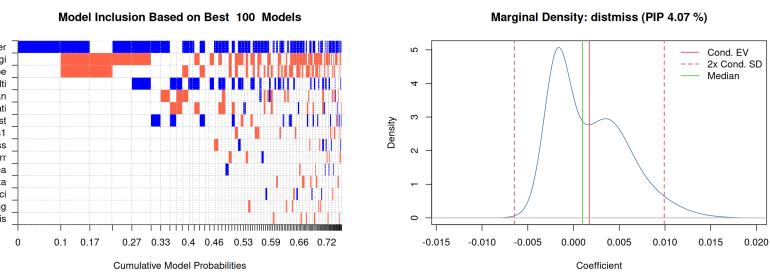
11/31

## Gaussian Distance



12/31

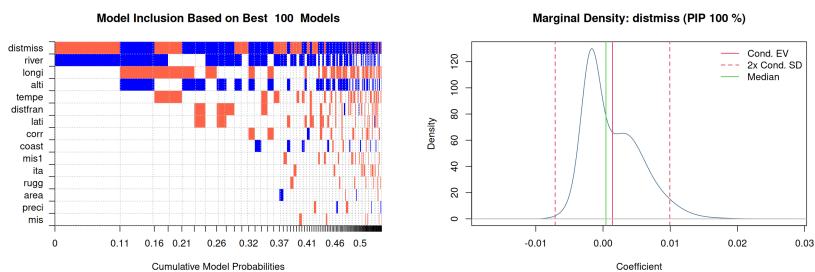
## Linear Distance



[show all tables and figures](#)

13/31

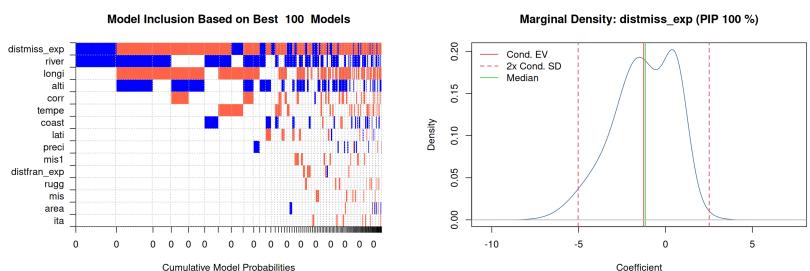
## Linear Distance, Fixed Inclusion



[show all tables and figures](#)

14/31

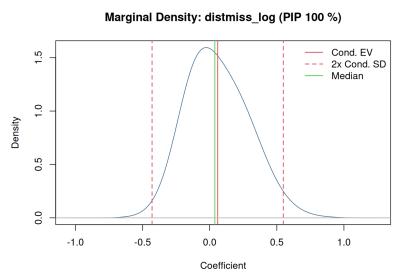
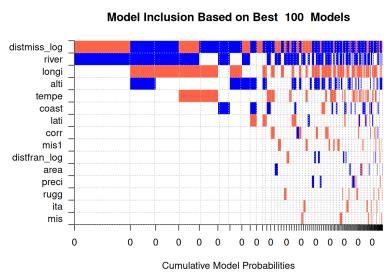
## Exponential Distance, Fixed Inclusion



[show all tables and figures](#)

15/31

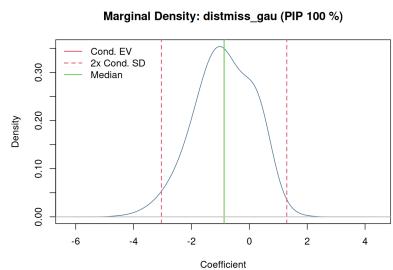
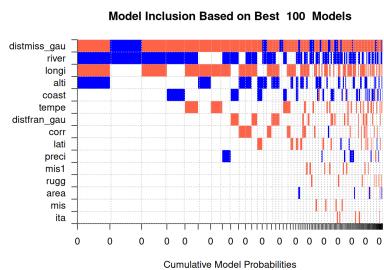
## Logarithmic Distance, Fixed Inclusion



[show all tables and figures](#)

16/31

## Gaussian Distance, Fixed Inclusion



[show all tables and figures](#)

17/31

## Recap

Using BMA to Evaluate the Model Setup

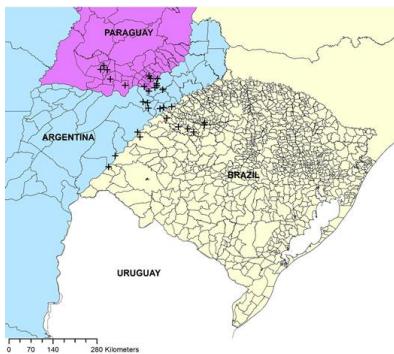
## Spatial Scope

Spatial Methods

What Lies Ahead

18/31

## Back to the Original Map



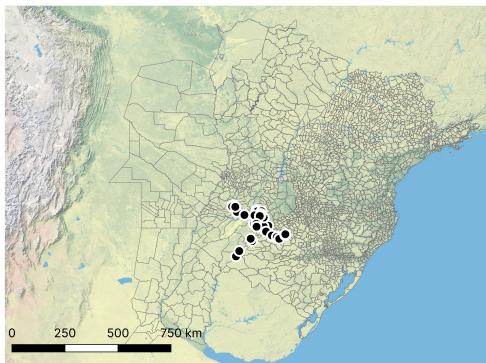
The original dataset contains a **spatially unbalanced** selection of localities.

Especially linear distance may be **overly sensitive** to observations far away from the missionary region.

Figure I from Valencia Caicedo (2019)

19/31

## Let's Extend the Sample



Spatially extended sample

Therefore, I **collected data** for surrounding states in Brazil and Argentina as well as the entirety of Paraguay.

(Uruguay might also be interesting.)

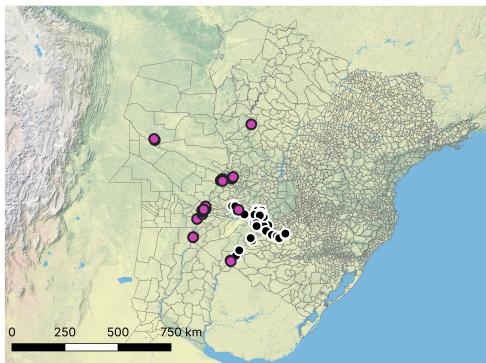
20/31

## Data Sources

- **Geographic characteristics** are taken from GADM (2023)
  - Area, longitude, latitude, distance calculation
- **Climate data** is from Fick and Hijmans (2017)
  - Temperature, precipitation
- **Elevation** is taken from Farr et al. (2007)
- **Ruggedness** is taken from Amatulli et al. (2020)
- **River streams** are taken from Allen and Pavelsky (2018)

21/31

## Other Things that Happened Around that Area



Black: Jesuit Missions; Pink: Battles

Missions are somewhat **clustered**, and municipalities farther away have similar distances to the closest mission.

Mapping **battles of the Paraguay War** (1864–1870) yields some overlap with the missionary area.

22/31

## Regression Results with Extended Data Set

(Dep.: illiteracy)	<b>Mission Distance</b>	<b>Both Variables</b>	<b>Battle Distance</b>
dismiss	0.0029** (0.0011)	0.0100*** (0.0015)	
distbattle		-1.510*** (0.1980)	-0.5942*** (0.1401)
Observations	2,202	2,202	2,202
R <sup>2</sup>	0.33385	0.35119	0.33646
Adjusted R <sup>2</sup>	0.32215	0.33949	0.32481

Heteroskedasticity-robust standard-errors in parentheses

Signif. Codes: \*\*\*: 0.01, \*\*: 0.05, \*: 0.1 [show full table](#)

23/31

Recap

Using BMA to Evaluate the Model Setup

Spatial Scope

## Spatial Methods

What Lies Ahead

24/31

## Ways to Treat Space

- **Space** occurs in the models from the paper in various instances: **coordinates** as explanatories, linear **distances**, and Conley **standard errors** as a robustness check.
- It may make sense to more **explicitly** account for a **spatial network**.
- The following models are estimated with an order-one queen **contiguity matrix** and otherwise unchanged from the original. They serve only as an illustration.
- Next steps are revising the **specification** and varying the **matrices** (e.g. distance decay matrix).

25/31

## SLX Results

(Dep.: illiteracy)	<b>SLX No Controls</b>	<b>SLX Geo Controls</b>
dismiss	-0.004 (0.007)	0.003 (0.009)
lag.dismiss	0.005*** (0.002)	0.005** (0.002)
Observations	556	556
R <sup>2</sup>	0.080	0.168
Adjusted R <sup>2</sup>	0.054	0.122
Residual Std. Error	3.913 (df = 540)	3.770 (df = 526)
F Statistic	3.120*** (df = 15; 540)	3.662*** (df = 29; 526)

Note:

\*p<0.1; \*\*p<0.05; \*\*\*p<0.01

26/31

## SDM Results

(Dep.: illiteracy)	<b>SDM No Controls</b>
dismiss	-0.007 (0.005)
lag.dismiss	0.002** (0.001)
Observations	556
Log Likelihood	-1,395.697
$\sigma^2$	7.805
Akaike Inf. Crit.	2,827.395
Wald Test	2,150.868*** (df = 1)
LR Test	287.349*** (df = 1)

27/31

## Recap

### Using BMA to Evaluate the Model Setup

#### Spatial Scope

#### Spatial Methods

## What Lies Ahead

28/31

## Still to Do

- To average over different distance measures, it would be beneficial to be able to **restrict** the model space to models with **only one distance measure**.
- It may make sense to impose a **prior on the decay parameter(s)** instead of fixing them.
- **SDM** and the **SLX** models require a more careful implementation.
- It would be nice to be able to **add variables**, but suitable variables (e.g. initial population density) are hard to find.

29/31

## References (1)

- Allen, G. H., & Pavelsky, T. M. (2018). Global River Widths from Landsat (GRWL) Database.  
<https://doi.org/10.5281/ZENODO.1297434>
- Amatulli, G., McInerney, D., Sethi, T., Strobl, P., & Domisch, S. (2020). Geomorpho90m, empirical evaluation and accuracy assessment of global high-resolution geomorphometric layers. *Scientific Data*, 7(1). <https://doi.org/10.1038/s41597-020-0479-6>
- Farr, T. G., Rosen, P. A., Caro, E., Crippen, R., Duren, R., Hensley, S., Kobrick, M., Paller, M., Rodriguez, E., Roth, L., Seal, D., Shaffer, S., Shimada, J., Umland, J., Werner, M., Oskin, M., Burbank, D., & Alsdorf, D. (2007). The Shuttle Radar Topography Mission. *Reviews of Geophysics*, 45(2). <https://doi.org/10.1029/2005rg000183>
- Fick, S. E., & Hijmans, R. J. (2017). WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology*, 37(12), 4302–4315.  
<https://doi.org/10.1002/joc.5086>

30/31

## References (2)

- GADM. (2023). GADM: Database of Global Administrative Areas [Accessed January 17, 2025]. Includes shapefiles for Brazil, Argentina, and Paraguay.].  
[https://gadm.org/download\\_country.html](https://gadm.org/download_country.html)
- Valencia Caicedo, F. (2019). The Mission: Human Capital Transmission, Economic Persistence, and Culture in South America\*. *The Quarterly Journal of Economics*, 134(1), 507–556.  
<https://doi.org/10.1093/qje/qjy024>

31/31

## Max Heinze

📍 1020 Vienna, Austria 📞 +43 1 31336-5424 📩 mheinze@wu.ac.at 🌐 LinkedIn 🌐 GitHub

### Professional Experience

<b>Research and Teaching Associate</b> Department of Economics, Vienna University of Economics and Business	since Sep. 2024 Vienna, Austria
• <b>Research Interests:</b> Environmental Economics, Development Economics, Applied Econometrics	
• <b>Teaching:</b> Macroeconomics (B.Sc. Level) – Winter Term 2024/25	
<b>Project Research Assistant</b> Institute of Ecological Economics, Vienna University of Economics and Business	Feb. 2023 – Aug. 2024 Vienna, Austria
<b>Teaching Assistant for Macroeconomics (B.Sc. Level)</b> Department of Economics, Vienna University of Economics and Business	Sep. 2021 – Jul. 2024 Vienna, Austria
<b>Teaching Assistant for Econometrics (B.Sc. Level)</b> Department of Economics, Vienna University of Economics and Business	Feb. 2022 – Feb. 2023 Vienna, Austria
<b>Junior Data Analyst</b> World Data Lab	Mar. 2021 – Sep. 2022 Vienna, Austria
<b>Teaching Assistant for Advanced Microeconomics (B.Sc. Level)</b> Department of Economics, Vienna University of Economics and Business	Mar. 2021 – Jun. 2021 Vienna, Austria
<b>Intern Group Controlling</b> BWT AG	Aug. 2020 – Sep. 2020 Mondsee, Austria

### Education

<b>Ph.D. in Economics</b> Vienna University of Economics and Business	since Sep. 2024 Vienna, Austria
<b>M.Sc. in Economics (Science Track)</b> Vienna University of Economics and Business · GPA: 1.21	Sep. 2022 – Aug. 2024 Vienna, Austria
• <b>Exchange Semester:</b> Stockholm School of Economics • <b>Thesis:</b> Does Mining Activity Make Food Insecurity Worse?	Stockholm, Sweden

<b>B.Sc. in Business, Economics and Social Sciences</b> Vienna University of Economics and Business · GPA: 1.09	Sep. 2018 – Sep. 2022 Vienna, Austria
--	--

### Other Activities

<b>Student Representative</b>	Jul. 2023 – Aug. 2024
• One of five elected representatives for WU's Economic and Socio-Economic Master's Programs	
<b>Member of WU Center of Excellence</b>	Dec. 2022 – Jun. 2024
• Two-year program for selected high-potential Master's students	
<b>Member of WU Top League</b>	Sep. 2018 – Jun. 2021
• Three-year program for high-potential Bachelor's students	

### Skills and Interests

**Languages:** German (Native), English (Proficient), Swedish (Interm.), Spanish (Basic), Italian (Basic)  
**Software:** R, LATEX, Tikz, Python (Basic), MATLAB (Basic), Dynare (Basic), Microsoft Office

# Mines–Rivers–Yields

## Downstream Mining Impacts on Agriculture in Africa

Lukas Vashold<sup>†\*</sup> Gustav Pirich<sup>†</sup> Maximilian Heinze<sup>†</sup>  
Nikolas Kuschnig<sup>†\*</sup>

<sup>†</sup> Vienna University of Business and Economics

Minerals are essential to fuel the green transition, can foster local employment and facilitate economic development. However, their extraction is linked to several negative social and environmental externalities. These are particularly poorly understood in a development context, undermining efforts to address and internalize them. In this paper, we exploit the discontinuous locations of mines along rivers and their basins to identify causal effects on agricultural yields in Africa. We find considerable impacts on vegetation and yields downstream, which are mediated by water pollution and only dissipate slowly with distance. Our findings suggest that pollution from mines may play a role in the limited adoption of intensive agriculture. They underscore an urgent need for domestic regulations and international governance to limit negative externalities from mining in vulnerable regions.

\*Correspondence to [nikolas.kuschnig@wu.ac.at](mailto:nikolas.kuschnig@wu.ac.at), [lukas.vashold@wu.ac.at](mailto:lukas.vashold@wu.ac.at), at Welthandelsplatz 1, 1020 Vienna, Austria. Equal contribution authors. Funding by the Oesterreichische Nationalbank (OeNB anniversary fund, project No. 18799) is gratefully acknowledged.

## 1. Introduction

Mines provide vital materials for global supply chains and can play an important role in local economies. The extracted minerals are essential manufacturing components and are required for the transition towards cleaner energy (Masson-Delmotte et al., 2022; Pörtner et al., 2022). As a result of the global push for green technologies, the demand for minerals is projected to increase considerably, reaching 1.8–3.5 billion tons of material extracted by 2050 (Hund et al., 2023). Mining also provides local economic opportunities, increasing incomes and the levels of wealth and asset ownership (Bazillier and Girard, 2020; Goltz and Barnwal, 2019; Ofosu et al., 2020). At the same time, mines are linked to severe negative externalities that are primarily borne locally. The resource wealth from mining may crowd out other industries, drive corruption and conflict,<sup>1</sup> and induce a *resource curse*. Crucially, mines are prolific sources of pollution (of water, air, soil, and food; see Awotwi et al., 2021; Macklin et al., 2023; Mwelwa et al., 2023) and require copious amounts of water for their operation (Northey et al., 2018), threatening local ecosystems and agriculture. Effective governance is vital to realize the benefits from mining (Ali et al., 2017), but relies on scarce evidence on the nature and extent of these externalities.

The negative externalities of mines are a particular concern for countries with relatively weak institutions, which might otherwise enable them to internalize costs, and impacts are likely to go unmitigated. For many countries on the African continent, where mining operations have been booming (ICMM, 2022), this is a major challenge. Artisanal and small-scale mining is prevalent (ASM Inventory, 2022; Girard et al., 2022), and many mines lack the containment facilities required to manage pollution (Kossoff et al., 2014; Macklin et al., 2023). In Africa, agriculture accounts for a considerable portion of economic value-added in these countries (World Bank, 2024), and subsistence farming remains prevalent. Pollution may thus cause large disruptions (economic and nutritional) due to fewer possibilities and limited opportunities for adaptation. Notably, food insecurity in Africa is severe and worsening, while other regions have made improvements (Food and Agriculture Organization of the United Nations et al., 2023). The problem is made worse by the scarcity of reliable

<sup>1</sup>Berman, Couttenier, and Girard, 2023; Berman, Couttenier, Rohner, et al., 2017; Knutzen et al., 2016; Rigterink et al., 2023.

and general insights into the external costs of resource extraction and mining in particular, due to the limited availability of data.

In this paper, we provide causal evidence for considerable effects of water pollution from mines on agriculture and natural vegetation in Africa. For identification, we exploit a discontinuity in the flow of water that arises from the location of mines along fine-grained river-basins (based on data by Lehner and Grill, 2013). We use the remotely sensed Enhanced Vegetation Index (EVI; see Didan, 2015) to measure agricultural productivity on croplands, and vegetation health in general. Coupled with a comprehensive dataset of industrial and artisanal mining sites (see Maus et al., 2022) and other remotely sensed covariates, this allows us to assess the impacts of mines across the African continent. An illustration of the discontinuity we exploit is provided in Figure 1. The presence of mines tends to cause a sharp drop in vegetation health downstream. By contrast, regions that are upstream of mines are unaffected by polluted water and can serve as a control group.

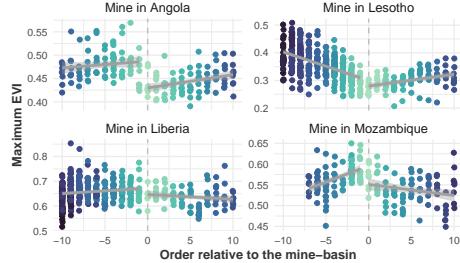


FIGURE 1: Vegetation index around four selected mines in Angola, Lesotho, Liberia, and Mozambique (over time). Mine locations are indicated with a dashed line (in the center); up- and downstream river basins (by discrete order) are plotted with linear trend lines. A variant with river distance on the horizontal axis is available in Figure B1 of the Appendix.

We find that mining sites have considerable effects on vegetation downstream. Overall vegetation health, as measured by the maximum annual EVI within

basins, is reduced by about 1.4–1.6 percent at the mean. The reduction in EVI is larger for agricultural land, which we identify using established land cover masks. This cropland-specific EVI is reduced by 1.9–2.1 percent at the mean. These effects are economically significant and robust to a multitude of robustness checks. Both impacts dissipate further downstream, in line with hydrological studies on the transport of sediment and other pollutants (see Macklin et al., 2023, and references therein). The speed of dissipation is heterogeneous — the effect halves at a distance of 79 km for general vegetation, and at 17 km for croplands. This suggests different mediators for general and cropland vegetation or adaptation behavior in the latter. We also assess the possibility of heterogeneous effects regarding the characteristics of mines, biomes, and in terms of regions. Larger, but not necessarily faster-growing mines appear to cause larger impacts. Effects are strongest for grassland-like biomes, and particularly pronounced in West Africa, where artisanal mining of gold is prevalent.

We contribute to the literature in two major ways. First, we provide reliable causal evidence for the widespread impact of mines on natural vegetation and agricultural yields in Africa. Macklin et al. (2023) investigate river contamination from metal mining, and estimate that 23 million people live in river basins that are affected by pollution from mining. However, *causal* evidence for the consequences of this pollution on agricultural productivity remains scarce and limited in scope. For instance, Aragón and Rud (2015) provide evidence that gold mining in Ghana reduced agricultural productivity by about 40 percent, while Mwelwa et al. (2023) track heavy metals from mines along the food chain in Zambia and find considerable impacts. Our analysis provides strong evidence for mines as one mediator for the relation between institutional strength and agricultural yields (Wuepper et al., 2023) in the African context.

Second, we add an important dimension to the impacts of pollution in a development context. While air pollution is comparatively well-researched, water pollution has received little attention in the literature (Keiser, 2019). The studies that do exist tend to focus on *drinking water* (Greenstone and Hanna, 2014; Keiser and Shapiro, 2018; Olmstead, 2010) and pollution *from* agriculture (Moss, 2007). Studies in a context with relatively weak institutions are rare (see Greenstone and Hanna, 2014; Liu et al., 2021, for two notable exceptions). We follow Keeler et al. (2012) in viewing water quality as a contributor to

ecosystem services, and show how natural vegetation and agricultural yields are impacted by mining-induced water pollution. This impact is likely to be especially pronounced and consequential in the context of our study, where environmental regulations are often lacking, subsistence farming is prevalent, and food insecurity remains a challenge. To the extent that high-income countries drive the demand for minerals and outsource polluting industries to low-income and developing countries, our results can be seen in the context of global environmental justice (Banzhaf et al., 2019; Hilson and McQuilken, 2014).

Our analysis has some notable parallels in the literature. The identification procedure, which relies on the direction of water flow along basins, has been previously exploited by Dias et al. (2023) to identify the impacts of glyphosate use (in the cultivation of soybean in Brazil) on birth outcomes. Another notable study by E. Strobl and R. O. Strobl (2011) uses a conceptually similar approach to assess the impact of dams on agricultural productivity in Africa. They find an overall productivity loss of 0.96 percent for the period 1981–2000. The water use of mining operations (and its implications for the local environment) is an important research strand in the environmental literature (Luckeneder et al., 2021; Moura et al., 2022; Northey et al., 2018), to which we add. Our study also relates to the literature on the role of institutions in agriculture (Wuepper et al., 2023) and water pollution. Sigman (2002) finds increased pollution levels upstream of international borders, providing evidence of freeriding at a national level. We show that this problem also exists within countries.

The remainder of this paper is structured as follows. In the next section on materials and methods, we introduce our conceptual framework, and the data and empirical strategy that we use to identify the causal effects of water pollution from mining on agricultural productivity. In Section 3, we present our main results, illustrate potential heterogeneities, and assess the robustness of them. In Section 4, we discuss our results and their implications and conclude.

