nature sustainability

Supplementary information https://doi.org/10.1038/s41893-024-01421-8

Deforestation triggered by artisanal mining in eastern Democratic Republic of the Congo

In the format provided by the authors and unedited

Deforestation trigger from artisanal mining in eastern DR Congo

Supplementary Information

A Altnerative control regimes

The DiD estimates reported in the article use not-yet-established mines as control units for calculating the group-time average treatment effects for already established mines. This strategy allows us to not rely on randomly chosen control units that may differ from mining locations in non-random ways, but also imposes a slightly stronger parallel trends assumption than estimation with never-treated control units[1]. Following Callaway & Sant'Anna (2021)[2], the parallel trends assumption for group g in time period $t > g$ under a never-treated control regime without treatment anticipation is formalised as

$$E[Y_t(0) - Y_{t-1}(0)|G_g = 1] = E[Y_t(0) - Y_{t-1}(0)|X, C = 1] \tag{1}$$

where $Y_t(0)$ is the outcome in year t, G_g indicates whether a unit received first treatment in period g, X is a vector that allows to include covariates if parallel trends only hold conditional on covariates and C indicates whether a unit is never-treated. It is important to note that the left side of the equation in assumption 1 involves a counterfactual that is not observable and typically inferred from the analysis of pre-trends.

In comparison, the parallel trends assumption with not-yet-treated control units is formalised as

$$E[Y_t(0) - Y_{t-1}(0)|G_g = 1] = E[Y_t(0) - Y_{t-1}(0)|X, D_s = 0] \tag{2}$$

with D_s being a treatment indicator and $s \geq t$. Comparing the two assumptions, it can be seen that the former restricts parallel trends for each group g to a comparison between g and the group of never-treated control units that remains constant over time, while the latter assumption imposes parallel trends between group g and the group of not-yet-treated units that changes over time.

The question which control regime to use then ultimately depends on whether one believes parallel trends are more reasonable to hold between mining sites and randomly chosen control units, conditional on covariates (assumption 1), or whether the assumption is more reasonable to hold for mining sites established in different years, as they serve as control units for each other (assumption 2). Given that assumption 1 in our case relies crucially on the inclusion of all relevant covariates that could lead to a violation of parallel trends to account for the heterogeneity between randomly chosen and mining locations, we opted for not-yet-established mines as the more comparable control units in the main specification[3].

The DiD design can be enhanced with the inclusion of covariates with the rational to correct for violations of the parallel trends assumption[2]. Although we interpreted the pseudo-effect estimates for the unconditional specifications with not-yet-treated control units in the article as evidence in favor of parallel trends, testing how results change after conditioning on covariates can provide a useful robustness test[4]. Therefore, we conditioned the pretrends on a number of covariates (Table 1) and reran the analysis. As reported in Figure 1, results were smaller than without covariates, now amounting to 2.7 pp additional forest lost compared to the 4 pp from before. However, the effect remained significant and clearly visible.

Table 1: List of covariates used as control variables under conditional parallel trends.

Variable	Description	Source
Tree cover in 2000	Share of pixels with >60% tree cover in spatial extent	Hansen et al. (2013)[5]
Accessibility	Travel duration via surface transport to nearest town of at least 50,000 inhabitants or contiguous area with 1,500 or more inhabitants per square kilometre	Weiss et al. (2018)[6]
Distance to roads	Distance to the closest road, classified as either *primary* or *secondary*	OpenStreetMap
Distance to rivers	Distance to the closest river	OpenStreetMap
Altitude	Altitude of the location	FAO's GAEZ[7]
Slope	Slope of the location	FAO's GAEZ[7]
Temperature	Average temperature	FAO's GAEZ[7]
Precipitation	Annual precipitation	FAO's GAEZ[7]
Distance to logging roads	Distance to logging roads established before 2003	Kleinschroth et al. (2019)[8]
Distance to border	Distance to the nearest national border	DRC Forest Atlas (WRI)
Distance to Protected Areas	Distance to the nearest protected area border, with negative values for locations inside.	DRC Forest Atlas (WRI)

In a second variation from our basline specification, we picked random locations with more than 30% pixels of > 60% tree cover within 5km buffers (Figure 4b). We then estimated the DiD for the "never-established" mines while conditioning on the covariates in Table 1. For this specification, we found an additional loss of 3 pp after 10 years, again slightly lower than the estimate from the main specification but larger than the covariate-matched approach with not-yet-established mines as control units. However, we note that the two of the pseudo coefficients are statistically different from 0 in this case.

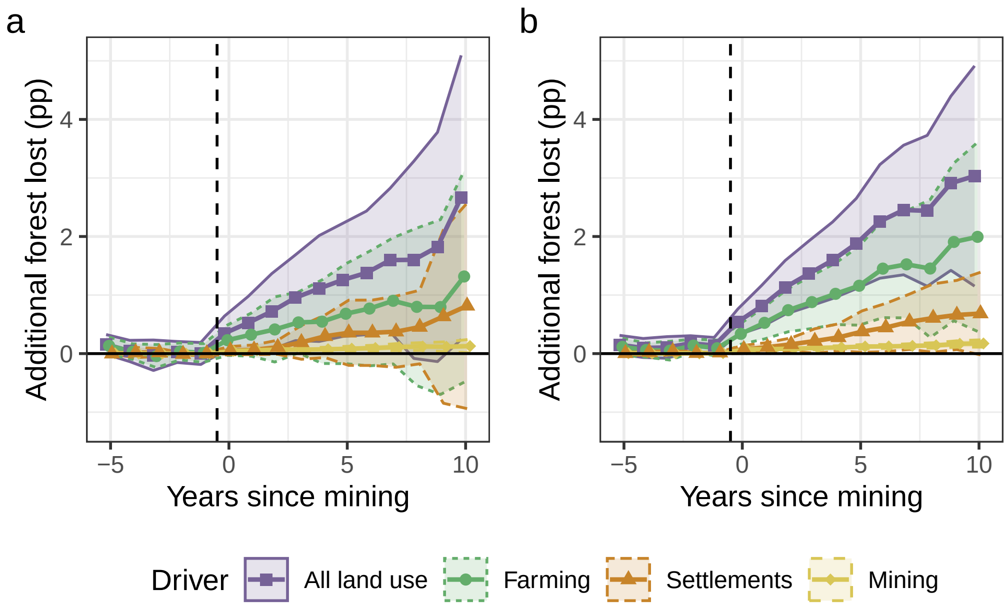

Figure 1: Estimates of impact on cumulative deforestation since 2000 as percentage points of additional forest cover lost, within a 5-km radius around mines, over time since treatment for various land uses following deforestation. All estimates conditional on the covariates tree cover in 2000, accessibility, distance to roads, distance to rivers, altitude, location, growing period length, distance to logging roads and distance to national borders. (a) Not-yet-established mines as controls (n=255) and (b) Never-treated randomly chosen control units (n=1255). Bootstrapped standard errors are clustered at the mine level. Points indicate ATT estimates following the DiD estimator proposed in Callaway & Sant'Anna (2021) and errorbars refer to 95% confidence bands.

B Alternative data

The data used in our main analysis is based on Masolele et al (2024)[9] . Although it relies on Global Forest Change (GFC) data[5] for the detection of deforestation events, it markedly differs from GFC. GFC identifies tree cover loss since 2001, which can be a consequence of deforestation or degradation, but also canopy cover loss on non-forest land[10]. In contrast, the data from our main analysis focuses on post-forest land use and therefore on persistent land-cover change, which excludes many forest loss events in GFC.

To illustrate this difference, we reran the analysis with GFC data and compared it with results estimated with the Tropical Moist Forest (TMF) data set developed by[11]. TMF has the same resolution as GFC but does not work with a canopy threshold to define forest loss. Instead, TMF tracks forest changes since 1986 and classifies them as undisturbed, degraded or deforested forest pixels, depending on the severity and duration of canopy cover disturbance. To be classified as deforested, disturbance within a forest pixel must exceed 2.5 years.