## 2. Materials and methods

In this section, we introduce the conceptual framework of our study, describe the data we use, and outline our empirical strategy. Our approach builds on the discontinuity from mines being located along a chain of river basins that

are connected by directed water flows. The flow of water is used to identify the causal effect of mine-induced water pollution, which reaches basins downstream, but not upstream of a mine. We do not consider other impacts of mines, such as air pollution or soil erosion. Vegetation is measured with the Enhanced Vegetation Index (EVI), which allows for consistency over time and regions while providing information at a granular level. The mine dataset we employ allows us to cover artisanal and small-scale mines in addition to larger industrial sites. The effect of interest is operationalized in two ways: via the up- and downstream order of basins, and via the river distance. With this setup at hand, we estimate the impact of mining-related water pollution on vegetation health and agricultural productivity across the African continent.

### 2.1. Conceptual framework

Mining metals and other minerals requires large amounts of water at the various stages of operation (Moura et al., 2022; Northey et al., 2018). In the process, the water is often contaminated with toxic chemicals. In the case of gold mining, that includes mercury, lead, and sodium cyanide. Afterward, the water is redirected into the stream it was taken from, often without proper treatment. In addition, mines produce a large amount of ‘tailings’ (waste rock and sediment), which are stored in large deposits called tailing dams. When oxidized by air and weathered by rain, these tailings steadily cause pollution of water resources (both above and below the surface) as they feed into rivers (Schwarzenbach et al., 2010). The structural failure of these tailing dams is a major environmental concern. While causal estimates of the effects of water pollution from mines are lacking, Macklin et al. (2023) estimate that almost 500,000 kilometers of river channels may be affected by metal mining in this way.

We seek to estimate the causal effect of water pollution from mining sites on both natural vegetation and agricultural yields. For this, we locate mines within river basins (i.e., their catchment areas), and exploit the discontinuity from the mine to compare basins up- and downstream of the mine. The unidirectional flow of water along these basins allows us to identify the water-mediated impact of mining sites. This is illustrated in Figure 2, which shows two mine-basins and their surroundings. The mine-basin itself and basins downstream are affected

by water pollution from the mine, while basins upstream can serve as a control group.

Identification relies on the outcome being *discontinuously affected* at the mine location only by the mine itself. There are few concerns for such impacts following alongside river basins, which are determined by geography (i.e., elevation and slope); we will discuss and attempt to soothe them below. There is also a possibility that farmers downstream of mines choose to relocate to unaffected regions, or adapt their farming practices. The potential relocation of farmers, which would reduce yields mechanistically, is understood as part of the impact. (This means, that part of the effect may be due to its perception by farmers.) Adaptation, in turn, would likely attenuate our estimates. Our second outcome — natural vegetation that is not specific to agricultural sites — is less likely to be impacted by either behavior.

There are two important considerations when it comes to *external validity*. First, the discontinuity that we exploit is tied to the direction of the water flow. Our estimates do not reflect other local impacts that are not transmitted along river-basins, but may still affect vegetation and agricultural productivity. These include adverse effects such as heightened air pollution (Pandey et al., 2014), soil erosion (Jarsj et al., 2017) or food contamination (Mwelwa et al., 2023), but also potential positive effects such as increased wealth levels, asset ownership, or incomes (Goltz and Barnwal, 2019; Oforu et al., 2020) that may drive irrigation or fertilizer use. Second, agriculture in Africa rarely makes use of irrigation, being rainfed instead. Our analysis is focused on impacts that are transmitted along river-basins, and only impacts on agricultural lands (and techniques) that use water from these streams are identified. While insights may not generalize to rainfed agricultural land, it is notable that the effect that we estimate may play a role in the slow adoption of intensive agriculture that relies on irrigation.

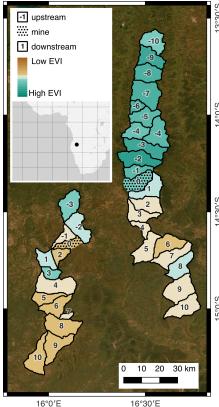


FIGURE 2: Two mines alongside their up and downstream basins.

## 2.2. Dataset

Our dataset is a panel of  $N = 14,327$  units that we observe over the period from 2016–2023 ( $T = 8$ ). The units of observation are hydrological basins in Africa (from the HydroBASINS dataset, see Lehner and Grill, 2013), which are delineated using the location of bodies of water and remotely sensed elevation, terrain slope, and stream gradient data.<sup>2</sup> Basins are nested in a hierarchical structure; each of the twelve levels features basins of roughly comparable size. The defining feature of basins is their directionality — water only flows downstream, and the water in an upstream basin will (generally) not be affected by anything that happens downstream. For our analysis, we split basins into three types. Mine-basins (of order zero) that contain a mine, and basins that are upstream of *all* and downstream of *any* mine-basin. The order of basins, i.e., the discrete distance in basins, is our primary measure for the treatment intensity. As an alternative, we rely on the distance along the river network (in kilometers). We observe 1,900 mine-basins, and consider up- and downstream basins up to order ten for our analysis, for a total of 6,698 upstream- and 5,729 downstream-basins (for a summary, see Table C1 in the Appendix).

For a given basin, the next (i.e., order one) downstream basin is always unique, while there may be multiple upstream basins of any given order. This is because streams can join, but not split in the direction of their flow (compare the basins labeled ‘−4’ in Figure 2). Not all mine-basins have a full set of up- or downstream basins, and the number generally decreases with the order. When mine-basins are located in or near the top or bottom branch of a river network, or if another mine-basin is situated up- or downstream of the mine-basin, there will be fewer relevant basins up- and downstream. Each basin occurs only once in our sample, and a basin that is, e.g., upstream of two mine-basins is only associated directly with one of them.

**Outcome** Crop statistics reported by official institutions lack the spatial granularity needed to measure localized impacts. Even when fine-scale data is available, institutional differences in how crop yields are calculated and reported would lead to inconsistencies across countries, threatening comparability, and

---

<sup>2</sup>We use the version of the dataset that specifically accounts for the position of lakes, delineating lake-adjacent basins similarly to coastal basins.

possibly biasing estimates. We thus opt for the Enhanced Vegetation Index (EVI) as a time-consistent and fine-scale spatially explicit measure of agricultural productivity as our outcome variable.<sup>3</sup> To produce a proxy for agricultural productivity, we first filter out low-quality pixels (due to cloud cover) and aggregate the mean EVI per basin for the available 16-day time frame. We use these values to compute the maximum annual EVI as a *peak vegetation index*, a proxy that has been shown to be tightly correlated with measures of gross primary production of vegetation (see Shi et al., 2017, for an assessment).

Vegetation indices on croplands specifically correlate strongly with crop yields and are frequently used in the literature.<sup>4</sup> To zoom in on agricultural productivity, we consider two cropland masks to filter for areas that are specifically used for agricultural purposes. Digital Earth Africa (2022) provides a cropland mask that is specifically trained and targeted to Africa, achieving a high overall accuracy, between 86.4% and 90.7%. However, this product is anchored at 2019 and not time-varying. Thus, as an additional robustness check, we use a time-varying cropland mask, provided by the European Space Agency (ESA; Defourny et al., 2024). The ESA land cover classification is broader (being on a global scale) and coarser in its resolution, and may thus be more susceptible to misclassification than the Digital Earth Africa product. However, it may help capture relocation of croplands in the period considered, which may help reflect productivity changes in agricultural areas. Not all basins contain croplands (for either mask), reducing our sample for the analysis of cropland-specific effects.

**Treatment** The treatment of interest is the location of mining sites, which we obtain from Maus et al. (2022). They provide locations and delineated sites of large-scale, industrial mines as well as artisanal and small-scale mining (ASM). Compared to other commonly used mining databases, such as the SNL Metals & Mining database (from which Maus et al. (2022) depart), it offers the advantage that the delineation of mines was conducted manually within a 10 kilometer buffer around the point locations of known mines. As a result, it also captures smaller mines in the vicinity. This is particularly important in Africa, where ASM is prevalent, and data records are often incomplete. This increases the

<sup>3</sup>Specifically, we use the product by Didan (2015) that is derived from MODIS satellite imagery.

<sup>4</sup>See, e.g., Azzari et al., 2017; Becker-Reshef et al., 2010; Bolton and Friedl, 2013; Johnson, 2016; Shammi and Meng, 2021.

coverage and accuracy of our treatment variable beyond what is usually found in the literature. While we cannot exclude the possibility that some mines are missing from our sample, the delineation approach (within a 10 km buffer) soothes concerns of effect attenuation from mines nearby.<sup>5</sup>

To construct our treatment, we intersect the mining sites of Maus et al. (2022) with the hydrological basins from Lehner and Grill (2013). Basins that contain one or multiple mines are directly treated (with a binary indicator), while the basins that are further downstream of the mine-basin are indexed via the distance along the river network or the order of the basin (with separate binary indicators). Basins that are upstream of the mine (and not downstream of any other mines) are considered as untreated and thus constitute our control group. In some cases, the up- and downstream areas of different mines overlap. To resolve this overlap and avoid duplicate basins in our sample, we consider each basin to only be associated with the closest mine-basin in terms of order.

**Other variables** We consider a number of covariates that may confound the effect of interest. Most importantly, that is elevation and the slope of land within a given basin. Both directly relate to the basin-level discontinuity and may thus suffer from imbalance across treatment and control basins, and could, at worst, bias estimates. We use detailed grid-level data (Amatulli et al., 2018a,b) for both, and aggregate information from the 30 arcsec (802–926 meters in the study area) cells to the basin level. We also consider differences in soil type, considering the primary soil class present in a given basin, obtained from the SoilGrids project (Hengl et al., 2017).

Other factors that may play a role, particularly at larger distances, include climate and socioeconomic characteristics. We account for climatological conditions by considering precipitation and maximum temperatures within basins. Both variables were retrieved from the Climatic Research Unit gridded Time Series (version 4.08), which are available on a monthly basis and a resolution of 0.5 degrees (Harris et al., 2020). We use the yearly sum of precipitation and the maximum of monthly temperatures, following previous studies that

<sup>5</sup>Some mines considered may be inactive, although mines that are considered to be inactive based on company reports are often mined illegally by individual miners. This has been shown to substantially contribute to water pollution even after their official decommissioning (Macklin et al., 2023). We investigate potential heterogeneity of active and inactive sites using a longitudinal dataset of mines by Sepin et al. (2024).

analyze vegetation dynamics in river basins (Na-U-Dom et al., 2017). In terms of socioeconomic circumstances, we control for a given basin’s population (derived from WorldPop, 2018) and its average accessibility, measured in minutes of travel time to the nearest city (derived from Weiss et al., 2018). Both datasets are available as grids with a resolution of approximately  $1 \times 1$  km; we anchor them in the year 2015 and aggregate them to the basin level by using total population and average travel time.

The effect of interest may be heterogeneous, depending on various circumstances. We specifically consider heterogeneity regarding persistent climatological conditions (as conveyed, e.g., by biomes) and mine characteristics. For the former, we allow the effect to vary across Ecoregions (following Dinerstein et al., 2017). As an alternative, we use the regional classification of the US Department of Agriculture, which groups countries based on primary crops and their varying crop calendar cycles.<sup>6</sup> In terms of mine characteristics, we investigate the intensity of mining activity as well as the activity over time. We proxy the former via the total mine area in a given mine-basin. For the activity of mines, we resort to the dataset provided by Sepin et al. (2024), which builds on previous mapping efforts (Maus et al., 2022; Tang and Werner, 2023) and uses machine learning techniques and satellite imagery to add temporal information on the evolution of mine sites over time.

**Summary statistics** Table 1 presents summary statistics of the data used. As remarked above, not all basins in our sample contain croplands, thus reducing our sample for the assessment of the impacts of mining on agricultural productivity. Both the maximum and the mean of the cropland-specific EVI (based on Digital Earth Africa, 2022) are slightly higher than the overall measure. The other covariates exhibit strong variation across our sample. Table C2 in the Appendix shows summary statistics split by their treatment status, i.e., location relative to the mine. Generally, up- and downstream basins are well-balanced with respect to the considered covariates. Downstream basins exhibit slightly higher precipitation, and are more highly populated, but less accessible, and lie at slightly lower altitudes. These minuscule differences across our treatment and control groups alleviate concerns of potential non-comparability. Still, we assess

<sup>6</sup>The regional classification can be found [here](#). We group North and East Africa in one region due to the low number of observations.

the robustness of our findings in various ways, including a coarsened exact matching as a way to increase balance across covariates and reduce model dependency, in Section 3.2.

TABLE 1: Summary statistics for the dataset used.

Variable	N	Mean	St. Dev.	Min.	Max.
Max. EVI	114,616	0.411	0.168	-0.112	0.993
Mean EVI	114,616	0.270	0.118	-0.112	0.578
Max. Cropland EVI	94,671	0.454	0.129	-0.112	0.990
Mean Cropland EVI	94,671	0.286	0.093	-0.114	0.734
Max. Temperature	114,616	33.80	4.047	20.00	45.40
Precipitation	114,616	882.3	606.3	0.555	4,375.3
Population	114,536	8,185	37,090	0.000	1,396,921
Elevation	114,616	804.6	482.0	-118.3	3,059.7
Slope	114,616	2.201	2.320	0.000	20.92
Accessibility	114,576	183.9	255.9	1.002	7,681

### 2.3. Empirical strategy

Formally, we use a quasi-experimental regression discontinuity (RD) design to estimate the causal effect of mines on vegetative health and agricultural productivity, as they are transmitted via water flows. That is, we estimate

$$y_{ijt} = \beta' f(d_{ij} \times \text{downstream}_j) + \delta' \mathbf{x}_{it} + \mu_j + \psi_t + \varepsilon_{ijt}, \quad (1)$$

where we relate the vegetation in basin  $i$ , located up- or downstream of mine  $j$  in year  $t$ , to a treatment in terms of distance to the nearest mine-basin. This distance is denoted as  $d_{ij}$ , and is operationalized in two different ways. Our preferred specification uses (indicators for) the basin order directly, while we use river distances in alternative specifications. Identification stems from the indicator  $\text{downstream}_{ij}$ , which is equal to one if a basin lies downstream of a mine (including the mine-basin). Basin-specific covariates are denoted with  $\mathbf{x}_{it}$ , mine- and year-fixed effects with  $\mu_j$  and  $\psi_t$ , and the error term,  $\varepsilon_{ijt}$ , is assumed to be homoskedastic.

The parameters contained in  $\beta$  are identified under the assumption that there are no *other* discontinuous changes in agricultural productivity at the

mine-basin. This assumption needs some justification. Fundamental differences of river-basins at each side of the discontinuity — e.g., in terms of size or other characteristics — would pose a threat to our identification strategy. In our analysis, we opt for the most granular layer of the hydrological basins (Lehner and Grill, 2013) that is available, meaning that the relevant discontinuity is more likely to be isolated. In the construction of the dataset, the original reference system for the delineation of hydrological basins (the Pfafstetter system, see e.g. K. L. Verdin and J. P. Verdin, 1999) was refined to enhance the comparability across basins in terms of size and connections to other basins. At Level 12, there is a total of 241,026 basins for the African continent, which cover an average ( $\{5, 50, 95\}^{\text{th}}$  percentile) area of 124.4 (11.6, 131.4, 218.9) km<sup>2</sup> (cf. Table C1 in the Appendix).

Some concerns regarding the non-comparability of higher-order basins (e.g., regarding geophysical or meteorological differences) may remain. We try to alleviate these concerns in various ways. First, we include and investigate the effect of potentially relevant covariates, including geophysical characteristics of basins (their elevation and slope), meteorological conditions (temperature and precipitation), and socioeconomic information (basins' accessibility and population). As we will show, the inclusion of these controls leaves the estimated effects largely unchanged. A matching exercise, where we induce balance without relying on a linear functional form, yields similar results. We also consider covariates as placebo outcomes, revealing no significant discontinuities at the mine-basins. Second, we subset our analysis and estimates to basins within a maximum order of ten (i.e., separated by at most nine basins from the mine-basin). On average, basins of the highest order are located roughly 100 kilometers away from the mine-basin (cf. Table C1), a distance at which most mining-related chemicals and sediment have dissipated according to hydrological studies (Macklin et al., 2023). Third, in a battery of robustness checks, we restrict our sample of basins even further, for example, by excluding the mine-basin itself or only considering basins immediately adjacent to the mine-basin (i.e., basins of order one). None of these restrictions refute the results of our baseline specifications.

Our preferred specification, where we use indicators for the basin-order, also remedies some concerns that would be inherent to RD designs with continuous

running variables (see, e.g. Cattaneo et al., 2019). It highlights the immediate discrepancy from basin to basin and reduces measurement error as it operationalizes distance at the level of our units of observations. Concomitantly, by relying solely on local information, it is relatively agnostic regarding assumptions of the functional form of distance and thus less susceptible to misspecification.<sup>7</sup> Furthermore, the basin-order allows for an intuitive assessment of the decay of the downstream effects of mines, that is independent of specific choices for the parameterization of distance.

Finally, we follow the recent literature on RD designs<sup>8</sup> in terms of best practices. We implement applicable procedures that alleviate concerns when using a continuous running variable, as is the case for our alternative distance specification. This includes routines that allow for the automatic selection of the bandwidth that determines the maximum value of the running variable for which observations are retained. Using the retained observations, polynomials of different orders can be fitted at each side of the discontinuity to compute bias-corrected and robust local average treatment effects. As we will show below, using these procedures leaves our results qualitatively unchanged.

### 3. Results

Table 2 presents our main results for the impact of mines on general (left two columns) and cropland-specific (right columns) vegetation. We present results for specifications that differ in the inclusion of covariates (one plain and one fully saturated version, first and second columns) and how they operationalize the effect of interest — namely via (a) indicators for all basin-orders, where the first upstream-basin is omitted (upper panel), and (b) the linear-quadratic river distance from the mine-basin and a downstream indicator (lower panel). The reported estimates are focused on parameters that are causally identified by our research design; complete estimates are provided in Tables C3–C4 of the Appendix.

<sup>7</sup>A notorious example concerns the choice of appropriate polynomials for continuous running variables (see Gelman and Imbens, 2019).

<sup>8</sup>See Calonico et al., 2014; Cattaneo et al., 2019; Imbens and Kalyanaraman, 2012; Kolesár and Rothe, 2018.

TABLE 2: Main estimation results.

Specification	<i>Max. EVI</i>		<i>Max. Cropland EVI</i>	
	(Plain)	(Full)	(Plain)	(Full)
<i>Order</i>				
Mine-basin (0 <sup>th</sup> )	-0.0064*** (0.0014)	-0.0059*** (0.0013)	-0.0093*** (0.0021)	-0.0095*** (0.0020)
Downstream (1 <sup>st</sup> )	-0.0060*** (0.0018)	-0.0057*** (0.0017)	-0.0049* (0.0026)	-0.0061** (0.0026)
Downstream (2 <sup>nd</sup> )	-0.0070*** (0.0021)	-0.0066*** (0.0021)	-0.0042 (0.0028)	-0.0062** (0.0030)
<i>Fit statistics</i>				
Sample mean	0.412	0.412	0.454	0.454
Observations	114,616	114,496	94,671	94,604
R <sup>2</sup>	0.912	0.924	0.780	0.786
<i>Distance</i>				
Downstream	-0.0065*** (0.0023)	-0.0058*** (0.0021)	-0.0086*** (0.0029)	-0.0087*** (0.0028)
Downstream × Distance	-2.0 × 10 <sup>-5</sup> (0.0001)	-2.0 × 10 <sup>-5</sup> (0.0001)	0.0003** (0.0001)	0.0002 (0.0001)
Downstream × Distance <sup>2</sup>	-4.0 × 10 <sup>-7</sup> (9.2 × 10 <sup>-7</sup> )	-9.8 × 10 <sup>-8</sup> (7.2 × 10 <sup>-7</sup> )	-2.2 × 10 <sup>-6**</sup> (1.1 × 10 <sup>-6</sup> )	-1.9 × 10 <sup>-6*</sup> (1.0 × 10 <sup>-6</sup> )
<i>Fit statistics</i>				
Sample mean	0.412	0.412	0.454	0.454
Observations	114,616	114,496	94,671	94,604
R <sup>2</sup>	0.918	0.924	0.780	0.786
<i>Controls</i>				
Geophysical	No	Yes	No	Yes
Meteorological Conditions	No	Yes	No	Yes
Socioeconomics	No	Yes	No	Yes
<i>Fixed-effects</i>				
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.

Note: The Table provides estimates of Equation 1, with treatment measured by indicators of basin-order in the upper panel, and linear-squared distance in the lower panel. The first two columns hold results for the overall EVI, the latter two columns results for the cropland-specific EVI. The first and third column include no covariates, whereas columns two and four include the full set of control variables. All specifications include mine and year fixed effects.

Our preferred basin-order specification indicates significant negative effect of being downstream of a mine, both for the general and the cropland-specific vegetation index. These impacts are economically meaningful; for the mine-basin, they correspond to a 1.4–1.6% reduction in the mine-basin for the general index, and a reduction of 1.9–2.1% for the cropland index relative to the sample mean. Estimates are robust to the inclusion of covariates. As visualized in Figure 3, this impact persists beyond the immediate mine-basin, providing strong evidence for water pollution as a mediator. Moreover, we can see that upstream basins are unaffected, allowing us to rule out a pretrend with reasonable confidence. For the general index, impacts appear to be more pervasive, only dissipating at the highest order. For the cropland index, the initial four basins (the mine and orders one to three) are significantly and considerably affected by the mine. Estimates for higher order basins (which are less well-represented in the data) become imprecise.

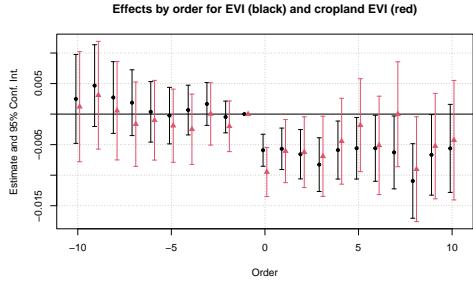


FIGURE 3: Estimated order coefficients for all up- and downstream basins (with the mine-basin in the center) for the overall EVI (in black) and the cropland EVI (in red, offset to the right) with full covariates. Whiskers represent 95%-confidence intervals.

For the second, distance-based, specification we find comparable results. The average impact of being located downstream of the mine is highly significant, and of a similar magnitude to the basin-order specification. For the general

vegetation index, the impact mirrors estimates for the mine-basin itself, reflecting the slow impact dissipation. For the cropland index, the imprecise higher-order estimates are pooled, and the overall impact is slightly below the one for the mine-basin. Notably, the operationalization via linear and linear-quadratic distance does not accurately reflect impact decay in the data. Estimates are only (barely) significant when no covariates are considered for the cropland index.