As the results in Figure 1 show, the deforestation estimates using TMF are similar to those from our main analysis, while TMF degradation and GFC forest loss are substantially higher.

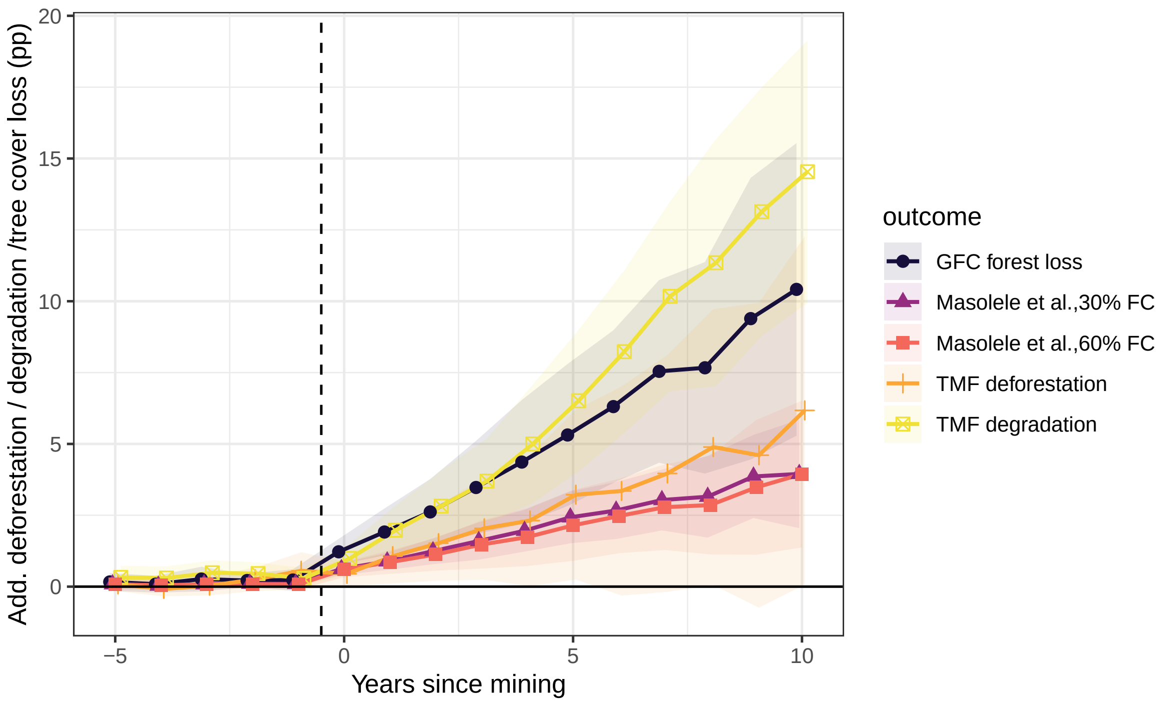

Figure 2: Additional disturbance effects by datasets. Points indicate ATT estimates following the DiD estimator proposed in Callaway & Sant'Anna (2021) and errorbars refer to 95% confidence bands. n=255 mining sites.

C Alternative estimator

In addition to Callaway and Sant'Anna's (2021)[2] (CSA) group-time specific estimator, we also used the two-stage procedure proposed by Gardner (2022)[12] (GAR) (see also Borusyak et al. (2024)[13]for a similar approach). In contrast to CSA, it follows a two-stage procedure. First, counterfactual trends that would occur in the absence of treatment are imputed by using the not-yet-treated units only:

$$Y_{it} = \mu_i + \lambda_t + u_{it} \tag{3}$$

Once Y is imputed, the ATTs can be estimated using a second stage regression of treatment dummies on the imputed outcome.

Lead coefficients for the years before treatment can be included into the second-stage regression to test the plausibility of the parallel trends assumption. As Roth et al. (2023)[14] note, the GAR estimator relies on a more stringent parallel pre-trends assumption than that of CAS to obtain unbiased estimates, since it requires parallel pre-trends to hold for all included pre-treatment periods to impute the counterfactual. In contrast, CSA's group-time weighted average estimator only relies on parallel trends in the period before treatment begins for a valid counterfactual, although it seems unlikely that parallel trends would have been observed in the absence of treatment in such a case. It can therefore be useful to limit the number of lead coefficients that are used to impute the first stage. In the absence of an established procedure by which to determine the number of leads, five pre-treatment periods were used throughout the analysis.

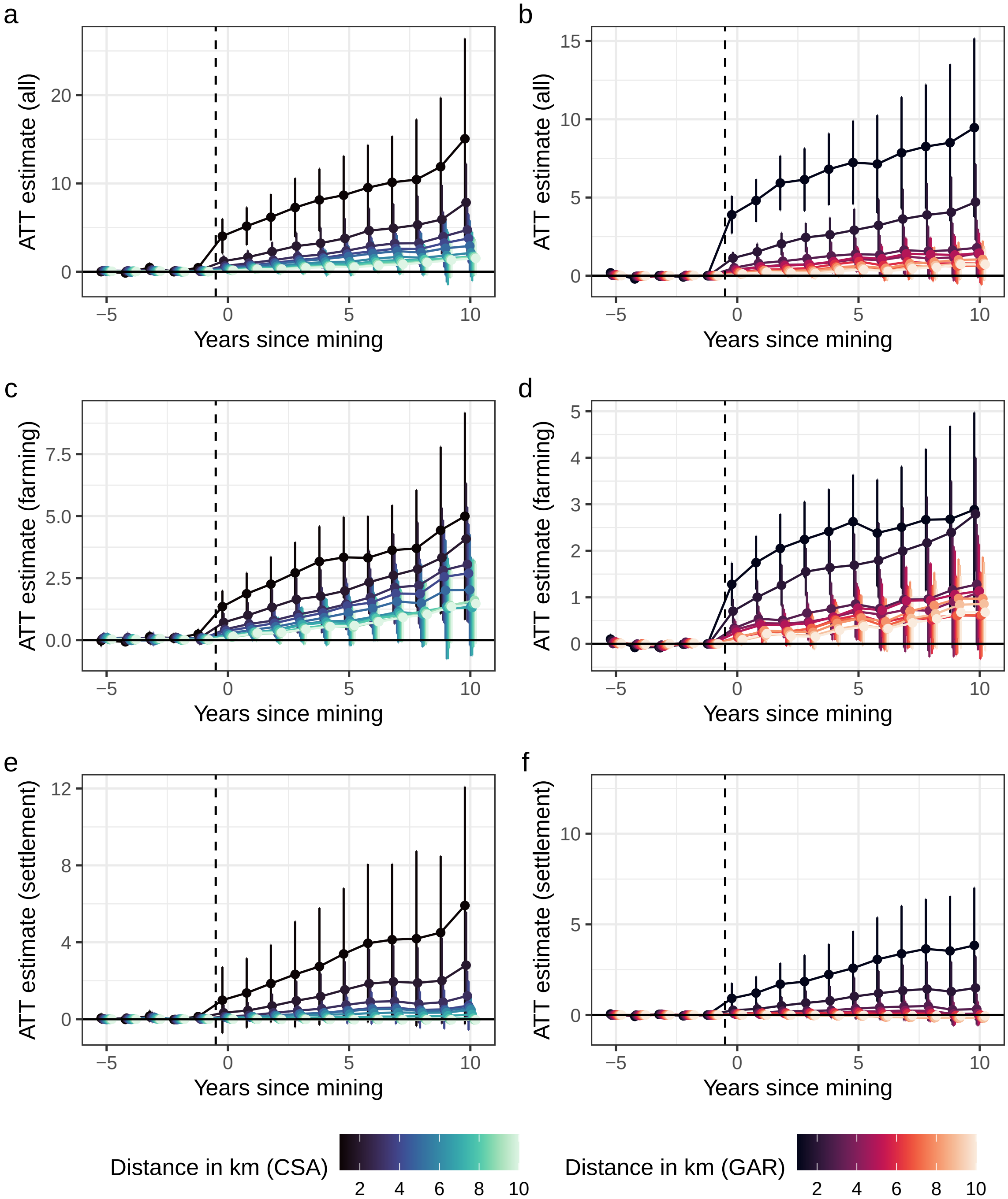