**Impact decay** To judge the extent and possible economic impact of the effect that we detect, we investigate the decay of effects next. The primary transmission channel is water pollution from tailings (leftover material) and chemicals that are used for the extraction and processing of target minerals. The dispersal of contaminants in water is unlikely to occur linearly along streams (Macklin et al., 2023). A more appropriate functional form to investigate the speed of decay is the exponential distance-decay function. Next, we consider this exponential formulation, where we additionally estimate the non-linear distance-decay parameter (see Section A in the Appendix for more details).

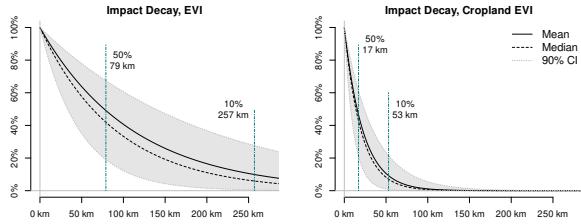


FIGURE 4: Effect decay over the distance, assuming an exponential decay function. The solid (dashed) black line denotes the mean (median) effect; the shaded area between the dotted lines denotes the 90% credible interval. The vertical lines denote the distance where the average effect is reduced by 50% and 90%.

In Figure 4, we see how impacts dissipate with distance along the river network. The differential speed of decay for general and agricultural vegetation, which is suggested by Figure 3, is even more pronounced. The effect on general vegetation is rather persistent, with the average effect decaying by 50% (90%)

after 79 (257) kilometers. By comparison, the average effect on the cropland vegetation decays by 50% (90%) after 17 (53) kilometers. The coefficients, i.e., effect at 100%, are  $-0.0028^{***}$  (0.0008) for general vegetation, and  $-0.0040^{***}$  (0.0011) for the agricultural index. Notably, these results are only informed by basins up to order ten, which lie at an average distance of 106.1 km. As a result, the tail impact decay for general vegetation relies on extrapolation.

### 3.1. Heterogeneity

Next, we investigate the heterogeneity of our results along several dimensions. We differentiate by (1) conditions that relate to the mines themselves, i.e., (1a) the total area mined and (1b) the activity of mining operations over time, and (2) spatial heterogeneities of basin systems, namely their (2a) geographical position in terms of region and (2b) biome. Figure 5 provides an overview of the main results, while Tables C5 and C6 in the Appendix report the full results.

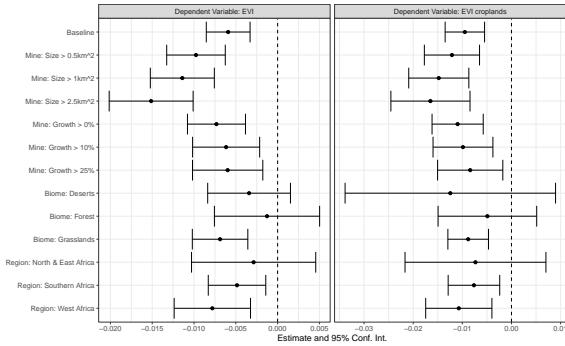


FIGURE 5: Average treatment effects and 95% confidence intervals in the first downstream basin for possibly heterogeneous subsets.

First, we consider heterogeneity along characteristics of the mine–basin in rows two to four of Figure 5. For the overall EVI, the left panel shows a clear increase

in the effect magnitude with increasing mine area. Compared to the baseline specification (which includes all mine-basins), the average effect is tripled for basins with a mining area of more than  $2.5 \text{ km}^2$ . This increase is less pronounced for the cropland-specific EVI (in the right panel). Although differences are insignificant, the increasing relationship seems to persist. Compared to the sample mean, the average reduction in the cropland-specific EVI increases from 2.1 percent in the baseline specification to 3.7 percent when only considering mine-basins with a mined area of more than  $2.5 \text{ km}^2$ . These findings do not come as a surprise, as larger mines are also more likely to produce more discharge material and contaminate nearby water reservoirs or flows.

Heterogeneity with respect to the activity of mines is investigated in rows five to seven. We use the growth of mining areas from 2017–2023 (based on Sepin et al., 2024) to approximate the activity of mining sites, and subset the dataset to mine-basins that exhibited any, at least 10%, or at least 25% growth in their mining area. Estimation results do not reveal substantial heterogeneities along this dimension; point estimates of the impacts remain stable, whereas the precision suffers from the decrease in sample size. Besides the noisy approximation and a true lack of heterogeneity, this result could be driven by a lack of maintenance of inactive mining sites. At the same time, precautions may be particularly prevalent during phases of expansion.

Second, we turn to spatial heterogeneities. In rows eight to ten of Figure 5, we split the sample by biome. We group the granular ecoregions of Dinerstein et al. (2017) into three broad biome groups that are present in our data — deserts, forests, and grasslands — and determine a primary group for each basin. For both, the overall and the cropland-specific EVI, significant effects are only detectable for basin systems in grasslands. Coefficients for the two other biomes are negative, but rather imprecise. This may be due to the limited number of observations for these biomes, or the rare practice of irrigated agriculture and applicability of the EVI in these biomes.

In the final three rows of Figure 5, we consider regional heterogeneities that reflect differences in crop cycles. We find somewhat larger effects in West Africa, slightly reduced effects in South Africa, and no significant effects in North and East Africa (which are grouped due to limited samples). These discrepancies are more pronounced for the overall EVI than for the cropland-specific one.

In many Western African countries, including Ghana, Mali, or Burkina Faso, gold mining is prevalent and often conducted in informal, small mines (Girard et al., 2022). The extraction of gold from raw ore uses toxic chemicals, such as mercury or sodium cyanide, and artisanal mining operations lack the necessary equipment to control their runoff (Hinton et al., 2003). This may contribute to larger impacts in West Africa, when compared to Southern Africa, where mining of other metals and coal is more widespread.

### 3.2. Robustness

Finally, we assess the robustness of our results. Figure 6 presents estimates for our main specification that are subject to various robustness checks. Specifically, we vary the exact sample considered, the outcome variable definition, and the level of fixed effects. Tables C7 and C8 in the Appendix report the full results of these exercises.

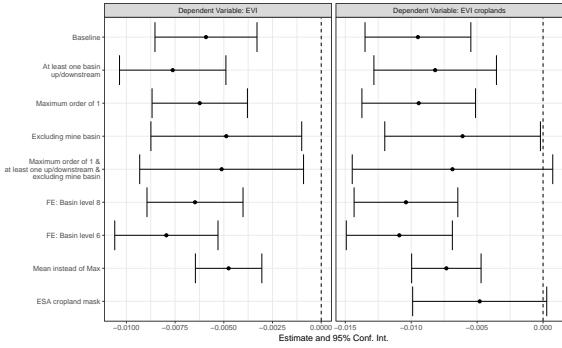


FIGURE 6: Average treatment effects and 95% confidence intervals in the first downstream basin under various robustness checks.

In rows two to five of Figure 5, we begin by restricting the sample in several ways. First, we only consider mine-basins that have at least one basin each that

is up- and downstream. Second, we restrict the maximum order of basins to one, disregarding any basins that are further away than the most immediate ones. Third, we exclude the mine-basin itself, only considering up- and downstream basins. Fourth and last, we combine the previous three restrictions. None of these restrictions invalidate our baseline results (which are shown in the first row). Only the combination of all restrictions leads to an estimate for the cropland-specific EVI that becomes insignificant at the 5% significance level, although the point estimate remains similar to the baseline. This loss in precision results from the rather drastic decrease in sample size through the combination of restrictions (less than 10% of the full sample).

In rows six and seven, we vary the specification in terms of fixed effects, i.e., we adjust the level at which we assume that time-invariant basin characteristics to manifest. We introduce fixed effects that relate to the mine-basin system itself — the hierarchical nature of basins allows us to identify all sub-basins within a given super-basin. Specifically, we introduce fixed effects for all sub-basins within (i) Level 6 and (ii) Level 8 super-basins. These contain on average 67 and 6 of the Level 12 basins that constitute our unit of observation. Estimation results are qualitatively unchanged.

Lastly, we vary our outcome variable — the proxy for vegetation productivity and agricultural yields — in rows eight and nine of Figure 5. First, we consider the annual *mean* of the EVI instead of the yearly maximum. Second, we use an alternative, time-varying cropland mask by ESA to help distill our agriculture-specific productivity.<sup>9</sup> The results are qualitatively similar for both of these variations, implying a negative effect on vegetation and agricultural yields. Note that both the estimates for the mean EVI, and the ESA cropland mask cannot be directly compared to the baseline estimate, as they have different sample moments (cf. Table 1). Compared to the sample mean, the mean overall EVI is reduced by 2.02 percent, whereas the reduction amounts to 2.35 percent for the cropland-specific EVI. While the effect on the EVI for croplands, as identified by the ESA mask, is only significant at the 10% level, it implies a one percent reduction compared to its sample mean. The reduction in precision of the

---

<sup>9</sup>For this, we rely on the land use classifications derived by the European Space Agency (Defourny et al., 2024). Specifically, we consider the land use classes for rainfed croplands (code 10) and irrigated or post-flooding cropland (code 20). We disregard mosaics of cropland with natural vegetation to reduce noise.

estimate stems from a lower number of observations, as the ESA mask identifies fewer agricultural areas in Africa compared to our region-specific baseline mask.

**Estimation methods** In addition to the aforementioned robustness check, we assess the robustness of our estimation method when operationalizing distance in kilometers along the river network. For that, we follow the recent literature that proposes data-driven and robust methods for inference in RD designs with a continuous running variable (see Cattaneo et al., 2019). This involves a data-driven bandwidth selection procedure, a weighting scheme for observations that are closer to the cutoff, and separately fitted local polynomials for untreated and treated units.<sup>10</sup> The results of this exercise are presented in Table C9 in the Appendix.

For the conventional estimates, the optimally chosen bandwidth falls between 20 and 38 kilometers. The bias-corrected estimates employ more observations, with the bandwidth ranging between 43 and 74 kilometers.<sup>11</sup> This is a narrower bandwidth than the one used in our main specification, where we implicitly set the distance threshold to about +/- 100 kilometers. Yet, the coefficient estimates from these routines for the effect of mining on vegetation downstream are very similar to our baseline specification reported in Table 2. With the full set of controls, and using either a linear or quadratic polynomial, the point estimates for the local average treatment effect range between -0.005 and -0.006 for the overall EVI and between -0.010 and -0.012 for the cropland-specific one. These estimates are robust to the bias-correction as proposed by Cattaneo et al. (2019) and statistically significant in all specifications.

Lastly, we assess the validity of our identification assumptions. First, we re-estimate the main specification using placebo outcomes; i.e., we change the dependent variable to covariates that are plausibly unaffected (recommended

<sup>10</sup>We follow the set of practices as outlined by Cattaneo et al. (2019) and employ a triangular kernel, which gives observations closer to the cutoff a greater weight. We chose the bandwidth by minimizing the mean squared prediction error and present the results for both linear and quadratic polynomials, which are fitted for up- and downstream basins separately (following Imbens and Kalyanaraman, 2012). The tables present the estimates for the local average treatment effects. We report the results for the conventional and bias-corrected estimation procedures, which provide an improved finite sample performance.

<sup>11</sup>These bandwidths are close to what theory, from hydrological studies, would suggest. Macklin et al. (2023), e.g., find elevated levels of toxic pollutants like zinc, lead, and arsenic between 10 and 80 kilometers downstream of mines.

by Cattaneo et al., 2019). The underlying assumption is that there are no other discontinuities that correlate with the treatment. We do not find a statistically significant effect for elevation, slope, temperature, precipitation, accessibility nor population (see Table C11 in the Appendix). Second, we induce balance among geophysical covariates using coarsened exact matching (Iacus et al., 2012) and re-estimate Equation 1 with the matched sample. Estimates are qualitatively similar to the main result. These results corroborate the validity of our identification assumption.

#### 4. Conclusion

In this paper, we identified the causal effects of mining on agricultural productivity mediated by water pollution. In a quasi-experimental research design, we used the location of mines along a river basin network as a discontinuity for identification. We compared agricultural productivity, measured by the Enhanced Vegetation Index (EVI), in basins upstream and downstream of mining sites. In our main specification, we found a reduction of the vegetation index downstream by 1.4–1.6% for all vegetation, and by 1.9–2.1% on croplands. Effects were stronger for larger mines, mines on grassland, and West African mines. Results were robust to changes in the sample considered, the definition of the outcome variable, the level of fixed effects, and the method of estimation.

The effect size can be further contextualized by comparing our estimate with effect sizes from the related literature. Adamopoulos et al. (2024) and Chen et al. (2022) find that institutions (in particular land tenure systems) affect agricultural productivity by about 20 to 40 percent. Aragón and Rud (2015) estimates the effect of industrial pollution from gold mining in Ghana on agricultural productivity at 40 percent. These papers use survey data to measure crop yields, which may be more responsive than our remotely sensed outcome measure. Several other papers use vegetation indices as a proxy for crop yields in related contexts, and find economically significant results that are comparable in size to our estimates.<sup>12</sup> In a related study, Wuepper et al. (2023) identify the impact of improving institutions on crop yields (also measured via the annual maximum EVI), and find an effect of 2.2 percent. The similar effect

<sup>12</sup>See, e.g. Asher and Novosad, 2020; Lobell et al., 2022; E. Strobl and R. O. Strobl, 2011.

size of our study suggests that the management of externalities from mines, and water pollution in particular, may be an important factor behind such institutional differences.

Our findings inform the discussion about resource extraction in Africa in general, and in countries with weak environmental governance in particular. The effects we found showcase a need to tackle the lack of containment facilities that have the potential to reduce negative impacts of extraction sites. Environmental legislation should introduce and enforce requirements for such facilities. Measures should not be limited to industrial mines, but extended to include closer monitoring of the informal mining sector, particularly in countries and regions where such activity is widespread. Since these mines play an important role in the livelihoods of many, stopping such operations altogether is arguably infeasible. Instead, it may be beneficial to provide information on potential countermeasures for water pollution and support workers in implementing them. While environmental externalities are often difficult to internalize due to their (economically) abstract nature, we show that water pollution from mines directly impacts agricultural yields. Addressing this externality is pressing, especially in regions with prevalent subsistence agriculture and ones that are plagued by food insecurity.

Our paper contributes causal estimates of mining impacts on agriculture in Africa, overcoming data scarcity (that has previously impeded large-scale analyses) by relying on remotely sensed data for both treatment and outcome. Nonetheless, there are some limitations and many future alleys for research. Remotely sensed data has important benefits, but also entails some limitations. Since vegetation indices derived from satellite data are an indirect measurement of agricultural productivity, the effects we find cannot directly be translated into agricultural yields without further information. Additionally, while our mine data is comprehensive and has broader coverage than conventionally used datasets, measuring treatment purely from the location and size of mines disregards differences in waste production and treatment that would be important to assess in future research. The evidence for broad effects presented in this paper calls for future work to provide more in-depth insights into the specifics of the mechanism, and, importantly, explores approaches to mitigate this externality of mines.

## References

- Adamopoulos, T. et al. (Mar. 2024). "Land security and mobility frictions". In: *The Quarterly Journal of Economics* 139.3, pp. 1941–1987. ISSN: 1531-4650. DOI: [10.1093/qje/qjae010](https://doi.org/10.1093/qje/qjae010) (page 23).
- Ali, S. H. et al. (Mar. 2017). "Mineral supply for sustainable development requires resource governance". In: *Nature* 543, pp. 367–372. ISSN: 1476-4687. DOI: [10.1038/nature21359](https://doi.org/10.1038/nature21359) (page 2).
- Amatulli, G. et al. (2018a). "A suite of global, cross-scale topographic variables for environmental and biodiversity modeling". In: *Scientific Data* 5.180040, pp. 1–15. ISSN: 2052-4463. DOI: [10.1038/sdata.2018.40](https://doi.org/10.1038/sdata.2018.40) (page 10).
- (2018b). "A suite of global, cross-scale topographic variables for environmental and biodiversity modeling, links to files in GeoTIFF format". In: *PANGAEA*. DOI: [10.1594/PANGAEA.867115](https://doi.org/10.1594/PANGAEA.867115) (page 10).
- Aragón, F. M. and J. P. Rud (Sept. 2015). "Polluting industries and agricultural productivity: Evidence from mining in Ghana". In: *The Economic Journal* 126.597, pp. 1980–2011. ISSN: 0013-0133. DOI: [10.1111/eco.12244](https://doi.org/10.1111/eco.12244) (pages 4, 23).
- Asher, S. and P. Novosad (Mar. 2020). "Rural roads and local economic development". In: *American Economic Review* 110.3, pp. 797–823. ISSN: 0002-8282. DOI: [10.1257/aer.20180268](https://doi.org/10.1257/aer.20180268) (page 23).
- ASM Inventory (2022). *World Maps of Artisanal and Small-scale Mining* (page 2).
- Awotwi, A. et al. (2021). "Impact of post-reclamation of soil by large-scale, small-scale and illegal mining on water balance components and sediment yield: Pra River Basin case study". In: *Soil and Tillage Research* 211, p. 105026. ISSN: 0167-1987. DOI: [10.1016/j.still.2021.105026](https://doi.org/10.1016/j.still.2021.105026) (page 2).
- Azzari, G., M. Jain, and D. B. Lobell (Dec. 2017). "Towards fine resolution global maps of crop yields: Testing multiple methods and satellites in three countries". In: *Remote Sensing of Environment* 202, pp. 129–141. ISSN: 0034-4257. DOI: [10.1016/j.rse.2017.04.014](https://doi.org/10.1016/j.rse.2017.04.014) (page 9).
- Banzhaf, S., L. Ma, and C. Timmins (Feb. 2019). "Environmental justice: The economics of race, place, and pollution". In: *Journal of Economic Perspectives* 33.1, pp. 185–208. ISSN: 0895-3309. DOI: [10.1257/jep.33.1.185](https://doi.org/10.1257/jep.33.1.185) (page 5).
- Bazillier, R. and V. Girard (Mar. 2020). "The gold digger and the machine. Evidence on the distributive effect of the artisanal and industrial gold rushes in Burkina Faso". In: *Journal of Development Economics* 143, p. 102411. ISSN: 0304-3878. DOI: [10.1016/j.jdeveco.2019.102411](https://doi.org/10.1016/j.jdeveco.2019.102411) (page 2).
- Becker-Reshef, I. et al. (June 2010). "A generalized regression-based model for forecasting winter wheat yields in Kansas and Ukraine using MODIS data". In: *Remote Sensing of*

- Environment* 114.6, pp. 1312–1323. ISSN: 0034-4257. DOI: [10.1016/j.rse.2010.01.010](https://doi.org/10.1016/j.rse.2010.01.010) (page 9).
- Berman, N., M. Couttenier, and V. Girard (2023). “Mineral resources and the salience of ethnic identities”. In: *Economic Journal* 133.653, pp. 1705–1737. ISSN: 0013-0133. DOI: [10.1093/ej/ued018](https://doi.org/10.1093/ej/ued018) (page 2).
- Berman, N., M. Couttenier, D. Rohner, et al. (2017). “This mine is mine! How minerals fuel conflicts in Africa”. In: *American Economic Review* 107.6, pp. 1564–1610. ISSN: 0002-8282. DOI: [10.1257/aer.20150774](https://doi.org/10.1257/aer.20150774) (page 2).
- Bolton, D. K. and M. A. Friedl (May 2013). “Forecasting crop yield using remotely sensed vegetation indices and crop phenology metrics”. In: *Agricultural and Forest Meteorology* 173, pp. 74–84. ISSN: 0168-1923. DOI: [10.1016/j.agrformet.2013.01.007](https://doi.org/10.1016/j.agrformet.2013.01.007) (page 9).
- Calanico, S., M. D. Cattaneo, and R. Titiunik (2014). “Robust nonparametric confidence intervals for regression-discontinuity designs”. In: *Econometrica* 82.6, pp. 2295–2326. ISSN: 0012-9682. DOI: [10.3982/ECTA11757](https://doi.org/10.3982/ECTA11757) (page 14).
- Cattaneo, M. D., N. Idrobo, and R. Titiunik (Nov. 2019). *A practical introduction to regression discontinuity designs: Foundations*. Cambridge University Press. ISBN: 9781108710206. DOI: [10.1017/9781108684606](https://doi.org/10.1017/9781108684606) (pages 14, 22, 23, xiii).
- Chen, C., D. Restuccia, and R. Santaella-Llopis (July 2022). “The effects of land markets on resource allocation and agricultural productivity”. In: *Review of Economic Dynamics* 45, pp. 41–54. ISSN: 1094-2025. DOI: [10.1016/j.red.2021.04.006](https://doi.org/10.1016/j.red.2021.04.006) (page 23).
- Defourny, P. et al. (2024). *Product user Guide and specification CDR and ICDR Sentinel-3 Land Cover (v2.1.1)*. Tech. rep. UCLouvain / European Space Agency. DOI: [10.24381/cds.00612c93a](https://doi.org/10.24381/cds.00612c93a) (pages 9, 21, xii).
- Dias, M., R. Rocha, and R. R. Soares (Feb. 2023). “Down the river: Glyphosate use in agriculture and birth outcomes of surrounding populations”. In: *Review of Economic Studies* 90.6, pp. 2943–2981. ISSN: 1467-937X. DOI: [10.1093/restud/rdad011](https://doi.org/10.1093/restud/rdad011) (page 5).
- Didan, K. (2015). *MOD13Q1 MODIS/Terra Vegetation Indices 16-Day L3 Global 250m SIN Grid V006*. DOI: [10.5067/MODIS/MOD13Q1.006](https://doi.org/10.5067/MODIS/MOD13Q1.006) (pages 3, 9).
- Digital Earth Africa (2022). *Cropland extent maps for Africa*. Tech. rep. Dataset. DE Africa Services (pages 9, 11).
- Dinerstein, E. et al. (2017). “An ecoregion-based approach to protecting half the terrestrial realm”. In: *Bioscience* 67.6, pp. 534–545. ISSN: 0006-3568. DOI: [10.1093/biosci/bix014](https://doi.org/10.1093/biosci/bix014). eprint: [28608869](https://doi.org/10.1093/biosci/bix014) (pages 11, 19).
- Food and Agriculture Organization of the United Nations et al. (2023). *The State of Food Security and Nutrition in the World 2023. Urbanization, agri-food systems transformation and healthy diets across the rural–urban continuum*. Report. DOI: [10.4060/cc3017en](https://doi.org/10.4060/cc3017en) (page 2).