Figure 3: Event-study plots showing the ATT estimated as additional percentage points of forest lost within rings of increasing distance from the mining sites, over time. Plots (a), (c) and (e) are estimated following Callaway & Sant'Anna (2021), while (b), (d) and (e) employed the Gardner (2022) two stage procedure. The colours refer to the increasing distance from the mines. Confidence intervals are displayed at 95% level. n = 255 mining sites.

Figure 2 compares the CSA and GAR estimates in relation to time since the first mining incidence and distance from the mine. It shows that estimates run using GAR are more conservative across all specifications, although they still clearly reveal the effects of mining on surrounding forests.

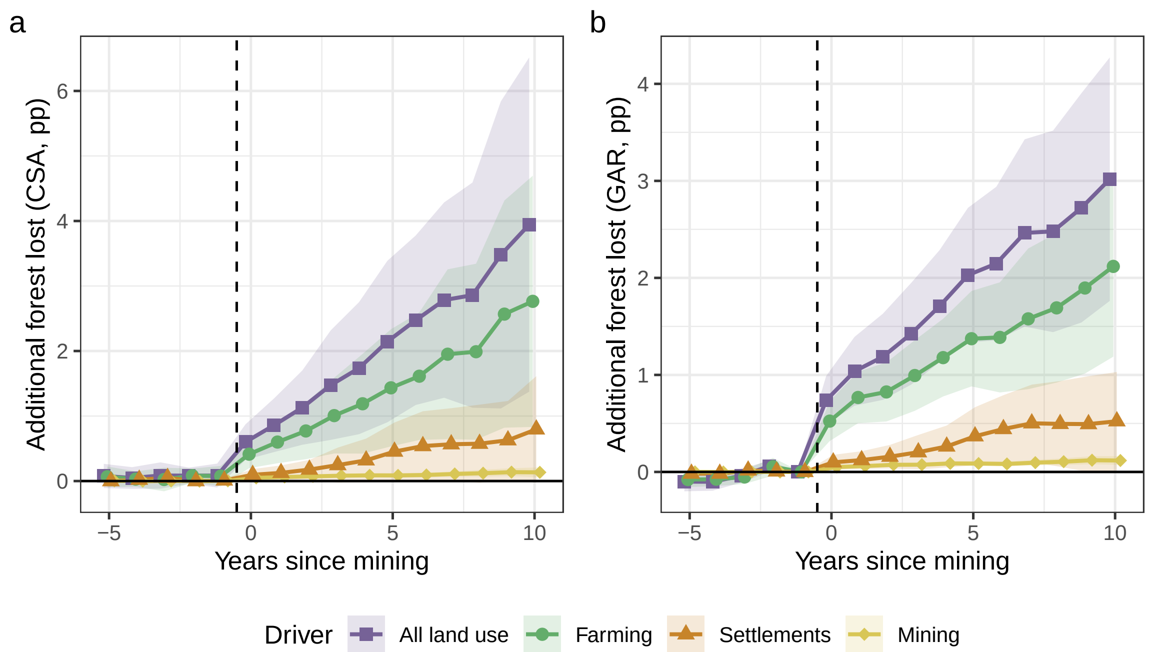

Figure 4: Estimates of impact on cumulative deforestation since 2000 as a share of forest cover in 2000 within a 5-km radius around mines for different periods following treatment, showing various land use following deforestation. For every additional kilometre from the mine, a new ring is estimated using (a) the estimator used by Callaway & Sant'Anna (2021), and (b) the estimator proposed in Gardner (2022). Displayed confidence intervals at 95%.n = 255 mining sites.

Figure 3 compares the estimates for the CSA estimator, shown in panel A, and the GAR estimator, panel B. As with the spatio-temporal estimation, results from GAR are more conservative than from CSA, but still convey a similar message: mining leads to substantial forest loss within a 5-km buffer around sites, and farming activities play an important role in explaining the forest loss.

	500m		1000m		1500m		3000m	
period	ATT	SE	ATT	SE	ATT	SE	ATT	SE
-5	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
-4	0.0	0.1	0.0	0.1	0.0	0.1	0.0	0.1
-3	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
-2	0.1	0.0	0.1	0.0	0.1	0.0	0.1	0.0
-1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
0	0.6*	0.1	0.6*	0.1	0.6*	0.1	0.6*	0.1
1	0.9*	0.1	0.9*	0.1	0.9*	0.2	0.9*	0.1
2	1.1*	0.2	1.1*	0.2	1.1*	0.2	1.1*	0.2
3	1.5*	0.3	1.5*	0.3	1.5*	0.3	1.5*	0.3
4	1.7*	0.4	1.7*	0.4	1.7*	0.4	1.7*	0.4
5	2.1*	0.4	2.1*	0.4	2.1*	0.4	2.1*	0.4
6	2.5*	0.5	2.5*	0.5	2.5*	0.5	2.5*	0.5
7	2.8*	0.6	2.8*	0.6	2.8*	0.5	2.8*	0.5
8	2.8*	0.6	2.9*	0.6	2.9*	0.7	2.8*	0.6
9	3.5*	0.9	3.5*	0.9	3.5*	0.9	3.5*	0.9
10	3.9*	0.9	3.9*	0.9	3.9*	0.9	3.9*	0.9

Notes. Impact estimates within 5-km buffers around mines using Callaway & Sant'Anna (2021) estimator for different cluster thresholds. ATT estimates without 95% confidence intervals spanning zero are indicated with *. n = 255 mining sites.

Table 2: Cluster threshold variations

D Cluster threshold variations

E Spillover-robust estimation

A challenge in the empirical strategy is the isolation of one mine's forest loss effect from another proximate mine's spillover effects. Such contamination could bias pre-trend estimates or lead to problems of "double counting", where the forest loss effect of one mine inflates the estimated effect of another. Although important to consider in many settings, DiD designs with spatial spillovers are still poorly understood[14]. We account for potential spillovers in a ring-estimator framework as outlined in[15], where control units can be exposed to treatment effects from nearby treated units. The spillover estimation is similar to the two-stage estimator of[12], the difference being that the first stage is imputed from observations that are not-yet-treated and, additionally, not-yet-spillover-exposed, i.e. they do not have any active mines located nearby. By including spillover-exposure dummies into the second-stage regression, the treatment effect can be spllit into the spillover deforestation from other mines and the actual deforestation from the mine itself, given that the parallel-trends assumption extends to the spillover area.

Figure 4 shows that the spillover-robust ATT estimates do not differ systematically from the ATTs estimated using the two-stage procedure outlined in Gardner (2022). The spillover effects appear insignificant and relatively constant over time. Therefore, we conclude that spillovers do not impact the validity of our findings.