- Gelman, A. and G. Imbens (2019). "Why high-order polynomials should not be used in regression discontinuity designs". In: *Journal of Business & Economic Statistics*. doi: [10.1080/07350015.2017.1366909](https://doi.org/10.1080/07350015.2017.1366909) (page 14).
- Girard, V., T. Molina-Millán, G. Vic, et al. (2022). "Artisanal mining in Africa". In: *Working Paper Series No. 2201*. Novafrica (pages 2, 20).
- Goltz, J. von der and P. Barnwal (June 2019). "Mines: The local wealth and health effects of mineral mining in developing countries". In: *Journal of Development Economics* 139, pp. 1–16. issn: 0304-3878. doi: [10.1016/j.jdeveco.2018.05.005](https://doi.org/10.1016/j.jdeveco.2018.05.005) (pages 2, 7).
- Greenstone, M. and R. Hanna (2014). "Environmental regulations, air and water pollution, and infant mortality in India". In: *American Economic Review* 104.10, pp. 3038–72. issn: 0002-8282. doi: [10.1257/aer.104.10.3038](https://doi.org/10.1257/aer.104.10.3038) (page 4).
- Harris, I. et al. (Apr. 2020). "Version 4 of the CRU TS monthly high-resolution gridded multivariate climate dataset". In: *Scientific Data* 7.109, pp. 1–18. issn: 2052-4463. doi: [10.1038/s41597-020-0453-3](https://doi.org/10.1038/s41597-020-0453-3) (page 10).
- Hengl, T. et al. (2017). "SoilGrids250m: Global gridded soil information based on machine learning". In: *PLOS ONE* 12.2, e0169748. issn: 1932-6203. doi: [10.1371/journal.pone.0169748](https://doi.org/10.1371/journal.pone.0169748) (page 10).
- Hilson, G. and J. McQuilken (Mar. 2014). "Four decades of support for artisanal and small-scale mining in sub-Saharan Africa: A critical review". In: *Extractive Industries and Society* 1.1, pp. 104–118. issn: 2214-790X. doi: [10.1016/j.exis.2014.01.002](https://doi.org/10.1016/j.exis.2014.01.002) (page 5).
- Hinton, J. J., M. M. Veiga, and A. T. C. Veiga (Mar. 2003). "Clean artisanal gold mining: A utopian approach?" In: *Journal of Cleaner Production* 11.2, pp. 99–115. issn: 0959-6526. doi: [10.1016/S0959-6526\(02\)00031-8](https://doi.org/10.1016/S0959-6526(02)00031-8) (page 20).
- Hund, K. et al. (2023). *Minerals for climate action: The mineral intensity of the clean energy transition*. Tech. rep. Washington DC, United States: World Bank Group (page 2).
- Iacus, S. M., G. King, and G. Porro (2012). "Causal inference without balance checking: Coarsened Exact Matching". In: *Political Analysis* 20.1, pp. 1–24. issn: 1047-1987. doi: [10.1093/pan/mpr013](https://doi.org/10.1093/pan/mpr013) (page 23).
- Imbens, G. and K. Kalyanaraman (2012). "Optimal bandwidth choice for the regression discontinuity estimator". In: *Review of Economic Studies* 79.3, pp. 933–959. issn: 0034-6527. doi: [10.1093/restud/rdr043](https://doi.org/10.1093/restud/rdr043) (pages 14, 22).
- International Council on Mining and Metals (2022). *Mining Contribution Index (MCI)* (page 2).
- Jarsj, J. et al. (Oct. 2017). "Patterns of soil contamination, erosion and river loading of metals in a gold mining region of northern Mongolia". In: *Regional Environmental Change* 17.7, pp. 1991–2005. issn: 1436-378X. doi: [10.1007/s10113-017-1169-6](https://doi.org/10.1007/s10113-017-1169-6) (page 7).

- Johnson, D. M. (Oct. 2016). "A comprehensive assessment of the correlations between field crop yields and commonly used MODIS products". In: *International Journal of Applied Earth Observation and Geoinformation* 52, pp. 65–81. ISSN: 1569-8432. doi: [10.1016/j.jag.2016.05.010](https://doi.org/10.1016/j.jag.2016.05.010) (page 9).
- Keeler, B. L. et al. (2012). "Linking water quality and well-being for improved assessment and valuation of ecosystem services". In: *Proceedings of the National Academy of Sciences* 109.45, pp. 18619–18624. doi: [10.1073/pnas.1215991109](https://doi.org/10.1073/pnas.1215991109) (page 4).
- Keiser, D. A. (2019). "The missing benefits of clean water and the role of mismeasured pollution". In: *Journal of the Association of Environmental and Resource Economists*. doi: [10.1086/703254](https://doi.org/10.1086/703254) (page 4).
- Keiser, D. A. and J. S. Shapiro (Sept. 2018). "Consequences of the Clean Water Act and the demand for water quality". In: *The Quarterly Journal of Economics* 134.1, pp. 349–396. ISSN: 1531-4650. doi: [10.1093/qje/qjy019](https://doi.org/10.1093/qje/qjy019) (page 4).
- Knutson, C. H. et al. (Sept. 2016). "Mining and local corruption in Africa". In: *American Journal of Political Science* 61.2, pp. 320–334. ISSN: 1540-5907. doi: [10.1111/aajs.12268](https://doi.org/10.1111/aajs.12268) (page 2).
- Kolesár, M. and C. Rothe (Aug. 2018). "Inference in regression discontinuity designs with a discrete running variable". In: *American Economic Review* 108.8, pp. 2277–2304. ISSN: 0002-8282. doi: [10.1257/aer.20160945](https://doi.org/10.1257/aer.20160945) (pages 14, xiv).
- Kossoff, D. et al. (2014). "Mine tailings dams: Characteristics, failure, environmental impacts, and remediation". In: *Applied Geochemistry* 51, pp. 229–245. ISSN: 0883-2927. doi: [10.1016/j.apgeochem.2014.09.010](https://doi.org/10.1016/j.apgeochem.2014.09.010) (page 2).
- Lechner, B. and G. Grill (Apr. 2013). "Global river hydrography and network routing: Baseline data and new approaches to study the world's large river systems". In: *Hydrological Processes* 27.15, pp. 2171–2186. ISSN: 1099-1085. doi: [10.1002/hyp.9740](https://doi.org/10.1002/hyp.9740) (pages 3, 8, 10, 13).
- Liu, Y. et al. (2021). "A review of water pollution arising from agriculture and mining activities in Central Asia: Facts, causes and effects". In: *Environmental Pollution* 291, p. 118209. ISSN: 0269-7491. doi: [10.1016/j.envpol.2021.118209](https://doi.org/10.1016/j.envpol.2021.118209) (page 4).
- Lobell, D. B., S. Di Tommaso, and J. A. Burney (June 2022). "Globally ubiquitous negative effects of nitrogen dioxide on crop growth". In: *Science Advances* 8.22. ISSN: 2375-2548. doi: [10.1126/sciadv.abm9909](https://doi.org/10.1126/sciadv.abm9909) (page 23).
- Luckeneder, S. et al. (2021). "Surge in global metal mining threatens vulnerable ecosystems". In: *Global Environmental Change* 69, p. 102303. ISSN: 0959-3780. doi: [10.1016/j.gloenvcha.2021.102303](https://doi.org/10.1016/j.gloenvcha.2021.102303) (page 5).
- Macklin, M. G. et al. (2023). "Impacts of metal mining on river systems: A global assessment". In: *Science* 381.6664, pp. 1345–1350. ISSN: 0036-8075. doi: [10.1126/science.adg6704](https://doi.org/10.1126/science.adg6704) (pages 2, 4, 6, 10, 13, 17, 22, ii).

- Masson-Delmotte, V. et al. (Aug. 2022). *Climate Change 2022: The Physical Science Basis*. [Online; accessed 15. Aug. 2022]. Cambridge: Cambridge University Press (page 2).
- Maus, V. et al. (2022). “An update on global mining land use”. In: *Scientific data* 9,1, pp. 1–11 (pages 3, 9–11).
- Moss, B. (2007). “Water pollution by agriculture”. In: *Philosophical Transactions of the Royal Society B: Biological Sciences* 363.1491, pp. 659–666. ISSN: 1471-2970 (page 4).
- Moura, A. et al. (2022). “Estimating water input in the mining industry in Brazil: A methodological proposal in a data-scarce context”. In: *Extractive Industries and Society* 9, p. 101015. ISSN: 2214-790X. doi: [10.1016/j.exis.2021.101015](https://doi.org/10.1016/j.exis.2021.101015) (pages 5, 6).
- Mwelwa, S. et al. (2023). “Biotransfer of heavy metals along the soil-plant-edible insect-human food chain in Africa”. In: *Science of The Total Environment* 881, p. 163150. ISSN: 0048-9697. doi: [10.1016/j.scitotenv.2023.163150](https://doi.org/10.1016/j.scitotenv.2023.163150) (pages 2, 4, 7).
- Neath, A. A. and J. E. Cavanaugh (Mar. 2012). “The Bayesian information criterion: background, derivation, and applications”. In: *WIREs Computational Statistics* 4,2, pp. 199–203. ISSN: 1939-5108. doi: [10.1002/wics.199](https://doi.org/10.1002/wics.199) (page ii).
- Northey, S. A. et al. (2018). “Production weighted water use impact characterisation factors for the global mining industry”. In: *Journal of Cleaner Production* 184, pp. 788–797. ISSN: 0959-6526. doi: [10.1016/j.jclepro.2018.02.307](https://doi.org/10.1016/j.jclepro.2018.02.307) (pages 2, 5, 6).
- Ofosu, G. et al. (2020). “Socio-economic and environmental implications of artisanal and small-scale mining (ASM) on agriculture and livelihoods”. In: *Environmental Science & Policy* 106, pp. 210–220. ISSN: 1462-9011. doi: [10.1016/j.envsci.2020.02.005](https://doi.org/10.1016/j.envsci.2020.02.005) (pages 2, 7).
- Olmstead, S. M. (2010). “The economics of water quality”. In: *Review of Environmental Economics and Policy*. doi: [10.1093/reep/rep016](https://doi.org/10.1093/reep/rep016) (page 4).
- Pandey, B., M. Agrawal, and S. Singh (Jan. 2014). “Assessment of air pollution around coal mining area: Emphasizing on spatial distributions, seasonal variations and heavy metals, using cluster and principal component analysis”. In: *Atmospheric Pollution Research* 5,1, pp. 79–86. ISSN: 1309-1042. doi: [10.5094/APR.2014.010](https://doi.org/10.5094/APR.2014.010) (page 7).
- Pörtner, H.-O. et al. (Aug. 2022). *Climate Change 2022: Impacts, Adaptation and Vulnerability*. [Online; accessed 15. Aug. 2022]. Cambridge: Cambridge University Press (page 2).
- Rigterink, A. S. et al. (2023). “Mining competition and violent conflict in Africa: Pitting against each other”. In: *Empirical Studies of Conflict Project (ESOC) Working Papers* 35 (page 2).
- Schwarzenbach, R. P. et al. (Nov. 2010). “Global water pollution and human health”. In: *Annual Review of Environment and Resources* Volume 35, 2010, pp. 109–136. doi: [10.1146/annurev-environ-100809-125342](https://doi.org/10.1146/annurev-environ-100809-125342) (page 6).

- Sepin, P., L. Vashold, and N. Kuschnig (2024). *Mapping Mining Areas in the Tropics from 2016–2024* (pages 10, 11, 19, ix).
- Shammi, S. A. and Q. Meng (Feb. 2021). “Use time series NDVI and EVI to develop dynamic crop growth metrics for yield modeling”. In: *Ecological Indicators* 121, p. 107124. ISSN: 1470-160X. doi: [10.1016/j.ecolind.2020.107124](https://doi.org/10.1016/j.ecolind.2020.107124) (page 9).
- Shi, H. et al. (Jan. 2017). “Assessing the ability of MODIS EVI to estimate terrestrial ecosystem gross primary production of multiple land cover types”. In: *Ecological Indicators* 72, pp. 153–164. ISSN: 1470-160X. doi: [10.1016/j.ecolind.2016.08.022](https://doi.org/10.1016/j.ecolind.2016.08.022) (page 9).
- Sigman, H. (2002). “International spillovers and water quality in rivers: Do countries free ride?” In: *American Economic Review* 92.4, pp. 1152–1159. ISSN: 0002-8282. doi: [10.1257/00028280260344687](https://doi.org/10.1257/00028280260344687) (page 5).
- Strobl, E. and R. O. Strobl (Nov. 2011). “The distributional impact of large dams: Evidence from cropland productivity in Africa”. In: *Journal of Development Economics* 96.2, pp. 432–450. ISSN: 0304-3878. doi: [10.1016/j.jdeveco.2010.08.005](https://doi.org/10.1016/j.jdeveco.2010.08.005) (pages 5, 23).
- Tang, L. and T. T. Werner (2023). “Global mining footprint mapped from high-resolution satellite imagery”. In: *Communications Earth & Environment* 4.1, p. 134 (page 11).
- Na-U-Dom, T., X. Mo, and M. Garcia (Feb. 2017). “Assessing the climatic effects on vegetation dynamics in the Mekong river basin”. In: *Environments* 4.1, p. 17. ISSN: 2076-3298. doi: [10.3390/environments4010017](https://doi.org/10.3390/environments4010017) (page 11).
- Verdin, K. L. and J. P. Verdin (May 1999). “A topological system for delineation and codification of the Earth’s river basins”. In: *Journal of Hydrology* 218.1, pp. 1–12. ISSN: 0022-1694. doi: [10.1016/S0022-1694\(99\)00011-6](https://doi.org/10.1016/S0022-1694(99)00011-6) (page 13).
- Weiss, D. J. et al. (2018). “A global map of travel time to cities to assess inequalities in accessibility in 2015”. In: *Nature* 553, pp. 333–336. ISSN: 1476-4687. doi: [10.1038/nature25181](https://doi.org/10.1038/nature25181) (page 11).
- World Bank (2024). *Agriculture, forestry, and fishing, value added (% of GDP)* (page 2).
- WorldPop (2018). *Global 1km Population*. Dataset. doi: [10.5258/SOTON/WP00647](https://doi.org/10.5258/SOTON/WP00647) (page 11).
- Wuepper, D. et al. (July 2023). “Institutions and global crop yields”. In: *NBER Working Paper*. doi: [10.3386/w31426](https://doi.org/10.3386/w31426) (pages 4, 5, 23).

## A. Impact decay assessment

The transport of chemicals used for the extraction and processing of target minerals and leftover material, called mine tailings (or ‘gangue’), via rivers and other water flows (e.g., groundwater flows or aquifers) is the primary transmission channel to downstream basins in our study. Hydrological studies on the dispersal and transportation of polluting materials from metal mining have shown that >90% of them are sediment associated and are transported 10–100 kilometers from their point of discharge by mining operations (see Macklin et al., 2023, and references therein). Furthermore, evidence from Macklin et al. (2023), based on a process-based model for the dispersal of contaminated sediment, suggests that for most elements associated with metal mining, concentrations decay non-linearly. As a result, it is unlikely that the effects on vegetation and agricultural productivity fade out linearly with increasing distance. While Figure 3 in the main text provides (agnostic) evidence of this decay based on the order of basins, we also investigate this decay with a parametric model.

We re-estimate the model introduced in Equation 1, assuming an exponential distance-decay function, i.e.  $\exp\{-\delta d_{ij}\}$ , where  $d_{ij}$  is the distance along the river in kilometers. We use an exponential decay parameterization because (i) hydrological studies provide evidence that this best fits the dispersion patterns of mining-related sediment, and (ii) it reduces the influence of far away observations, and (iii) the exponential function is commonly used for similar processes in other statistical applications. As the speed of decay, reflected by  $\delta$ , is not known a-priori, we conduct a grid search over the interval [0.001, 2] (in steps of 0.001), which would allow for both extremely rapid and slow decay.

For the results, we adopt a Bayesian model averaging approach, where we use the Bayesian information criterion (BIC) to approximate the marginal likelihood (and thus quantify model fit).<sup>13</sup> We then compute the mean speed of effect decay at increasing distances, depicted by the solid black line in Figure 4, and additionally provide quantile-based estimates. These are the  $(\{5, 50, 95\})^{\text{th}}$  quantiles for the effect decay parameter, which are depicted by dotted lines in Figure 4.

<sup>13</sup>We compute the posterior probability for model  $j$  with decay parameter  $\delta_j$  and Bayesian information criterion  $BIC_j$  as  $PP_j = \exp\{-BIC_j/2\} / \sum_j \exp\{-BIC_j/2\}$  (see e.g. Neath and Cavanaugh, 2012). The implicit prior is uniform across the searched interval.

## B. Additional Figures

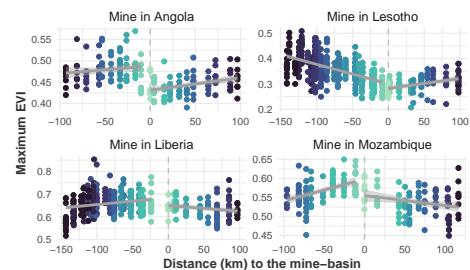


FIGURE B1: Enhanced vegetation index (over time) around four selected mines in Angola, Lesotho, Liberia, and Mozambique. Mine locations are indicated with a dashed line (in the center); up- and downstream river basins are plotted with linear trend lines.

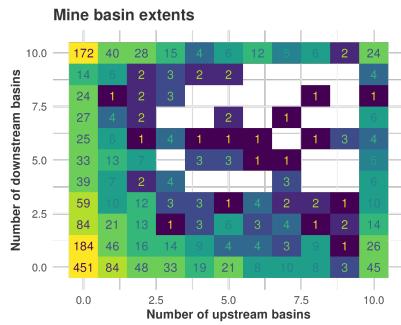


FIGURE B2: Number of mine-basins with  $Y$  upstream and  $X$  downstream basins in the dataset.

### C. Additional Tables

TABLE C1: Number and average distance of basins by order.

Order	Downstream		Upstream	
	N	Distance (km)	N	Distance (km)
0	1900	0.0	-	-
1	1162	10.7	987	14.5
2	841	22.2	865	24.2
3	695	32.9	778	34.7
4	591	43.7	738	44.7
5	531	54.4	681	55.1
6	462	64.8	593	65.9
7	418	74.3	575	75.6
8	376	85.1	530	86.6
9	343	95.9	499	95.7
10	310	106.1	452	104.2

TABLE C2: Summary statistics split by upstream/downstream status.

Upstream Basins					
Variable	N	Mean	St. Dev.	Min.	Max.
Max. EVI	53,584	0.417	0.169	0.021	0.983
Mean EVI	53,584	0.276	0.120	0.020	0.578
Max. Cropland EVI	44,389	0.459	0.127	0.057	0.990
Mean Cropland EVI	44,389	0.291	0.093	-0.002	0.637
Max. Temperature	53,584	33.83	4.003	20.00	45.10
Precipitation	53,584	905.4	606.5	0.851	3,976.0
Population	53,584	6,693.8	27,878.2	0.000	1,396,921.0
Elevation	53,584	840.5	471.2	10.53	3,059.7
Slope	53,584	2.295	2.256	0.086	20.91
Accessibility	53,584	192.0	242.3	3.000	7,542.8
Downstream Basins (incl. Mine Basins)					
Variable	N	Mean	St. Dev.	Min.	Max.
Max. EVI	61,032	0.406	0.167	-0.112	0.993
Mean EVI	61,032	0.264	0.116	-0.112	0.563
Max. Cropland EVI	50,282	0.450	0.130	-0.112	0.981
Mean Cropland EVI	50,282	0.283	0.093	-0.114	0.734
Max. Temperature	61,032	33.78	4.085	20.00	45.40
Precipitation	61,032	862.0	605.4	0.555	4,375.3
Population	60,952	9,497.1	43,568.1	0.000	1,244,492.0
Elevation	61,032	773.1	489.1	-118.3	3,047.1
Slope	61,032	2.119	2.371	0.000	20.456
Accessibility	60,992	176.9	267.1	1.002	7,681.8

TABLE C3: Main estimation results for Order specification.

Dependent Variables: Model:	Maximum EVI			Maximum Cropland EVI		
	(1)	(2)	(3)	(4)	(5)	(6)
<i>Variables</i>						
Downstream x Order = 0	-0.0064*** (0.0014)	-0.0063*** (0.0014)	-0.0059*** (0.0013)	-0.0093*** (0.0021)	-0.0097*** (0.0021)	-0.0095*** (0.0020)
Downstream x Order = 1	-0.0060*** (0.0018)	-0.0048*** (0.0018)	-0.0057*** (0.0017)	-0.0049*** (0.0026)	-0.0027*** (0.0027)	-0.0061*** (0.0026)
Downstream x Order = 2	-0.0070*** (0.0021)	-0.0052*** (0.0021)	-0.0060*** (0.0021)	-0.0042*** (0.0028)	-0.0040*** (0.0029)	-0.0042*** (0.0030)
Downstream x Order = 3	-0.0094*** (0.0023)	-0.0069*** (0.0023)	-0.0083*** (0.0022)	-0.0049*** (0.0032)	-0.0049*** (0.0033)	-0.0069*** (0.0033)
Downstream x Order = 4	-0.0071*** (0.0025)	-0.0053*** (0.0024)	-0.0059*** (0.0024)	-0.0027*** (0.0034)	-0.0036*** (0.0036)	-0.0044*** (0.0036)
Downstream x Order = 5	-0.0077*** (0.0028)	-0.0052*** (0.0026)	-0.0056*** (0.0026)	-0.0009*** (0.0037)	-0.0013*** (0.0038)	-0.0018*** (0.0039)
Downstream x Order = 6	-0.0084*** (0.0031)	-0.0054*** (0.0028)	-0.0056*** (0.0028)	-0.0042*** (0.0039)	-0.0044*** (0.0041)	-0.0051*** (0.0041)
Downstream x Order = 7	-0.0095*** (0.0033)	-0.0063*** (0.0031)	-0.0063*** (0.0030)	0.0008 (0.0041)	0.0003 (0.0043)	-2.53 × 10 <sup>-5</sup> (0.0044)
Downstream x Order = 8	-0.0140*** (0.0036)	-0.0110*** (0.0036)	-0.0109*** (0.0036)	-0.0074*** (0.0036)	-0.0085*** (0.0043)	-0.0090*** (0.0044)
Downstream x Order = 9	-0.0147*** (0.0035)	-0.0095*** (0.0034)	-0.0097*** (0.0034)	-0.0032*** (0.0039)	-0.0040*** (0.0043)	-0.0052*** (0.0044)
Downstream x Order = 10	-0.0107*** (0.0037)	-0.0056*** (0.0037)	-0.0056*** (0.0037)	-0.0038*** (0.0045)	-0.0038*** (0.0049)	-0.0043*** (0.0050)
Elevation	-7.77 × 10 <sup>-6</sup> (6.08 × 10 <sup>-6</sup> )	-2.3 × 10 <sup>-5***</sup> (6.29 × 10 <sup>-6</sup> )	-2.3 × 10 <sup>-5***</sup> (7.19 × 10 <sup>-6</sup> )	-1.59 × 10 <sup>-5***</sup> (7.35 × 10 <sup>-6</sup> )	-3.86 × 10 <sup>-5***</sup> (7.023 × 10 <sup>-6</sup> )	
Slope	0.0034*** (0.0005)	0.0033*** (0.0005)	0.0033*** (0.0006)	0.0023*** (0.0006)	0.0023*** (0.0006)	
Yearly Max. Temperature		-0.0053*** (0.0007)		-0.0071*** (0.0007)		
Yearly Precipitation		3.33 × 10 <sup>-5***</sup> (3.61 × 10 <sup>-6</sup> )		2.86 × 10 <sup>-5***</sup> (3.95 × 10 <sup>-6</sup> )		
Accessibility in 2015		-9.97 × 10 <sup>-6</sup> (5.28 × 10 <sup>-6</sup> )		-3.78 × 10 <sup>-6</sup> (1.18 × 10 <sup>-5</sup> )		
Population in 2015		-1.51 × 10 <sup>-7***</sup> (2.75 × 10 <sup>-8</sup> )		-1.06 × 10 <sup>-7***</sup> (2.04 × 10 <sup>-8</sup> )		
Sample Mean Effect	-1.567	-1.531	-1.438	-2.042	-2.127	-2.089
<i>Fixed-effects</i>						
Year	Yes	Yes	Yes	Yes	Yes	Yes
Mine	Yes	Yes	Yes	Yes	Yes	Yes
<i>Fit statistics</i>						
Observations	114,616	114,616	114,496	94,671	94,671	94,604
R <sup>2</sup>	0.91808	0.92156	0.92395	0.77981	0.78184	0.78597
Within R <sup>2</sup>	0.00393	0.04627	0.05582	0.00180	0.01099	0.02531

*Clustered (Mine) standard-errors in parentheses*  
*Signif. Codes: \*\*\*: 0.01, \*\*: 0.05, \*: 0.1*  
Note: Table shows results for estimation of Equation 1, with distance included as measured by the ordering of basins with respect to the mine basin. Columns (1)–(3) hold results from models for the overall EVI as proxy measure for vegetative health within basins, columns (4)–(6) for the cropland-specific EVI as proxy measure for agricultural productivity. Models in columns (1) and (4) include no additional covariates, models (2) and (5) control for geophysical variables (elevation and slope), models (3) and (6) additionally control for meteorological (yearly sum of precipitation and yearly maximum temperature) and socioeconomic (accessibility to city in minutes and total population in 2015) conditions. All models include mine and year fixed effects. Standard errors are clustered at the mine basin system level.