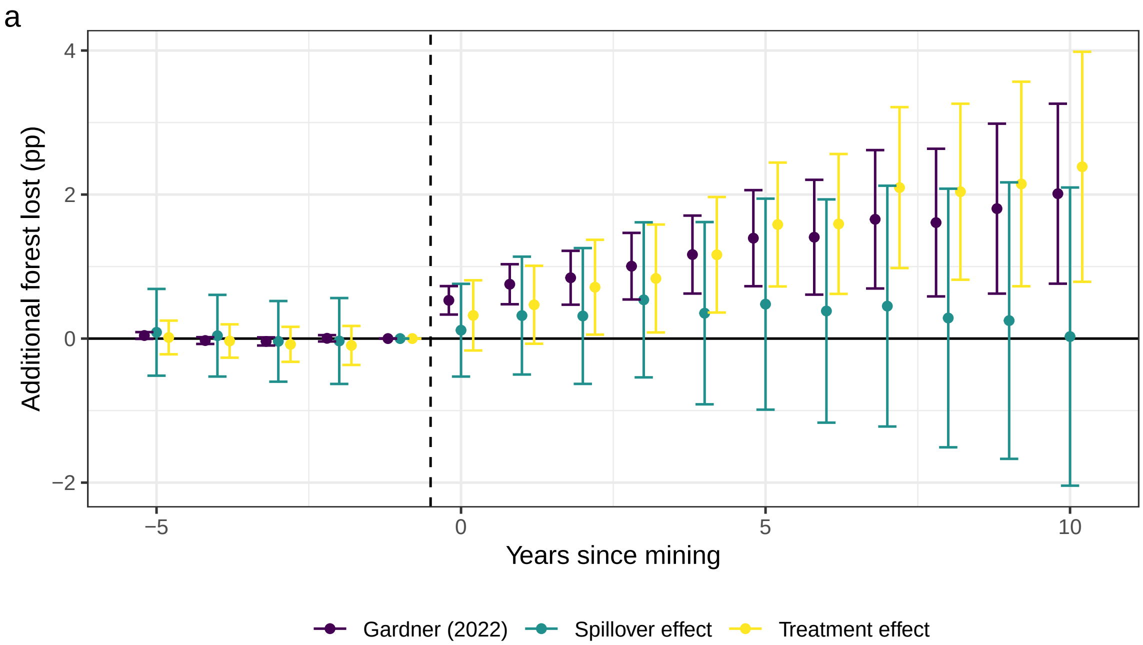

Figure 5: Spillover-robust estimation. Points represent effect estimates with ' Gardner (2022)' referring to ATTs estimated with the two-stage estimator proposed in Gardner (2022), while 'Spillover effect' and 'Treatment effect' refer to estimated parameters as suggested in Butts (2023). Bootstraped 95% confidence intervals are displayed.

	Accessibility	Agr. Suitability	River distance	Primary road dist.	Main road dist.	Forest cover in 2000	Conflict events
Accessibility	1	.	.	.	.	.	.
Agr. Suitability	-.08	1	.	.	.	.	.
River distance	-.05	.02	1	.	.	.	.
Primary road dist.	.39	-.15	-.21	1	.	.	.
Main road dist.	.57	-.18	-.17	.36	1	.	.
Forest cover in 2000	.32	-.15	.11	.10	.14	1	.
Conflict events	-.02	-.05	-.01	-.09	-.02	-.16	1

Table 3: Pearson correlation coefficients for the different variables tested in the heterogeneity analysis.