TABLE C4: Main estimation results for Distance specification.

Dependent Variables:	Maximum EVI			Maximum Cropland EVI		
Model:	(1)	(2)	(3)	(4)	(5)	(6)
<i>Variables</i>						
Downstream	-0.0065*** (0.0023)	-0.0060*** (0.0021)	-0.0058*** (0.0021)	-0.0086*** (0.0029)	-0.0088*** (0.0029)	-0.0087*** (0.0028)
Downstream × Distance	-2.02 × 10 <sup>-5</sup> (-0.0001)	1.05 × 10 <sup>-5</sup> (0.0001)	-2.02 × 10 <sup>-5</sup> (0.0001)	0.0003** (0.0000)	0.0002* (0.0000)	0.0002 (0.0001)
Downstream × Distance <sup>2</sup>	-3.98 × 10 <sup>-7</sup> (9.17 × 10 <sup>-7</sup> )	-3.37 × 10 <sup>-7</sup> (7.25 × 10 <sup>-7</sup> )	-3.98 × 10 <sup>-8</sup> (7.19 × 10 <sup>-8</sup> )	-2.15 × 10 <sup>-6**</sup> (1.06 × 10 <sup>-6</sup> )	-2.34 × 10 <sup>-6**</sup> (1.03 × 10 <sup>-6</sup> )	-1.47 × 10 <sup>-6</sup> (1.03 × 10 <sup>-6</sup> )
Distance	4.05 × 10 <sup>-5</sup> (9.03 × 10 <sup>-5</sup> )	2.98 × 10 <sup>-5</sup> (8.4 × 10 <sup>-5</sup> )	2.56 × 10 <sup>-5</sup> (8.19 × 10 <sup>-5</sup> )	-7.01 × 10 <sup>-5</sup> (0.0001)	-5.62 × 10 <sup>-5</sup> (0.0001)	-4.6 × 10 <sup>-5</sup> (0.0001)
Distance <sup>2</sup>	-1.87 × 10 <sup>-7</sup> (6.27 × 10 <sup>-7</sup> )	-9.18 × 10 <sup>-8</sup> (5.68 × 10 <sup>-7</sup> )	2.1 × 10 <sup>-8</sup> (5.56 × 10 <sup>-7</sup> )	6.93 × 10 <sup>-7</sup> (8.38 × 10 <sup>-7</sup> )	8 × 10 <sup>-7</sup> (8.23 × 10 <sup>-7</sup> )	6.08 × 10 <sup>-7</sup> (8.22 × 10 <sup>-7</sup> )
Elevation	-7.45 × 10 <sup>-4</sup> (6.56 × 10 <sup>-6</sup> )	-2.22 × 10 <sup>-4**</sup> (6.1 × 10 <sup>-6</sup> )	-	-1.83 × 10 <sup>-4**</sup> (7.55 × 10 <sup>-6</sup> )	-1.03 × 10 <sup>-5**</sup> (7.61 × 10 <sup>-6</sup> )	-
Slope	0.0034*** (0.0005)	0.0034*** (0.0005)	0.0034*** (0.0005)	0.0023*** (0.0006)	0.0023*** (0.0006)	0.0023*** (0.0006)
Yearly Max. Temperature	-0.0053*** (0.0007)	-0.0053*** (0.0007)	-0.0053*** (0.0007)	-0.0070*** (0.0007)	-0.0070*** (0.0007)	-0.0070*** (0.0007)
Yearly Precipitation	-	3.33 × 10 <sup>-5***</sup> (3.6 × 10 <sup>-6</sup> )	-	2.88 × 10 <sup>-5***</sup> (3.91 × 10 <sup>-6</sup> )	-	-
Accessibility in 2015	-	-1.01 × 10 <sup>-5*</sup> (5.31 × 10 <sup>-6</sup> )	-	-4.03 × 10 <sup>-6</sup> (1.19 × 10 <sup>-5</sup> )	-	-
Population in 2015	-	-1.51 × 10 <sup>-7***</sup> (2.77 × 10 <sup>-8</sup> )	-	-1.06 × 10 <sup>-7***</sup> (2.03 × 10 <sup>-8</sup> )	-	-
<i>Fixed-effects</i>						
Year	Yes	Yes	Yes	Yes	Yes	Yes
Mine	Yes	Yes	Yes	Yes	Yes	Yes
<i>Fit statistics</i>						
Observations	114,616	114,616	114,496	94,671	94,671	94,604
R <sup>2</sup>	0.91804	0.92152	0.92300	0.77071	0.78175	0.78587
Within R <sup>2</sup>	0.00346	0.04573	0.05524	0.00138	0.01069	0.02485

Clustered (Mine) standard-errors in parentheses

Signif. Codes: \*\*\*, \*\*, \*, .05, .01

Note: Table shows results of estimation of Equation 1, with squared distance included as measured in kilometers along the river network. Columns (1)–(3) hold results from models for the overall EVI as proxy measure for vegetative health within basins, columns (4)–(6) for the cropland-specific EVI as proxy measure for agricultural productivity. Models in column (1) and (4) include no additional covariates, models (2) and (5) control for geophysical variables (elevation and slope), models (3) and (6) additionally control for meteorological (yearly sum of precipitation and yearly maximum temperature) and socioeconomic (accessibility to city in minutes and total population in 2015) conditions. All models include mine and year fixed effects. Standard errors are clustered at the mine basin system level.

## C1. Heterogeneity analysis

### C1.1. Mine size and growth

TABLE C5: Estimation results for heterogeneity analysis: Mine characteristics

Mine Size	> 0.5km <sup>2</sup>	> 1km <sup>2</sup>	> 2.5km <sup>2</sup>	> 0%	> 10%	> 25%	
Mine Growth	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Maximum EVI							
<i>Model:</i>							
Downstream x Order = 0	-0.0059*** (0.0013)	-0.0098*** (0.0018)	-0.0114*** (0.0020)	-0.0151*** (0.0026)	-0.0073*** (0.0018)	-0.0062*** (0.0020)	-0.0060*** (0.0021)
Downstream x Order = 1	-0.0057*** (0.0017)	-0.0078*** (0.0024)	-0.0090*** (0.0026)	-0.0094*** (0.0033)	-0.0064*** (0.0024)	-0.0043 (0.0024)	-0.0044 (0.0028)
Downstream x Order = 2	-0.0066*** (0.0021)	-0.0069** (0.0027)	-0.0063** (0.0029)	-0.0064* (0.0036)	-0.0060** (0.0027)	-0.0035 (0.0031)	-0.0032 (0.0032)
<i>Fit statistics</i>							
Observations	114,496	58,696	46,440	30,128	59,080	44,376	34,280
R <sup>2</sup>	0.92305	0.93068	0.93189	0.93393	0.90491	0.91660	0.91505
Within R <sup>2</sup>	0.05582	0.05630	0.06180	0.07193	0.05844	0.05212	0.05557
Maximum Cropland EVI							
Downstream x Order = 0	-0.0095*** (0.0020)	-0.0121*** (0.0029)	-0.0148*** (0.0031)	-0.0165*** (0.0041)	-0.0110*** (0.0027)	-0.0099*** (0.0031)	-0.0084*** (0.0034)
Downstream x Order = 1	-0.0061** (0.0026)	-0.0067* (0.0038)	-0.0100** (0.0042)	-0.0093 (0.0057)	-0.0051 (0.0036)	-0.0006 (0.0041)	0.0021 (0.0043)
Downstream x Order = 2	-0.0062** (0.0030)	-0.0059 (0.0041)	-0.0061 (0.0044)	-0.0062 (0.0059)	-0.0059 (0.0038)	0.0006 (0.0043)	0.0036 (0.0046)
<i>Fit statistics</i>							
Observations	94,604	49,492	39,462	25,265	51,191	37,151	28,621
R <sup>2</sup>	0.78597	0.80545	0.81184	0.81068	0.75736	0.76813	0.74775
Within R <sup>2</sup>	0.02531	0.02960	0.03412	0.04540	0.02771	0.03095	0.03355
<i>Fixed-effects</i>							
Year	Yes						
Mine	Yes						

Clustered (Mine) standard-errors in parentheses

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

Note: Table shows results for estimation of [Equation 1](#), with distance included as measured by the ordering of basins with respect to the mine basin, with the overall EVI as outcome in the upper panel and the cropland-specific EVI as outcome in the lower panel. Model in column (1) reports results for the baseline specification, models in columns (2)–(4) for subsets of mine basins with increasing total area of mined area, models in columns (5)–(7) for subsets of mine basins with increasing growth in mined area in the period from 2017 to 2023 based on [Sépin et al., 2024](#). All specifications include the full set of controls and mine and year fixed effects. Standard errors are clustered at the mine basin system level.

**C1.2. Biome and region**

TABLE C6: Estimation results for heterogeneity analysis: Biome and Region

Biome Region Model:	Deserts	Forests	Grasslands	N. & E. Africa	S. Africa	W. Africa
	(1)	(2)	(3)	(4)	(5)	(7)
Maximum EVI						
Downstream x Order = 0	-0.0050*** (0.0013)	-0.0034 (0.0025)	-0.0013 (0.0032)	-0.0069*** (0.0017)	-0.0029 (0.0017)	-0.0049*** (0.0018)
Downstream x Order = 1	-0.0057*** (0.0017)	-0.0015 (0.0033)	-0.0054 (0.0046)	-0.0067*** (0.0021)	-0.0020 (0.0046)	-0.0033 (0.0022)
Downstream x Order = 2	-0.0066*** (0.0021)	-0.0008 (0.0037)	-0.0085 (0.0055)	-0.0075*** (0.0026)	-0.0013 (0.0055)	-0.0045* (0.0026)
<i>Fit statistics</i>						
Observations	114,496	20,408	16,872	77,216	12,456	72,032
R <sup>2</sup>	0.92395	0.89783	0.92623	0.84199	0.94780	0.89978
Within R <sup>2</sup>	0.05582	0.05791	0.06954	0.05889	0.10409	0.05297
Maximum Cropland EVI						
Downstream x Order = 0	-0.0095*** (0.0020)	-0.0124 (0.0109)	-0.0049 (0.0051)	-0.0088*** (0.0021)	-0.0073 (0.0073)	-0.0107*** (0.0027)
Downstream x Order = 1	-0.0061** (0.0026)	-0.0162 (0.0158)	0.0005 (0.0064)	-0.0059** (0.0027)	-0.0064 (0.0078)	-0.0016 (0.0034)
Downstream x Order = 2	-0.0062** (0.0030)	-0.0050 (0.0175)	-0.0036 (0.0083)	-0.0066** (0.0030)	-0.0015 (0.0088)	-0.0006 (0.0040)
<i>Fit statistics</i>						
Observations	94,604	8,297	15,083	71,224	9,720	58,882
R <sup>2</sup>	0.78597	0.69727	0.85804	0.71390	0.86094	0.72147
Within R <sup>2</sup>	0.02531	0.06634	0.03601	0.02846	0.05931	0.02276
<i>Fixed-effects</i>						
Year	Yes	Yes	Yes	Yes	Yes	Yes
Mine	Yes	Yes	Yes	Yes	Yes	Yes

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

Note: Table shows results for estimation of [Equation 1](#), with distance included as measured by the ordering of basins with respect to the mine basin, with the overall EVI as outcome in the upper panel and the cropland-specific EVI as outcome in the lower panel. Model in column (1) reports results for the full sample, models in columns (2)–(4) for sample splits by primary biome of mine basin system, and models in columns (5)–(7) for sample splits by regions based on the USDA crop classifications. All specifications include the full set of controls and mine and year fixed effects. Standard errors are clustered at the mine basin system level.

## C2. Robustness analysis

### C2.1. Sample variations

TABLE C7: Estimation results for varying sample definition

Dependent Variables:	(1)	(2)	(3)	(4)	(5)	(6)	Maximum Cropland EVI			(7)	(8)	(9)	(10)
<i>Variables:</i>													
Downstream x Order = 0	-0.0059*** (0.0013)	-0.0070*** (0.0014)	-0.0062*** (0.0012)				-0.0095*** (0.0020)	-0.0082*** (0.0024)	-0.0094*** (0.0022)				
Downstream x Order = 1	-0.0057*** (0.0017)	-0.0053*** (0.0017)	-0.0053*** (0.0017)	0.0049** (0.0020)	0.0051** (0.0021)	0.0051** (0.0021)	-0.0061** (0.0026)	-0.0049 (0.0026)	-0.0051** (0.0030)	-0.0061** (0.0029)	-0.0069* (0.0039)		
Downstream x Order = 2	-0.0060*** (0.0021)	0.0041** (0.0026)		-0.0056** (0.0023)	0.0060** (0.0023)	0.0060** (0.0023)	-0.0067** (0.0030)	-0.0067** (0.0037)	-0.0067** (0.0033)	-0.0062* (0.0033)			
<i>Fixed effects</i>													
Year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	
Mine	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	
<i>Fit statistics</i>													
Observations	114,496	61,712	32,360	99,320	9,168	94,604	50,914	27,280	81,275	7,623			
R <sup>2</sup>	0.9239	0.9156	0.9099	0.9202	0.93378	0.78597	0.76613	0.84832	0.78332	0.81766			
Within R <sup>2</sup>	0.0582	0.05702	0.05650	0.05511	0.07964	0.02531	0.02382	0.03446	0.02322	0.03884			

Clustered (Mine) standard errors in parentheses

Signif. Codes: \*\*\*, \*\*, \*, .05, .01, .1

Note: Table shows results for estimation of Equation 1, with distance included as measured by the ordering of basins with respect to the mine basin. Columns (1)–(5) hold results from models for the overall EVI as proxy measure for vegetative health within basins, columns (6)–(10) for the cropland-specific EVI as proxy measure for agricultural productivity. Models in columns (1) and (6) are the baseline specification, models in columns (2) and (7) only include basin systems with at least one up- and downstream basin, models in columns (3) and (8) include maximum order one up- and downstream basins, models in columns (4) and (9) exclude the mine basin itself, models in columns (5) and (10) include only the first order basins of basins systems with at least one basin up- and downstream and excludes the mine basin. All specifications include the full set of controls and mine and year fixed effects. Standard errors are clustered at the mine basin system level.

## C2.2. Outcome and fixed effects variations

TABLE C8: Estimation results for varying fixed effects and outcome variables

Dependent Variable:	Maximum EVI	Mean EVI	Maximum Cropland EVI	Mean C EVI	ESA C EVI				
Model:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
<i>Variables</i>									
Downstream x Order = 0	-0.0059*** (0.0013)	0.0065*** (0.0013)	0.0079*** (0.0014)	-0.0048*** (0.0009)	0.0065*** (0.0020)	-0.0104*** (0.0020)	0.0109*** (0.0021)	0.0073*** (0.0013)	-0.0048* (0.0036)
Downstream x Order = 1	-0.0057*** (0.0017)	-0.0069*** (0.0016)	-0.0066*** (0.0017)	-0.0035*** (0.0011)	-0.0061*** (0.0026)	-0.0062*** (0.0025)	-0.0064*** (0.0025)	-0.0043*** (0.0017)	-0.0035 (0.0032)
Downstream x Order = 2	-0.0066*** (0.0021)	-0.0064*** (0.0020)	-0.0067*** (0.0013)	-0.0038*** (0.0030)	-0.0062*** (0.0029)	-0.0058*** (0.0028)	-0.0064*** (0.0019)	-0.0055*** (0.0035)	-0.0015 (0.0035)
<i>Fixed-effects</i>									
Year	Yes	Yes							
Mine	Yes	Yes							
Pfaffstetter basin level 8									
Pfaffstetter basin level 6			Yes			Yes			
<i>Fit statistics</i>									
Observations	114,496	114,496	114,496	114,496	94,604	94,604	94,604	94,604	67,649
R <sup>2</sup>	0.9293	0.91954	0.96419	0.95707	0.7850	0.77091	0.74193	0.8844	0.80154
Within R <sup>2</sup>	0.05582	0.06500	0.08647	0.11783	0.02531	0.02537	0.04285	0.04478	0.02553

Clustered (Mine) standard errors in parentheses

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

Note: Table shows results for estimation of Equation 1, with distance included as measured by the ordering of basins with respect to the mine basin. Columns (1)–(4) hold results from models for the overall EVI as proxy measure for vegetative health within basins, columns (5)–(9) for the cropland-specific EVI as proxy measure for agricultural productivity. Models in columns (1) and (4) are the baseline specification for the overall maximum EVI and the cropland-specific maximum EVI, respectively, with mine fixed effects. Models in columns (2) and (5) use fixed effects at Pfaffstetter level 8 basins, models in columns (3) and (6) fixed effects at Pfaffstetter level 6 basins. Models in columns (4) and (8) report results for the yearly mean of the overall EVI and the cropland-specific EVI instead of the maximum, respectively. Model in column (9) reports result for the cropland-specific EVI based on the time-varying cropland mask retrieved from Defourny et al., 2024. All specifications include the full set of controls. Standard errors are clustered at the mine basin system level.

### C2.3. Automatic Bandwidth Selection

Table C9: Distance specification using automatic bandwidth selection

	Max EVI	Max C EVI		
No Controls				
Conventional	-0.0050*** (0.0015)	-0.0069*** (0.0019)	-0.0112*** (0.0020)	-0.0112*** (0.0025)
Bias-Corrected	-0.0056*** (0.0015)	-0.0073*** (0.0019)	-0.0118*** (0.0020)	-0.0116*** (0.0025)
Observations	37880	55072	32813	52964
Bandwidth (conv)	20.3	34.9	20.7	41.7
Bandwidth (bias)	46.4	67.9	47.4	82.6
With full Controls				
Conventional	-0.0045*** (0.0015)	-0.0055*** (0.0019)	-0.0100*** (0.0020)	-0.0115*** (0.0026)
Bias-Corrected	-0.0049*** (0.0015)	-0.0059*** (0.0019)	-0.0105*** (0.0020)	-0.0118*** (0.0026)
Observations	38200	53384	32629	49968
Bandwidth (conv)	20.6	33.3	20.5	38.4
Bandwidth (bias)	43.4	63.3	45.4	73.7
<i>Settings</i>				
Kernel	Triangular	Triangular	Triangular	Triangular
BW.Criterion	mserd	mserd	mserd	mserd
Polynomial	Linear	Quadratic	Linear	Quadratic

*Clustered (Mine) standard-errors in parentheses*

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

Note: Table shows results for estimation of [Equation 1](#), 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.

Table C10: Order specification automatic bandwidth selection

	Max EVI	Max C EVI		
No Controls				
I(order>0)	-0.0048 (0.0013)	-0.0048 (0.0019)	-0.0090*** (0.0018)	-0.0090** (0.0030)
Observations	45613	45613	38537	38537
Bandwidth	2	2	2	2
With full Controls				
I(order>0)	-0.0048** (0.0012)	-0.0048 (0.0018)	-0.0090*** (0.0017)	-0.0090*** (0.0029)
Observations	45580	45580	38504	38504
Bandwidth	2	2	2	2
<i>Settings</i>				
Kernel	Triangular	Triangular	Triangular	Triangular
BW.Criterion	MSE	MSE	MSE	MSE
Cluster SEs	No	Mine basin	No	Mine basin

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

Note: Table shows results for estimation of [Equation 1](#), 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.

## C2.4. Placebo Outcomes

TABLE C11: Estimation results using covariates as placebo outcomes

Dependent Variables:	Elevation (1)	Slope (2)	Max. Temp (3)	Precipitation (4)	Accessibility in 2015 (5)	Population in 2015 (6)
<i>Variables</i>						
Downstream	-6.852 (8.509)	-0.0538 (0.0912)	-0.0137 (0.0567)	0.6025 (3.934)	-5.427 (5.531)	2,125.7 (1,589.8)
Distance × Downstream	-5.006 (0.0041)	-0.0060 (0.0040)	0.0135** (0.0130)	0.1942 (0.2278)	0.0839 (0.2278)	182.9** (58.0)
Distance <sup>2</sup> × Downstream	0.0043 (0.0039)	-8.25 × 10 <sup>-6</sup> (4.01 × 10 <sup>-5</sup> )	2.32 × 10 <sup>-6</sup> (3.26 × 10 <sup>-5</sup> )	0.0003 (0.0020)	0.0004 (0.0020)	1.081*** (0.9463)
Distance	2.326*** (0.4215)	0.0025 (0.0039)	-0.0067** (0.0032)	0.0879 (0.2129)	0.7557*** (0.2587)	-54.72 (45.17)
Distance <sup>2</sup>	0.0005 (0.0033)	1.12 × 10 <sup>-4</sup> (3.49 × 10 <sup>-5</sup> )	-5.34 × 10 <sup>-6</sup> (3.1 × 10 <sup>-5</sup> )	-0.0005 (0.0015)	-0.0013 (0.0021)	0.3439 (0.2724)
<i>Fixed-effects</i>						
Year	Yes	Yes	Yes	Yes	Yes	Yes
Mine	Yes	Yes	Yes	Yes	Yes	Yes
<i>Fit statistics</i>						
Observations	114,616	114,616	114,616	114,616	114,576	114,536
R <sup>2</sup>	0.95627	0.70192	0.95579	0.96187	0.88768	0.59121
Within R <sup>2</sup>	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

Note: Table shows results for estimation of [Equation 1](#), with distance included as measured by the ordering of basins with respect to the mine basin using the additionally used covariates as placebo outcomes for the full sample. All specifications control for the remaining covariates except the one used as placebo outcome as well as mine and year fixed effects. Standard errors are clustered at the mine basin system level.

# Mines – Rivers – Yields

## Downstream Mining Impacts on Agriculture in Africa

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

Institute for Ecological Economics Seminar

November 13, 2024

# Snapshot

Background

Methods and Data

Results

2/58

## Mines – Curse or Blessing?

### A Blessing?

- More **minerals** are needed for the **green transition**.
- **Africa** has large reserves of some of these minerals.
- Extraction entails **local economic opportunities**.

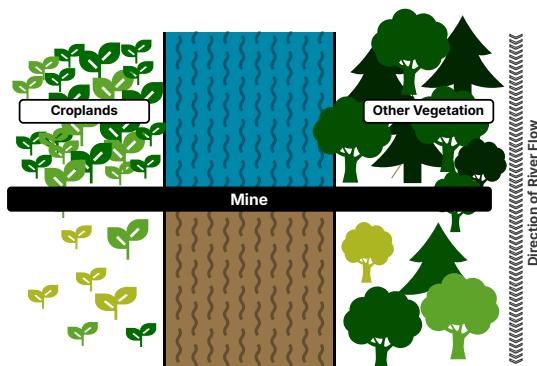
### A Curse?

- Resource extraction causes **negative externalities**.
- **Ecological** effects are well-documented and can impact livelihoods.
- Ideally, we can inform **policy** in order to mitigate them.

3/58

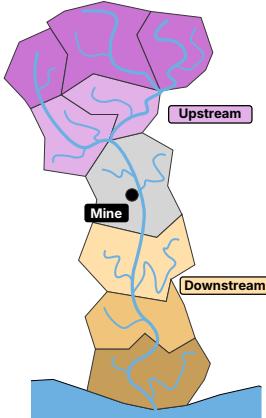
## How Pollution Travels

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



4/58

## How to Find Downstream Areas



Using data on **river basins** (Lehner & Grill, 2013), we know where water flows from a given location.

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

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

5/58

Snapshot

## Background

Methods and Data

Results

6/58

## Economic Benefits from Mines

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

7/58

## Mines' Negative Externalities

- Mines cause a flurry of **negative externalities**.
- Mines negatively impact the **institutional environment**, causing
  - **conflict** (Berman et al., 2017) and
  - **corruption** (Knutson et al., 2016).
- An especially extensive cluster of externalities revolves around **pollution**.

8/58

### Mines' Negative Externalities: Pollution

- Mines **use water** and produce **sediments and tailings** (Moura et al., 2022).
- Pollutants include **mercury** and **lead** (Schwarzenbach et al., 2010).
- 23 million people live in **polluted river basins** (Macklin et al., 2023).
- Industrial pollution **harms plant growth** (Yang et al., 2021).

**Mines pollute water, this water travels, and it can harm plants.**

9/58

## 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** (ICMM, 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**.

10/58

Snapshot

Background

## Methods and Data

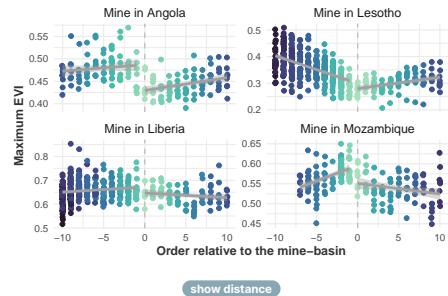
Results

11/58

## Intuition

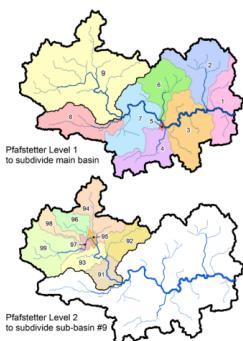
The **four mines** depicted give an intuition for what we expect.

Following the river “flow” from left to right, we can see **discontinuities at the mine basin**.



12/58

## Basins



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**.

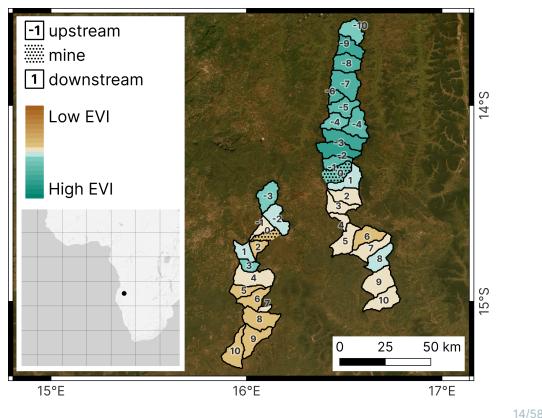
Illustration from Lehner and Grill (2013)

13/58

## Mines

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

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



14/58

## A Remotely Sensed Outcome Measure

- We cannot use official **agricultural production statistics** due to a
  - lack of spatial **granularity** and the
  - **institutional differences** across countries.
- Instead, we use the **Enhanced Vegetation Index (EVI)**, which
  - is **remotely sensed** as the difference between red and near infrared light,
  - ranges **between -1** (water) **and 1** (dense vegetation), and
  - can be seen as a measure of **vegetation greenness**.

15/58

## 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** (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).

16/58

## Observations

- We observe  $N = 14,327$  basins over a period of  $T = 8$  years:

- **6,698 upstream** basins,
- **1,900 mine** basins, and
- **5,729 downstream** basins.

[show order × up-/downstream numbers](#)

- In addition to treatment and outcome, we observe **covariates** concerning:

- **topography** (elevation and slope),
- **soil type**,
- **climate** (precipitation, temperature), and
- **socioeconomic characteristics** (population, accessibility).

[show balance](#)

17/58

## Summary Statistics

Variable	N	Mean	St. Dev.	Min.	Max.
<b>Max. EVI</b>	114,616	0.411	0.168	-0.112	0.993
<b>Mean EVI</b>	114,616	0.270	0.118	-0.112	0.578
<b>Max. Cropland EVI</b>	94,671	0.454	0.129	-0.112	0.990
<b>Mean Cropland EVI</b>	94,671	0.286	0.093	-0.114	0.734
<b>Max. Temperature</b>	114,616	33.80	4.047	20.00	45.40
<b>Precipitation</b>	114,616	882.3	606.3	0.555	4,375.3
<b>Population</b>	114,536	8,185	37,090	0.000	1,396,921
<b>Elevation</b>	114,616	804.6	482.0	-118.3	3,059.7
<b>Slope</b>	114,616	2.201	2.320	0.000	20.92
<b>Accessibility</b>	114,576	183.9	255.9	1.002	7,681

[show balance](#)

18/58

## Empirical Strategy

We employ a **quasi-experimental regression discontinuity design (RDD)**:

$$y_{ijt} = \beta_1 d_{ij} + \beta_2 d_{ij} \times \text{downstream}_j + \beta_3 \text{downstream}_j + \delta' \mathbf{x}_{it} + \mu_j + \psi_t + \varepsilon_{ijt},$$

- $y_{ijt}$ : **Outcome** of basin  $i$  near mine  $j$  in year  $t$ ,
- $d_{ij}$ : **Distance** to nearest mine,
  - either the **basin's order** relative to the mine,
  - or kilometers along the **river stream**.
- $\mu_j$ ,  $\psi_t$ : Mine and year **fixed effects**,
- $\mathbf{x}_{it}$ : Basin specific **covariates**:
  - topographic, climate, soil type, socioeconomic,

19/58

## Identification

$$y_{ijt} = \beta_1 d_{ij} + \beta_2 d_{ij} \times \text{downstream}_j + \beta_3 \text{downstream}_j + \delta' \mathbf{x}_{it} + \mu_j + \psi_t + \varepsilon_{ijt},$$

- Parameter  $\beta_3$  is identified under the assumption that there are no **other discontinuous changes** at the mine basin.
- To assess the **validity** of this assumption, we
  - check **balance** of up- and downstream basins,
  - include meteorological, geographical, and socioeconomic **controls**, and
  - use plausibly unaffected covariates as **placebo** outcomes .

20/58

Snapshot

Background

Methods and Data

## Results

21/58

### Results Overview

- We find a **significant reduction** in vegetation health **downstream** of mines.
- The magnitude of this effect is **greater on croplands**.
- Impacts **dissipate slowly** the farther we move from a mine.
- These results are **robust** to varying the sample, the outcome measurement, and the level of fixed effects.

22/58

## Order Specification Results (1)

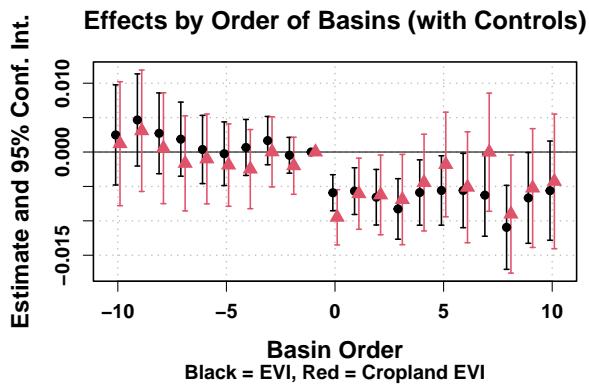
	<b>Max. EVI</b> (Plain)	<b>Max. EVI</b> (Full)	<b>Max. Cropland EVI</b> (Plain)	<b>Max. Cropland EVI</b> (Full)
<b>Order</b>				
Mine-basin (0 <sup>th</sup> )	-0.0064*** (0.0014)	<b>-0.0059***</b> (0.0013)	-0.0093*** (0.0021)	<b>-0.0095***</b> (0.0020)
Downstream (1 <sup>st</sup> )	-0.0060*** (0.0018)	<b>-0.0057***</b> (0.0017)	-0.0049* (0.0026)	<b>-0.0061**</b> (0.0026)
Downstream (2 <sup>nd</sup> )	-0.0070*** (0.0021)	<b>-0.0066***</b> (0.0021)	-0.0042 (0.0028)	<b>-0.0062**</b> (0.0030)
Sample mean	0.412	0.412	0.454	0.454
Observations	114,616	114,496	94,671	94,604
R <sup>2</sup>	0.912	0.924	0.780	0.786
Controls	No	Yes	No	Yes
Year F.E.	Yes	Yes	Yes	Yes
Mine F.E.	Yes	Yes	Yes	Yes

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

Significance levels: \*\*\*: 0.01, \*\*: 0.05, \*: 0.1. [show full results](#)

23/58

## Order Specification Results (2)



24/58

## Order Specification Results (3)

- We can see that **upstream** basins are unaffected, while **downstream** basins experience a significant negative effect.
- At the sample mean, the effect for the
  - Max. EVI** corresponds to an EVI reduction of **1.4%**.
  - Max. Cropland EVI** corresponds to an EVI reduction of **2.1%**.
- The effect **persists** beyond the mine basin.
- At higher order basins, impacts become imprecise.

25/58

## Distance Specification Results

	<i>Max. EVI</i> (Plain)	<i>Max. EVI</i> (Full)	<i>Max. Cropland EVI</i> (Plain)	<i>Max. Cropland EVI</i> (Full)
<b>Distance</b>				
Downstream	-0.0065*** (0.0023)	-0.0058*** (0.0021)	-0.0086*** (0.0029)	-0.0087*** (0.0028)
Downstream × Distance	-2.0 × 10 <sup>-5</sup> (0.0001)	-2.0 × 10 <sup>-5</sup> (0.0001)	0.0003** (0.0001)	0.0002 (0.0001)
Downstream × Distance <sup>2</sup>	-4.0 × 10 <sup>-7</sup> (9.2 × 10 <sup>-7</sup> )	-9.8 × 10 <sup>-8</sup> (7.2 × 10 <sup>-7</sup> )	-2.2 × 10 <sup>-6</sup> ** (1.1 × 10 <sup>-6</sup> )	-1.9 × 10 <sup>-6</sup> * (1.0 × 10 <sup>-6</sup> )
Sample mean	0.412	0.412	0.454	0.454
Observations	114,616	114,496	94,671	94,604
R <sup>2</sup>	0.918	0.924	0.780	0.786
Controls	No	Yes	No	Yes
Year F.E.	Yes	Yes	Yes	Yes
Mine F.E.	Yes	Yes	Yes	Yes

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

Significance levels: \*\*\*: 0.01, \*\*: 0.05, \*: 0.1. [show full results](#)

26/58

## Impact Decay (1)

- The decay of the impact is **not accurately reflected** by linear and linear-quadratic distance operationalizations.
- Dispersal of toxic mine tailings occurs non-linearly (Macklin et al., 2023).
- We therefore re-estimate the main specification using an **exponential distance decay** function:

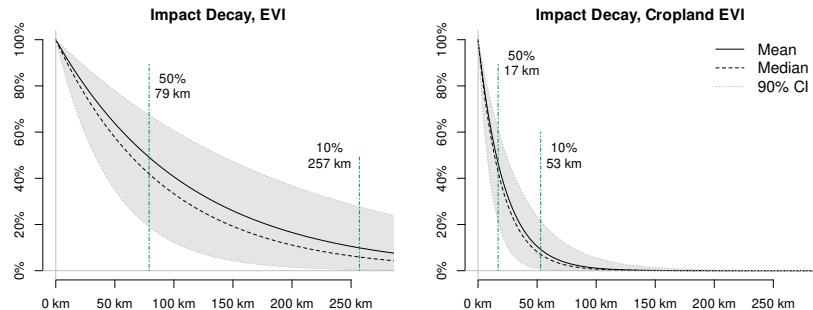
$$\exp(-\delta d_{ij}),$$

where  $d_{ij}$  is the distance along the river from a mine.

[show details](#)

27/58

## Impact Decay (2)



28/58

## Heterogeneity (1)

- We investigate heterogeneity along three main dimensions.

- **Characteristics of mine basin :**

- greater effect for larger mines,
- no differential effect for mines that grew faster over time

- **Biome :**

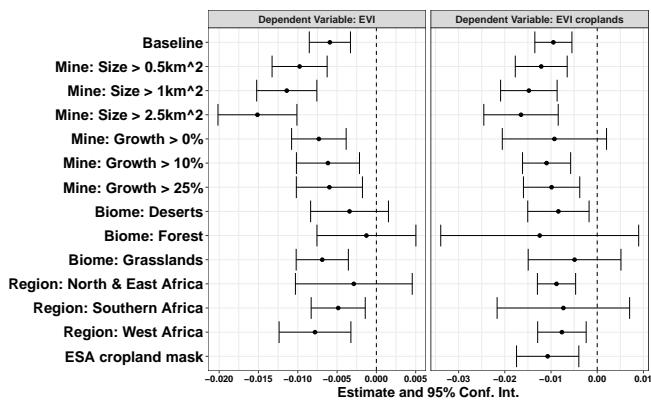
- larger effect for mines located in grasslands,
- no significant effect in deserts and forests.

- **Region :**

- greater effect in West Africa,
- smaller effect in Southern Africa,
- no significant effect in North & East Africa.

29/58

## Heterogeneity (2)



30/58

## Robustness Checks

- We check the robustness of our results in multiple ways.

- **Outcome Variable :**

- mean instead of maximum EVI, different cropland mask.

- **Sample Definition**

- **Level of Fixed Effects**

- **Estimation Methods :**

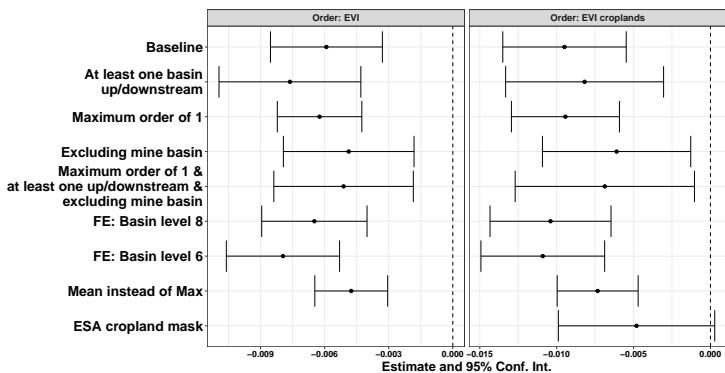
- data-driven estimation methods (Cattaneo et al., 2019),
- inducing balance using coarsened exact matching (Iacus et al., 2012).

- **Placebo Outcomes :**

- temperature, elevation, slope, precipitation, accessibility, population.

31/58

## Robustness



[show varying sample estimates](#) [show varying outcome/FE estimates](#) [show placebo estimates](#) [show automatic BW sel. estimates](#)

32/58

## Discussion

- We find **negative impacts** on vegetation health by about 1.4-2.1% at the sample mean.
- Various treatments have been found to have large effects on crop yields.
  - Land tenure reforms: +20% (Adamopoulos et al., 2024)
  - Gold mining: about -40% (Aragón & Rud, 2015)
- Studies that use remotely sensed outcome measures usually find smaller effects.
  - Institutional changes reduce crop yields by about 2% as measured through EVI (Wuepper et al., 2023)

33/58

## Relevance

- Our findings inform the **discussion** about **resource extraction**,
  - particularly in countries with **weak environmental governance**.
- There is a need to
  - tackle the lack of **containment facilities** and
  - improve **environmental legislation**,
  - both for **industrial** and **informal** mines.
- This is an especially urgent issue due to its potential to endanger **food supply**.

34/58

## Limitations and Future Alleys for Research

- Using **remotely sensed measures** helped us overcome data scarcity.
- However, they only represent crop yields **indirectly**.
- Our **treatment indicator** relied only on mine location.
- Differences in **waste management** are not accounted for, but may affect outcomes.
- **Adaptive behavior** by farmers is not covered and may attenuate results.

35/58

## 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,
  - on grasslands, and
  - in West Africa.
- The results were **robust** to various changes of treatment, outcome or sample definition.

36/58

## References (1)

- Adamopoulos, T., Brandt, L., Chen, C., Restuccia, D., & Wei, X. (2024). Land security and mobility frictions. *The Quarterly Journal of Economics*, 139(3), 1941–1987.  
<https://doi.org/10.1093/qje/qjae010>
- Aragón, F. M., & Rud, J. P. (2015). Polluting industries and agricultural productivity: Evidence from mining in Ghana. *The Economic Journal*, 126(597), 1980–2011.  
<https://doi.org/10.1111/ecoj.12244>
- 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>

37/58

## 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>
- Berman, N., Couttenier, M., Rohner, D., & Thoenig, M. (2017). This mine is mine! How minerals fuel conflicts in Africa. *American Economic Review*, 107(6), 1564–1610. <https://doi.org/10.1257/aer.20150774>
- 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., & Titunik, R. (2019, November). *A practical introduction to regression discontinuity designs: Foundations*. Cambridge University Press. <https://doi.org/10.1017/9781108684606>

38/58

## References (3)

- 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](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.
- 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>
- Iacus, S. M., King, G., & Porro, G. (2012). Causal inference without balance checking: Coarsened Exact Matching. *Political Analysis*, 20(1), 1–24. <https://doi.org/10.1093/pan/mpr013>

39/58

## References (4)

- ICMM. (2022). Mining Contribution Index (MCI) (6th Edition). [https://www.icmm.com/website/publications/pdfs/social-performance/2022/research\\_mci-6-ed.pdf?cb=16134](https://www.icmm.com/website/publications/pdfs/social-performance/2022/research_mci-6-ed.pdf?cb=16134)
- 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>
- Knutsen, C. H., Kotsadam, A., Olsen, E. H., & Wig, T. (2016). Mining and local corruption in Africa. *American Journal of Political Science*, 61(2), 320–334. <https://doi.org/10.1111/ajps.12268>
- 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>

40/58

## 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., Mudhatkal, 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>

41/58

## References (6)

- 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>
- 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>
- Wuepper, D., Wang, H., Schlenker, W., Jain, M., & Finger, R. (2023). Institutions and global crop yields. *NBER Working Paper*. <https://doi.org/10.3386/w31426>
- Yang, J., Sun, H., Qin, J., Wang, X., & Chen, W. (2021). Impacts of Cd on Temporal Dynamics of Nutrient Distribution Pattern of *Bletilla striata*, a Traditional Chinese Medicine Plant. *Agriculture*, 11(7), 594. <https://doi.org/10.3390/agriculture11070594>

42/58

## Appendix

43/58

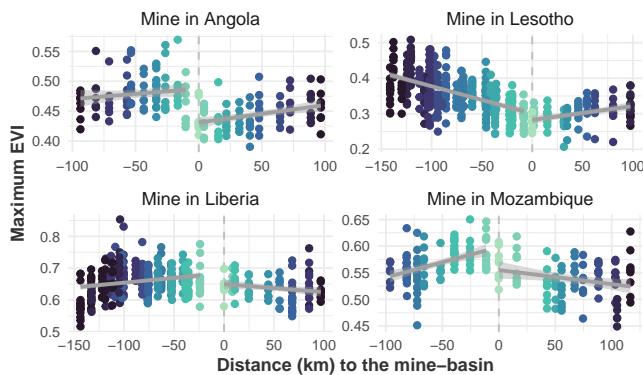
## Appendix Impact Decay Assessment

- We re-estimate our main specification using an **exponential decay** function  $\exp\{-\delta d_j\}$ .
- **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.

[go back](#)

44/58

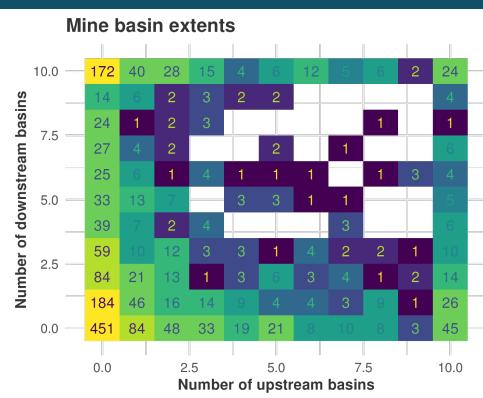
## Appendix Four Selected Mines, Distance



[go back](#)

45/58

## Appendix Basin Numbers



[go back to observations overview](#)

46/58

## Appendix Basins by Order

Order	Downstream N	Distance (km)	Upstream N	Distance (km)
0	1900	0.0	-	-
1	1162	10.7	987	14.5
2	841	22.2	865	24.2
3	695	32.9	778	34.7
4	591	43.7	738	44.7
5	531	54.4	681	55.1
6	462	64.8	593	65.9
7	418	74.3	575	75.6
8	376	85.1	530	86.6
9	343	95.9	499	95.7
10	310	106.1	452	104.2

[go back to observations overview](#)

47/58

## Appendix Summary Statistics for Upstream Basins

Upstream Basins					
Variable	N	Mean	St. Dev.	Min.	Max.
Max. EVI	53,584	0.417	0.169	0.021	0.983
Mean EVI	53,584	0.276	0.120	0.020	0.578
Max. Cropland EVI	44,389	0.459	0.127	0.057	0.990
Mean Cropland EVI	44,389	0.291	0.093	-0.002	0.637
Max. Temperature	53,584	33.83	4.003	20.00	45.10
Precipitation	53,584	905.4	606.5	0.851	3,976.0
Population	53,584	6,693.8	27,878.2	0.000	1,396,921.0
Elevation	53,584	840.5	471.2	10.53	3,059.7
Slope	53,584	2.295	2.256	0.086	20.91
Accessibility	53,584	192.0	242.3	3.000	7,542.8

[go back to covariate overview](#) [go back to summary statistics](#)

48/58

## Appendix Summary Statistics for Downstream Basins

Downstream Basins (incl. Mine Basins)					
Variable	N	Mean	St. Dev.	Min.	Max.
Max. EVI	61,032	0.406	0.167	-0.112	0.993
Mean EVI	61,032	0.264	0.116	-0.112	0.563
Max. Cropland EVI	50,282	0.450	0.130	-0.112	0.981
Mean Cropland EVI	50,282	0.283	0.093	-0.114	0.734
Max. Temperature	61,032	33.78	4.085	20.00	45.40
Precipitation	61,032	862.0	605.4	0.555	4,375.3
Population	60,952	9,497.1	43,568.1	0.000	1,244,492.0
Elevation	61,032	773.1	489.1	-118.3	3,047.1
Slope	61,032	2.119	2.371	0.000	20.456
Accessibility	60,992	176.9	267.1	1.002	7,681.8

[go back to covariate overview](#) [go back to summary statistics](#)

49/58

## Appendix Full Order Specification Results

Dependent Variables:	Maximum EVI						Maximum Cropland EVI					
Model:	(1)	(2)	(3)	(4)	(5)	(6)						
Variables												
Downstream x Order = 0	-0.0064*** (0.0014)	-0.0053*** (0.0014)	-0.0051*** (0.0014)									
Downstream x Order = 1	-0.0057*** (0.0017)	-0.0053*** (0.0017)										
Downstream x Order = 2	-0.0070*** (0.0021)	-0.0063*** (0.0021)										
Downstream x Order = 3	-0.0054*** (0.0021)	-0.0059*** (0.0021)	-0.0058*** (0.0021)									
Downstream x Order = 4	-0.0071*** (0.0025)	-0.0053*** (0.0024)										
Downstream x Order = 5	-0.0074*** (0.0028)	-0.0052*** (0.0026)										
Downstream x Order = 6	-0.0076*** (0.0031)	-0.0054*** (0.0028)										
Downstream x Order = 7	-0.0079*** (0.0033)	-0.0056*** (0.0031)										
Downstream x Order = 8	-0.0082*** (0.0033)	-0.0058*** (0.0031)										
Downstream x Order = 9	-0.0084*** (0.0033)	-0.0059*** (0.0034)										
Downstream x Order = 10	-0.0102*** (0.0037)	-0.0061*** (0.0037)										
Elevation	-2.77 × 10 <sup>-4</sup> (8.97 × 10 <sup>-5</sup> )	-2.3 × 10 <sup>-4</sup> *** (7.81 × 10 <sup>-5</sup> )	-2.3 × 10 <sup>-4</sup> *** (7.81 × 10 <sup>-5</sup> )	-2.3 × 10 <sup>-4</sup> *** (7.81 × 10 <sup>-5</sup> )	-2.3 × 10 <sup>-4</sup> *** (7.81 × 10 <sup>-5</sup> )	-2.3 × 10 <sup>-4</sup> *** (7.81 × 10 <sup>-5</sup> )	-2.3 × 10 <sup>-4</sup> *** (7.81 × 10 <sup>-5</sup> )	-2.3 × 10 <sup>-4</sup> *** (7.81 × 10 <sup>-5</sup> )	-2.3 × 10 <sup>-4</sup> *** (7.81 × 10 <sup>-5</sup> )	-2.3 × 10 <sup>-4</sup> *** (7.81 × 10 <sup>-5</sup> )	-2.3 × 10 <sup>-4</sup> *** (7.81 × 10 <sup>-5</sup> )	
Slope	0.0034*** (0.0009)	0.0033*** (0.0009)										
Yearly Max. Temperature	-0.0053*** (0.0017)											
Yearly Precipitation	3.33 × 10 <sup>-4</sup> *** (3.81 × 10 <sup>-5</sup> )	2.88 × 10 <sup>-4</sup> *** (3.81 × 10 <sup>-5</sup> )	2.88 × 10 <sup>-4</sup> *** (3.81 × 10 <sup>-5</sup> )	2.88 × 10 <sup>-4</sup> *** (3.81 × 10 <sup>-5</sup> )	2.88 × 10 <sup>-4</sup> *** (3.81 × 10 <sup>-5</sup> )	2.88 × 10 <sup>-4</sup> *** (3.81 × 10 <sup>-5</sup> )	2.88 × 10 <sup>-4</sup> *** (3.81 × 10 <sup>-5</sup> )	2.88 × 10 <sup>-4</sup> *** (3.81 × 10 <sup>-5</sup> )	2.88 × 10 <sup>-4</sup> *** (3.81 × 10 <sup>-5</sup> )	2.88 × 10 <sup>-4</sup> *** (3.81 × 10 <sup>-5</sup> )	2.88 × 10 <sup>-4</sup> *** (3.81 × 10 <sup>-5</sup> )	2.88 × 10 <sup>-4</sup> *** (3.81 × 10 <sup>-5</sup> )
Accessibility in 2015	0.97 × 10 <sup>-4</sup> (5.28 × 10 <sup>-5</sup> )	1.32 × 10 <sup>-4</sup> (5.31 × 10 <sup>-5</sup> )	1.32 × 10 <sup>-4</sup> (5.31 × 10 <sup>-5</sup> )	1.32 × 10 <sup>-4</sup> (5.31 × 10 <sup>-5</sup> )	1.32 × 10 <sup>-4</sup> (5.31 × 10 <sup>-5</sup> )	1.32 × 10 <sup>-4</sup> (5.31 × 10 <sup>-5</sup> )	1.32 × 10 <sup>-4</sup> (5.31 × 10 <sup>-5</sup> )	1.32 × 10 <sup>-4</sup> (5.31 × 10 <sup>-5</sup> )	1.32 × 10 <sup>-4</sup> (5.31 × 10 <sup>-5</sup> )	1.32 × 10 <sup>-4</sup> (5.31 × 10 <sup>-5</sup> )	1.32 × 10 <sup>-4</sup> (5.31 × 10 <sup>-5</sup> )	1.32 × 10 <sup>-4</sup> (5.31 × 10 <sup>-5</sup> )
Population in 2015	-1.51 × 10 <sup>-4</sup> *** (2.77 × 10 <sup>-5</sup> )	-1.06 × 10 <sup>-4</sup> *** (2.77 × 10 <sup>-5</sup> )	-1.06 × 10 <sup>-4</sup> *** (2.77 × 10 <sup>-5</sup> )	-1.06 × 10 <sup>-4</sup> *** (2.77 × 10 <sup>-5</sup> )	-1.06 × 10 <sup>-4</sup> *** (2.77 × 10 <sup>-5</sup> )	-1.06 × 10 <sup>-4</sup> *** (2.77 × 10 <sup>-5</sup> )	-1.06 × 10 <sup>-4</sup> *** (2.77 × 10 <sup>-5</sup> )	-1.06 × 10 <sup>-4</sup> *** (2.77 × 10 <sup>-5</sup> )	-1.06 × 10 <sup>-4</sup> *** (2.77 × 10 <sup>-5</sup> )	-1.06 × 10 <sup>-4</sup> *** (2.77 × 10 <sup>-5</sup> )	-1.06 × 10 <sup>-4</sup> *** (2.77 × 10 <sup>-5</sup> )	-1.06 × 10 <sup>-4</sup> *** (2.77 × 10 <sup>-5</sup> )
Sample Mean Effect	-1.567	-1.531	-1.438	-2.042	-2.127	-2.089						
Fixed-effects												
Year	Yes											
Mine	Yes											
Fit statistics												
Observations	114,616	114,616	114,496	94,671	94,671	94,604						
R <sup>2</sup>	0.91804	0.92152	0.92390	0.77981	0.78184	0.78587						
Within R <sup>2</sup>	0.00346	0.04573	0.05524	0.00138	0.01060	0.02485						

Clustered (Mine) standard-errors in parentheses

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

[go back](#)

50/58

## Appendix Full Distance Specification Results

Dependent Variables:	Maximum EVI						Maximum Cropland EVI					
Model:	(1)	(2)	(3)	(4)	(5)	(6)						
Variables												
Downstream	-0.0035*** (0.0023)	-0.0050*** (0.0021)	-0.0052*** (0.0021)									
Downstream x Distance	-2.02 × 10 <sup>-5</sup> (0.0001)	1.05 × 10 <sup>-5</sup> (0.0001)	-2.02 × 10 <sup>-5</sup> (0.0001)	-2.02 × 10 <sup>-5</sup> (0.0001)	-2.02 × 10 <sup>-5</sup> (0.0001)	-2.02 × 10 <sup>-5</sup> (0.0001)	-2.02 × 10 <sup>-5</sup> (0.0001)	-2.02 × 10 <sup>-5</sup> (0.0001)	-2.02 × 10 <sup>-5</sup> (0.0001)	-2.02 × 10 <sup>-5</sup> (0.0001)	-2.02 × 10 <sup>-5</sup> (0.0001)	-2.02 × 10 <sup>-5</sup> (0.0001)
Downstream x Distance <sup>2</sup>	(8.37 × 10 <sup>-7</sup> )											
Distance	4.05 × 10 <sup>-3</sup> (9.03 × 10 <sup>-5</sup> )	2.98 × 10 <sup>-3</sup> (8.4 × 10 <sup>-5</sup> )	2.56 × 10 <sup>-3</sup> (8.4 × 10 <sup>-5</sup> )	-7.01 × 10 <sup>-3</sup> (8.19 × 10 <sup>-5</sup> )	6.93 × 10 <sup>-3</sup> (8.00 × 10 <sup>-5</sup> )	6.93 × 10 <sup>-3</sup> (8.00 × 10 <sup>-5</sup> )	6.93 × 10 <sup>-3</sup> (8.00 × 10 <sup>-5</sup> )	6.93 × 10 <sup>-3</sup> (8.00 × 10 <sup>-5</sup> )	6.93 × 10 <sup>-3</sup> (8.00 × 10 <sup>-5</sup> )	6.93 × 10 <sup>-3</sup> (8.00 × 10 <sup>-5</sup> )	6.93 × 10 <sup>-3</sup> (8.00 × 10 <sup>-5</sup> )	6.93 × 10 <sup>-3</sup> (8.00 × 10 <sup>-5</sup> )
Distance <sup>3</sup>	-1.87 × 10 <sup>-7</sup> (6.27 × 10 <sup>-7</sup> )	-9.18 × 10 <sup>-8</sup> (5.31 × 10 <sup>-8</sup> )	2.1 × 10 <sup>-7</sup> (5.31 × 10 <sup>-8</sup> )	6.93 × 10 <sup>-7</sup> (5.31 × 10 <sup>-8</sup> )	6.93 × 10 <sup>-7</sup> (5.31 × 10 <sup>-8</sup> )	6.93 × 10 <sup>-7</sup> (5.31 × 10 <sup>-8</sup> )	6.93 × 10 <sup>-7</sup> (5.31 × 10 <sup>-8</sup> )	6.93 × 10 <sup>-7</sup> (5.31 × 10 <sup>-8</sup> )	6.93 × 10 <sup>-7</sup> (5.31 × 10 <sup>-8</sup> )	6.93 × 10 <sup>-7</sup> (5.31 × 10 <sup>-8</sup> )	6.93 × 10 <sup>-7</sup> (5.31 × 10 <sup>-8</sup> )	6.93 × 10 <sup>-7</sup> (5.31 × 10 <sup>-8</sup> )
Elevation	-7.45 × 10 <sup>-6</sup> (6.56 × 10 <sup>-6</sup> )	-2.22 × 10 <sup>-5</sup> (6.71 × 10 <sup>-6</sup> )	(7.55 × 10 <sup>-6</sup> ) (6.71 × 10 <sup>-6</sup> )	(7.55 × 10 <sup>-6</sup> ) (6.71 × 10 <sup>-6</sup> )	(7.55 × 10 <sup>-6</sup> ) (6.71 × 10 <sup>-6</sup> )	(7.55 × 10 <sup>-6</sup> ) (6.71 × 10 <sup>-6</sup> )	(7.55 × 10 <sup>-6</sup> ) (6.71 × 10 <sup>-6</sup> )	(7.55 × 10 <sup>-6</sup> ) (6.71 × 10 <sup>-6</sup> )	(7.55 × 10 <sup>-6</sup> ) (6.71 × 10 <sup>-6</sup> )	(7.55 × 10 <sup>-6</sup> ) (6.71 × 10 <sup>-6</sup> )	(7.55 × 10 <sup>-6</sup> ) (6.71 × 10 <sup>-6</sup> )	(7.55 × 10 <sup>-6</sup> ) (6.71 × 10 <sup>-6</sup> )
Slope	0.0034*** (0.0005)	0.0033*** (0.0005)										
Yearly Max. Temperature	-0.0053*** (0.0007)											
Yearly Precipitation	3.33 × 10 <sup>-4</sup> *** (3.6 × 10 <sup>-5</sup> )	2.88 × 10 <sup>-4</sup> *** (3.6 × 10 <sup>-5</sup> )	2.88 × 10 <sup>-4</sup> *** (3.6 × 10 <sup>-5</sup> )	2.88 × 10 <sup>-4</sup> *** (3.6 × 10 <sup>-5</sup> )	2.88 × 10 <sup>-4</sup> *** (3.6 × 10 <sup>-5</sup> )	2.88 × 10 <sup>-4</sup> *** (3.6 × 10 <sup>-5</sup> )	2.88 × 10 <sup>-4</sup> *** (3.6 × 10 <sup>-5</sup> )	2.88 × 10 <sup>-4</sup> *** (3.6 × 10 <sup>-5</sup> )	2.88 × 10 <sup>-4</sup> *** (3.6 × 10 <sup>-5</sup> )	2.88 × 10 <sup>-4</sup> *** (3.6 × 10 <sup>-5</sup> )	2.88 × 10 <sup>-4</sup> *** (3.6 × 10 <sup>-5</sup> )	2.88 × 10 <sup>-4</sup> *** (3.6 × 10 <sup>-5</sup> )
Accessibility in 2015	-1.01 × 10 <sup>-4</sup> (5.31 × 10 <sup>-5</sup> )	-4.03 × 10 <sup>-5</sup> (1.19 × 10 <sup>-5</sup> )	-4.03 × 10 <sup>-5</sup> (1.19 × 10 <sup>-5</sup> )	-4								

## Appendix Varying Outcome / Fixed Effects

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

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

[go back](#)

53/58

## Appendix Placebo Outcomes

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

54/58

## Appendix Dist. Spec. w/ Aut. Bandwidth Selection (No Controls)

	Max EVI		Max C EVI	
	Conv.	Bias-Corr.	Conv.	Bias-Corr.
	No Controls			
Conventional	-0.0050*** (0.0015)	-0.0056*** (0.0015)	-0.0112*** (0.0020)	-0.0116*** (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 20, 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.

55/58

## Appendix Dist. Spec. w/ Aut. Bandwith Selection (Full Controls)

	Max EVI		Max C EVI	
	Conv.	Bias-Corr.	Conv.	Bias-Corr.
<b>With Full Controls</b>				
Conventional	-0.0045*** (0.0015)	-0.0049*** (0.0015)	-0.0100*** (0.0020)	-0.0118*** (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 20, 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.

56/58

## Appendix Ord. Spec. w/ Aut. Bandwith Selection (Full Controls)

	Max EVI		Max C EVI	
	No Cluster	Cluster (Mine Basin)	No Cluster	Cluster (Mine Basin)
<b>No Controls</b>				
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 20, 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

57/58

## Appendix Ord. Spec. w/ Aut. Bandwith Selection (Full Controls)

	Max EVI		Max C EVI	
	No Cluster	Cluster (Mine Basin)	No Cluster	Cluster (Mine Basin)
<b>With Full Controls</b>				
I(order>0)	-0.0048** (0.0012)	-0.0048 (0.0018)	-0.0090*** (0.0017)	-0.0090*** (0.0029)
Observations	45,580	45,580	38,504	38,504
Bandwidth	2	2	2	2

**Note:** Table shows results for estimation of 20, 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

go back

58/58

# Mines – Rivers – Yields

## Downstream Mining Impacts on Agriculture in Africa

Lukas Vashold ([lukas.vashold@wu.ac.at](mailto:lukas.vashold@wu.ac.at))

Gustav Pirich ([gustav.pirich@wu.ac.at](mailto:gustav.pirich@wu.ac.at))

**Max Heinze** ([maximilian.heinze@wu.ac.at](mailto:maximilian.heinze@wu.ac.at))

Nikolas Kuschnig ([nikolas.kuschnig@wu.ac.at](mailto:nikolas.kuschnig@wu.ac.at))

NOeG Winter Workshop

December 18, 2024

1/46

## Background

Methods and Data

Results

2/46

## Mines – Curse or Blessing?

### A Blessing?

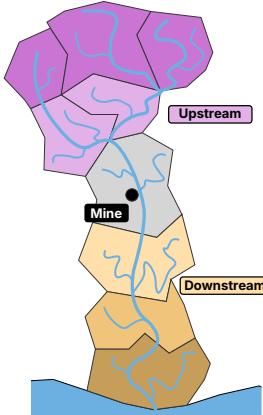
- Demand for relevant **minerals** is projected to increase **fourfold** until 2050 (Hund et al., 2023).
- **Extraction Benefits** include:
  - enabling the **green transition**,
  - increasing local **incomes** (Bazillier & Girard, 2020),
  - and improving **wealth** and **asset ownership** (von der 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** and **lead** (Schwarzenbach et al., 2010).
  - Industrial pollution **harms plant growth** (Yang et al., 2021).

3/46

## How to Find Affected Areas



Using data on **river basins** (Lehner & Grill, 2013), we know where water flows from a given location.

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

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

[show schematic depiction](#) [show more on basins](#)

4/46

## 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** (ICMM, 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).

5/46

Background

## Methods and Data

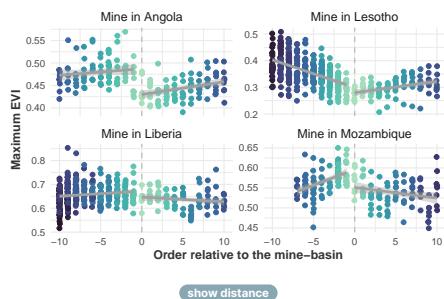
Results

6/46

## Intuition

The **four mines** depicted give an intuition for what we expect.

Following the river “flow” from left to right, we can see **discontinuities at the mine basin**.

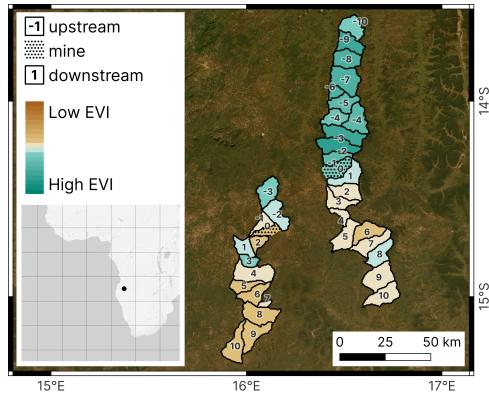


7/46

## Mines

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

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



8/46

## 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 the
  - **entire basin area**, and
  - **only on croplands** within the basin.

### Observations and Covariates

- We observe **6,698 upstream** basins, **1,900 mine** basins, and **5,729 downstream** basins over  $T = 8$  years. [\(Show order × status\)](#)
- We observe **covariates** on:
  - topography,
  - soil type,
  - climate, and
  - socioeconomic characteristics.

[\(Show summary statistics\)](#) [\(Show balance\)](#)

9/46

## Empirical Strategy (Spatial RDD), Identification

$$y_{ijt} = \beta_1 d_{ij} + \beta_2 d_{ij} \times \text{downstream}_j + \beta_3 \text{downstream}_j + \delta' \mathbf{x}_{it} + \mu_j + \psi_t + \varepsilon_{ijt},$$

- $y_{ijt}$ : **Outcome** of basin  $i$  near mine  $j$  in year  $t$ ,
- $\mu_j, \psi_t$ : Mine and year **fixed effects**,
- $\mathbf{x}_{it}$ : Basin specific **covariates**,
- $d_{ij}$ : **Distance** to nearest mine (as order or river stream length).
- Parameter  $\beta_3$  is identified under the assumption that there are no **other discontinuous changes** at the mine basin.
  - We check balance, include controls, conduct placebo checks.

10/46

### Background

### Methods and Data

## Results

11/46

## Results Overview

- We find a **significant reduction** in vegetation health **downstream** of mines.
- The magnitude of this effect is **greater on croplands**.
- Impacts **dissipate slowly** the farther we move from a mine.
- These results are **robust** to varying the sample, the outcome measurement, and the level of fixed effects.

12/46

## Order Specification Results (1)

	<i>Max. EVI</i> (Plain)	<i>Max. EVI</i> (Full)	<i>Max. Cropland EVI</i> (Plain)	<i>Max. Cropland EVI</i> (Full)
<b>Order</b>				
Mine-basin (0 <sup>th</sup> )	-0.0064*** (0.0014)	<b>-0.0059***</b> (0.0013)	-0.0093*** (0.0021)	<b>-0.0095***</b> (0.0020)
Downstream (1 <sup>st</sup> )	-0.0060*** (0.0018)	<b>-0.0057***</b> (0.0017)	-0.0049* (0.0026)	<b>-0.0061**</b> (0.0026)
Downstream (2 <sup>nd</sup> )	-0.0070*** (0.0021)	<b>-0.0066***</b> (0.0021)	-0.0042 (0.0028)	<b>-0.0062**</b> (0.0030)
Sample mean	0.412	0.412	0.454	0.454
Observations	114,616	114,496	94,671	94,604
R <sup>2</sup>	0.912	0.924	0.780	0.786
Controls	No	Yes	No	Yes
Year F.E.	Yes	Yes	Yes	Yes
Mine F.E.	Yes	Yes	Yes	Yes

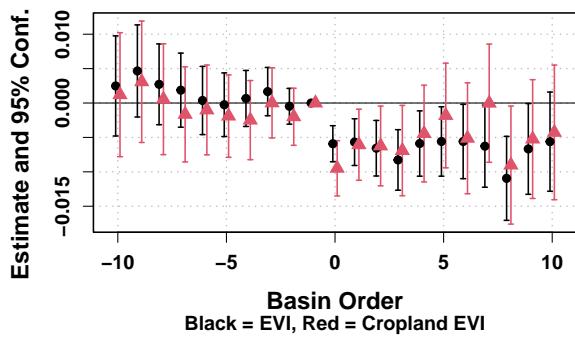
Clustered (by mine-basin) standard errors in parentheses.

Significance levels: \*\*\*: 0.01, \*\*: 0.05, \*: 0.1. [show full results](#) [show interpretation](#)

13/46

## Order Specification Results (2)

### Effects by Order of Basins (with Controls)



14/46

## Distance Specification Results

	<i>Max. EVI</i> (Plain)	<i>Max. EVI</i> (Full)	<i>Max. Cropland EVI</i> (Plain)	<i>Max. Cropland EVI</i> (Full)
<b>Distance</b>				
Downstream	<b>-0.0065***</b> (0.0023)	<b>-0.0058***</b> (0.0021)	<b>-0.0086***</b> (0.0029)	<b>-0.0087***</b> (0.0028)
Downstream × Distance	$-2.0 \times 10^{-5}$ (0.0001)	$-2.0 \times 10^{-5}$ (0.0001)	0.0003** (0.0001)	0.0002 (0.0001)
Downstream × Distance <sup>2</sup>	$-4.0 \times 10^{-7}$ (9.2 × 10 <sup>-7</sup> )	$-9.8 \times 10^{-8}$ (7.2 × 10 <sup>-7</sup> )	$-2.2 \times 10^{-6}**$ (1.1 × 10 <sup>-6</sup> )	$-1.9 \times 10^{-6}*$ (1.0 × 10 <sup>-6</sup> )
Sample mean	0.412	0.412	0.454	0.454
Observations	114,616	114,496	94,671	94,604
R <sup>2</sup>	0.918	0.924	0.780	0.786
Controls	No	Yes	No	Yes
Year F.E.	Yes	Yes	Yes	Yes
Mine F.E.	Yes	Yes	Yes	Yes

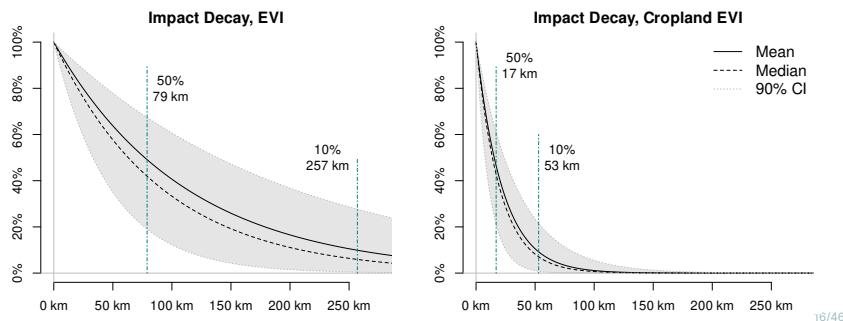
Clustered (by mine-basin) standard errors in parentheses.

Significance levels: \*\*\*: 0.01, \*\*: 0.05, \*: 0.1. [show full results](#)

15/46

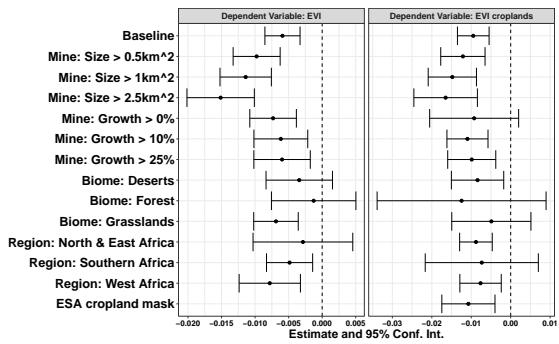
## Impact Decay

- 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. [details](#)



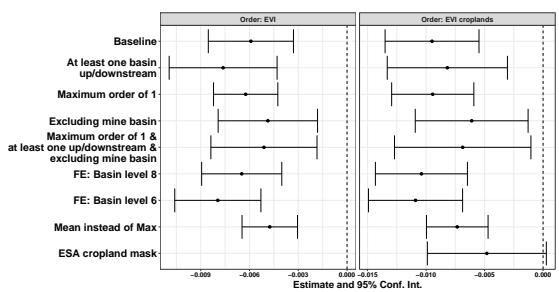
## Heterogeneity

- We investigate heterogeneity re.: **mine characteristics, biome, and region**.



## Robustness

- We check robustness by **varying the specification, estimation methods**, and checking **placebo** outcomes.



[show varying sample estimates](#) [show varying outcome/FE estimates](#) [show placebo estimates](#) [show automatic BW sel. estimates](#) 18/46

## Discussion

### Results

- We find **negative impacts** on vegetation health by about 1.4-2.1% at the sample mean.
- There is a need to
  - tackle the lack of **containment facilities** and improve environmental governance,
  - both for **industrial** and **informal** mines.

### Limitations

- **Remotely sensed measures** only represent crop yields **indirectly**.
- Our **treatment indicator** relied only on mine location.
- Differences in **waste management** are not accounted for.
- **Adaptive behavior** by farmers is not covered.

19/46

## 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, on grasslands, and in West Africa.

Results were **robust** to changes of treatment, outcome or sample definition.

20/46

## References (1)

- 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>

21/46

## References (2)

- 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](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*. Nováfrica.
- 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>

22/46

## References (3)

- ICMM. (2022). Mining Contribution Index (MCI) (6th Edition).  
[https://www.icmm.com/website/publications/pdfs/social-performance/2022/research\\_mci-6-ed.pdf?cb=16134](https://www.icmm.com/website/publications/pdfs/social-performance/2022/research_mci-6-ed.pdf?cb=16134)
- 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>
- 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>

23/46

## References (4)

- 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., Mudbhakal, 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., Siebert, 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>

24/46

## References (5)

- 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>
- 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>
- Yang, J., Sun, H., Qin, J., Wang, X., & Chen, W. (2021). Impacts of Cd on Temporal Dynamics of Nutrient Distribution Pattern of *Bletilla striata*, a Traditional Chinese Medicine Plant. *Agriculture*, 11(7), 594. <https://doi.org/10.3390/agriculture11070594>

25/46

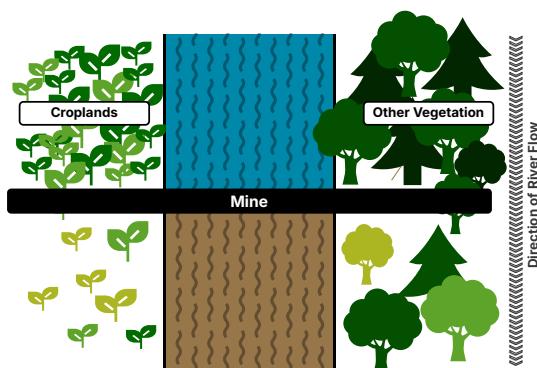
## Appendix

26/46

### Appendix How Pollution Travels

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

[go back](#)



27/46

## Appendix Basins



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**.

[go back](#)

Illustration from Lehner and Grill (2013)

28/46

## 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** (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).

29/46

## Appendix Summary Statistics

Variable	N	Mean	St. Dev.	Min.	Max.
<b>Max. EVI</b>	114,616	0.411	0.168	-0.112	0.993
<b>Mean EVI</b>	114,616	0.270	0.118	-0.112	0.578
<b>Max. Cropland EVI</b>	94,671	0.454	0.129	-0.112	0.990
<b>Mean Cropland EVI</b>	94,671	0.286	0.093	-0.114	0.734
<b>Max. Temperature</b>	114,616	33.80	4.047	20.00	45.40
<b>Precipitation</b>	114,616	882.3	606.3	0.555	4,375.3
<b>Population</b>	114,536	8,185	37,090	0.000	1,396,921
<b>Elevation</b>	114,616	804.6	482.0	-118.3	3,059.7
<b>Slope</b>	114,616	2.201	2.320	0.000	20.92
<b>Accessibility</b>	114,576	183.9	255.9	1.002	7,681

[go back to observations overview](#) [show balance](#)

30/46

## Appendix Order Specification Results

- We can see that **upstream** basins are unaffected, while **downstream** basins experience a significant negative effect.
- At the sample mean, the effect for the
  - **Max. EVI** corresponds to an EVI reduction of **1.4%**.
  - **Max. Cropland EVI** corresponds to an EVI reduction of **2.1%**.
- The effect **persists** beyond the mine basin.
- At higher order basins, impacts become imprecise.

[go back](#)

31/46

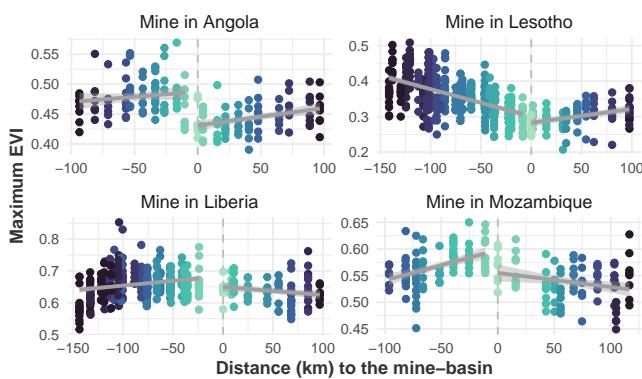
## Appendix Impact Decay Assessment

- We re-estimate our main specification using an **exponential decay** function  $\exp\{-\delta d_j\}$ .
- **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.

[go back](#)

32/46

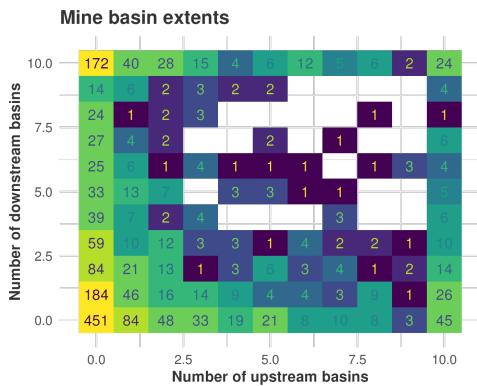
## Appendix Four Selected Mines, Distance



[go back](#)

33/46

## Appendix Basin Numbers



[go back to observations overview](#)

34/46

## Appendix Basins by Order

Order	Downstream N	Distance (km)	Upstream N	Distance (km)
0	1900	0.0	-	-
1	1162	10.7	987	14.5
2	841	22.2	865	24.2
3	695	32.9	778	34.7
4	591	43.7	738	44.7
5	531	54.4	681	55.1
6	462	64.8	593	65.9
7	418	74.3	575	75.6
8	376	85.1	530	86.6
9	343	95.9	499	95.7
10	310	106.1	452	104.2

[go back to observations overview](#)

35/46

## Appendix Summary Statistics for Upstream Basins

Variable	Upstream Basins				
	N	Mean	St. Dev.	Min.	Max.
Max. EVI	53,584	0.417	0.169	0.021	0.983
Mean EVI	53,584	0.276	0.120	0.020	0.578
Max. Cropland EVI	44,389	0.459	0.127	0.057	0.990
Mean Cropland EVI	44,389	0.291	0.093	-0.002	0.637
Max. Temperature	53,584	33.83	4.003	20.00	45.10
Precipitation	53,584	905.4	606.5	0.851	3,976.0
Population	53,584	6,693.8	27,878.2	0.000	1,396,921.0
Elevation	53,584	840.5	471.2	10.53	3,059.7
Slope	53,584	2.295	2.256	0.086	20.91
Accessibility	53,584	192.0	242.3	3.000	7,542.8

[go back to covariate overview](#) [go back to summary statistics](#)

36/46

## Appendix Summary Statistics for Downstream Basins

Downstream Basins (incl. Mine Basins)

Variable	N	Mean	St. Dev.	Min.	Max.
Max. EVI	61,032	0.406	0.167	-0.112	0.993
Mean EVI	61,032	0.264	0.116	-0.112	0.563
Max. Cropland EVI	50,282	0.450	0.130	-0.112	0.981
Mean Cropland EVI	50,282	0.283	0.093	-0.114	0.734
Max. Temperature	61,032	33.78	4.085	20.00	45.40
Precipitation	61,032	862.0	605.4	0.555	4,375.3
Population	60,952	9,497.1	43,568.1	0.000	1,244,492.0
Elevation	61,032	773.1	489.1	-118.3	3,047.1
Slope	61,032	2.119	2.371	0.000	20.456
Accessibility	60,992	176.9	267.1	1.002	7,681.8

[go back to covariate overview](#) [go back to summary statistics](#)

37/46

## Appendix Full Order Specification Results

Dependent Variables:	(1)	(2)	(3)	(4)	(5)	(6)
Model:						
Variables						
Downstream x Order = 0	-0.0064*** (0.0014)	-0.0053*** (0.0014)	-0.0059*** (0.0013)	-0.0073*** (0.0021)	-0.0097*** (0.0021)	-0.0095*** (0.0020)
Downstream x Order = 1						
Downstream x Order = 2						
Downstream x Order = 3						
Downstream x Order = 4						
Downstream x Order = 5						
Downstream x Order = 6						
Downstream x Order = 7						
Downstream x Order = 8						
Downstream x Order = 9						
Downstream x Order = 10						
Elevation						
Slope						
Yearly Max. Temperature						
Yearly Precipitation						
Accessibility in 2015						
Population in 2015						
Sample Mean Effect	-1.567	-1.531	-1.438	-2.042	-2.127	-2.089
Fixed-effects						
Year	Yes	Yes	Yes	Yes	Yes	Yes
Mine	Yes	Yes	Yes	Yes	Yes	Yes
Fit statistics						
Observations	114,616	114,616	114,496	94,671	94,671	94,604
R <sup>2</sup>	0.91804	0.92152	0.92390	0.77971	0.78175	0.76587
Within R <sup>2</sup>	0.00393	0.04627	0.05582	0.00180	0.01099	0.02531

Clustered (Mine) standard-errors in parentheses

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

[go back](#)

38/46

## Appendix Full Distance Specification Results

Dependent Variables:	(1)	(2)	(3)	(4)	(5)	(6)
Model:						
Variables						
Downstream	-0.0065*** (0.0023)	-0.0060*** (0.0021)	-0.0058*** (0.0021)	-0.0086*** (0.0029)	-0.0088*** (0.0029)	-0.0087*** (0.0028)
Downstream x Distance	-2.02 × 10 <sup>-5</sup> (0.0001)	1.05 × 10 <sup>-5</sup> (0.0001)	-2.02 × 10 <sup>-5</sup> (0.0001)	0.0003** (0.0001)	0.0002* (0.0001)	0.0002* (0.0001)
Downstream x Distance <sup>2</sup>	-3.98 × 10 <sup>-7</sup> (9.17 × 10 <sup>-7</sup> )	-4.37 × 10 <sup>-7</sup> (7.35 × 10 <sup>-7</sup> )	-9.8 × 10 <sup>-8</sup> (7.19 × 10 <sup>-7</sup> )	-2.15 × 10 <sup>-6*</sup> (1.06 × 10 <sup>-6</sup> )	-2.34 × 10 <sup>-6*</sup> (1.03 × 10 <sup>-6</sup> )	-1.94 × 10 <sup>-6</sup> (1.03 × 10 <sup>-6</sup> )
Distance	4.05 × 10 <sup>-3</sup> (9.03 × 10 <sup>-4</sup> )	2.98 × 10 <sup>-3</sup> (8.4 × 10 <sup>-4</sup> )	2.56 × 10 <sup>-3</sup> (8.19 × 10 <sup>-4</sup> )	-7.01 × 10 <sup>-5</sup> (2.00 × 10 <sup>-5</sup> )	-5.62 × 10 <sup>-5</sup> (1.98 × 10 <sup>-5</sup> )	-4.6 × 10 <sup>-5</sup> (1.96 × 10 <sup>-5</sup> )
Distance <sup>2</sup>	-1.92 × 10 <sup>-7</sup> (6.27 × 10 <sup>-7</sup> )	5.68 × 10 <sup>-7</sup> (5.68 × 10 <sup>-7</sup> )	5.56 × 10 <sup>-7</sup> (5.56 × 10 <sup>-7</sup> )	8.93 × 10 <sup>-7</sup> (8.38 × 10 <sup>-7</sup> )	8 × 10 <sup>-7</sup> (8.22 × 10 <sup>-7</sup> )	6.98 × 10 <sup>-7</sup> (8.22 × 10 <sup>-7</sup> )
Elevation	-7.45 × 10 <sup>-6</sup> (6.58 × 10 <sup>-6</sup> )	-2.22 × 10 <sup>-5</sup> *** (0.00324)	-2.22 × 10 <sup>-5</sup> *** (0.00324)	-1.83 × 10 <sup>-5</sup> *** (0.00232)	-1.83 × 10 <sup>-5</sup> *** (0.00232)	-4.03 × 10 <sup>-5</sup> *** (0.0006)
Slope						
Yearly Max. Temperature						
Yearly Precipitation						
Accessibility in 2015						
Population in 2015						
Sample Mean Effect	-0.0055*** (0.0007)	-0.0055*** (0.0007)	-0.0055*** (0.0007)	-0.0070*** (0.0007)	-0.0070*** (0.0007)	-0.0070*** (0.0007)
Fixed-effects						
Year	Yes	Yes	Yes	Yes	Yes	Yes
Mine	Yes	Yes	Yes	Yes	Yes	Yes
Fit statistics						
Observations	114,616	114,616	114,496	94,671	94,671	94,604
R <sup>2</sup>	0.91804	0.92152	0.92390	0.77971	0.78175	0.76587
Within R <sup>2</sup>	0.00346	0.04573	0.05524	0.00138	0.01060	0.02485

Clustered (Mine) standard-errors in parentheses

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

[go back](#)

39/46

## Appendix Varying Sample Definition

Dependent Variables:	(1)	(2)	Maximum EVI	(3)	(4)	(5)	(6)	Maximum Cropland EVI	(7)	(8)	(9)	(10)
<b>Model:</b>												
Variables												
Downstream x Order = 0	-0.0059*** (0.0013)	-0.0076*** (0.0014)	-0.0062*** (0.0012)			-0.0095*** (0.0020)	-0.0082*** (0.0024)	-0.0094*** (0.0022)				
Downstream x Order = 1	-0.0057*** (0.0017)	-0.0053*** (0.0020)	-0.0053*** (0.0017)	-0.0049** (0.0020)	-0.0051** (0.0021)	-0.0061** (0.0026)	-0.0049 (0.0032)	-0.0051* (0.0030)	-0.0061** (0.0030)	-0.0069* (0.0039)		
Downstream x Order = 2	-0.0066*** (0.0021)	-0.0054** (0.0026)		-0.0056** (0.0023)		-0.0062** (0.0030)	-0.0057 (0.0037)		-0.0062* (0.0033)			
<b>Fixed-effects</b>												
Year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mine	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
<b>Fit statistics</b>												
Observations	114,496	61,712	32,360	99,320	9,168	94,604	50,914	27,589	81,278	7,623		
R <sup>2</sup>	0.92395	0.91954	0.90419	0.92392	0.93378	0.78597	0.76613	0.84032	0.78332	0.81766		
Within R <sup>2</sup>	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](#)

40/46

## Appendix Varying Outcome / Fixed Effects

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

[go back](#)

41/46

## Appendix Placebo Outcomes

Dependent Variables:	Elevation	Slope	Max. Temp	Precipitation	Accessibility in 2015	Population in 2015	(1)	(2)	(3)	(4)	(5)	(6)	
<b>Model:</b>													
Variables													
Downstream	-6.852 (8.509)	-0.0538 (0.0912)	-0.0137 (0.0567)	0.6025 (3.934)	-5.427 (5.531)	2,125.7 (1,589.8)							
Distance x Downstream	-5.008*** (0.4814)	-0.0080 (0.0044)	0.0135*** (0.0036)	-0.1942 (0.2860)	-0.00839 (0.3278)	-182.9*** (55.80)							
Distance <sup>2</sup> x Downstream	0.0043 (0.0039)	-8.25 × 10 <sup>-6</sup> (4.01 × 10 <sup>-5</sup> )	2.12 × 10 <sup>-6</sup> (3.36 × 10 <sup>-5</sup> )	0.0003 (0.0020)	0.0004 (0.0028)	1.081*** (0.3463)							
Distance	2.326*** (0.4215)	0.0025 (0.0039)	-0.0067** (0.0032)	0.0879 (0.2129)	0.7557*** (0.2587)	-54.72 (45.17)							
Distance <sup>2</sup>	0.0005 (0.0033)	1.12 × 10 <sup>-6</sup> (3.49 × 10 <sup>-5</sup> )	-5.34 × 10 <sup>-6</sup> (3.1 × 10 <sup>-5</sup> )	-0.0005 (0.0015)	-0.0013 (0.0021)	0.3439 (0.2724)							
<b>Fixed-effects</b>													
Year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mine	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
<b>Fit statistics</b>													
Observations	114,616	114,616	114,616	114,616	114,576	114,536							
R <sup>2</sup>	0.95627	0.70192	0.95579	0.96187	0.88768	0.59121							
Within R <sup>2</sup>	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](#)

42/46

## Appendix Dist. Spec. w/ Aut. Bandwidth Selection (No Controls)

	Max EVI		Max C EVI	
	Conv.	Bias-Corr.	Conv.	Bias-Corr.
<b>No Controls</b>				
Conventional	-0.0050*** (0.0015)	-0.0056*** (0.0015)	-0.0112*** (0.0020)	-0.0116*** (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 10, 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.

43/46

## Appendix Dist. Spec. w/ Aut. Bandwidth Selection (Full Controls)

	Max EVI		Max C EVI	
	Conv.	Bias-Corr.	Conv.	Bias-Corr.
<b>With Full Controls</b>				
Conventional	-0.0045*** (0.0015)	-0.0049*** (0.0015)	-0.0100*** (0.0020)	-0.0118*** (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 10, 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.

44/46

## Appendix Ord. Spec. w/ Aut. Bandwidth Selection (Full Controls)

	Max EVI		Max C EVI	
	No Cluster	Cluster (Mine Basin)	No Cluster	Cluster (Mine Basin)
	<b>No Controls</b>			
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 10, 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

45/46

## Appendix Ord. Spec. w/ Aut. Bandwith Selection (Full Controls)

Max EVI		Max C EVI		
No Cluster	Cluster (Mine Basin)	No Cluster	Cluster (Mine Basin)	
<b>With Full Controls</b>				
I(order>0)	-0.0048** (0.0012)	-0.0048 (0.0018)	-0.0090*** (0.0017)	-0.0090*** (0.0029)
Observations	45,580	45,580	38,504	38,504
Bandwidth	2	2	2	2

**Note:** Table shows results for estimation of 10, 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

[go back](#)

46/46