F Descriptive statistics

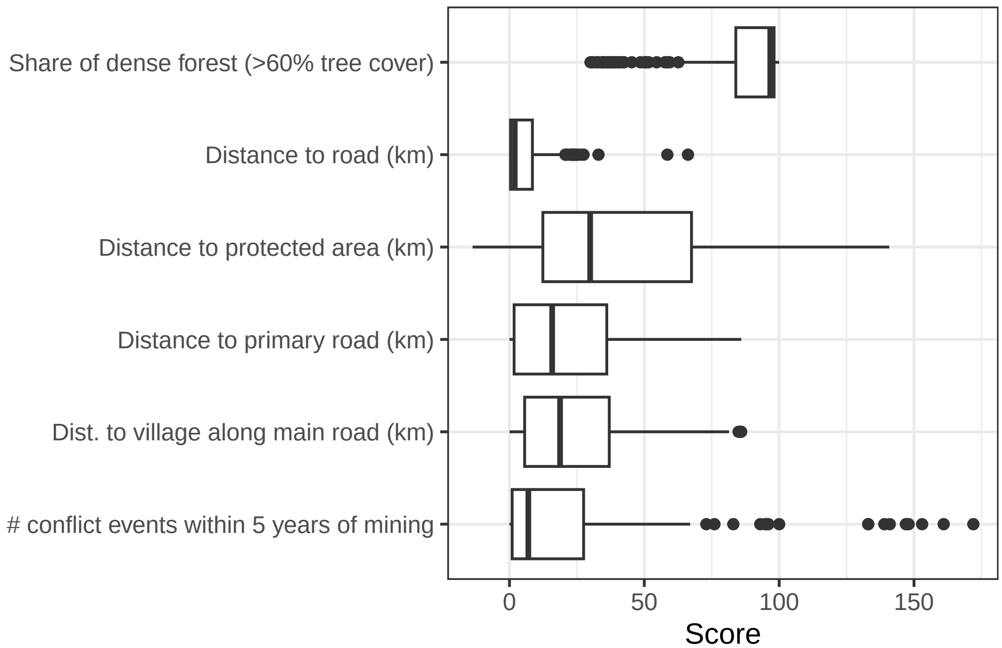

Figure 6: Boxplot diagrams for identified mining sites (n=255) drawn for different covariates. The unit of the score on the x-axis changes with covariates as in indicated in the variable discription on the y-axis.

G Software

The analysis was conducted in R version 4.3.2. For the data processing, we used the packages *terra*[16] and *sf*[17]. The statistical analysis relied on the packages *did*[18]and *did2s*[19].

References

1. Marcus, M. & Sant'anna, P. H. C. The role of parallel trends in event study settings: An application to environmental economics. *Journal of the Association of Environmental and Resource Economists* **8**, 235–275 (2021).
2. Callaway, B. & Sant'Anna, P. H. C. Difference-in-differences with multiple time periods. *Journal of Econometrics* **225**, 200–230 (2021).
3. de Chaisemartin, C. & D'Haultfœuille, X. Two-way fixed effects and differences-in-differences with heterogeneous treatment effects: A survey. *The Econometrics Journal* utac017 (2022) doi:10.1093/ectj/utac017.
4. Roth, J. Pretest with caution: Event-study estimates after testing for parallel trends. *American Economic Review: Insights* **4**, 305–322 (2022).
5. Hansen, M. C. *et al.* High-resolution global maps of 21st-century forest cover change. *Science* **342**, 850–854 (2013).
6. Weiss, D. J. *et al.* A global map of travel time to cities to assess inequalities in accessibility in 2015. *Nature* **553**, 333–336 (2018).
7. Fischer, G. *et al.* *Global Agro Ecological Zones V4 - Model Documentation.* (FAO, Rome, Italy, 2021).
8. Kleinschroth, F., Laporte, N., Laurance, W. F., Goetz, S. J. & Ghazoul, J. Road expansion and persistence in forests of the congo basin. *Nature Sustainability* **2**, 628–634 (2019).
9. Masolele, R. N. *et al.* Mapping the diversity of land uses following deforestation across Africa. *Scientific Reports* **14**, 1681 (2024).
10. Pendrill, F. *et al.* Disentangling the numbers behind agriculture-driven tropical deforestation. *Science* **377**, (2022).
11. Vancutsem, C. *et al.* Long-term (1990-2019) monitoring of forest cover changes in the humid tropics. *Science Advances* **7**, 1–22 (2021).
12. Gardner, J. *Two-Stage Differences in Differences.* http://arxiv.org/abs/2207.05943 (2022).
13. Borusyak, K., Jaravel, X. & Spiess, J. Revisiting event-study designs: Robust and efficient estimation. *The Review of Economic Studies* rdae007 (2024) doi:10.1093/restud/rdae007.
14. Roth, J., Sant'Anna, P. H. C., Bilinski, A. & Poe, J. What's trending in difference-in-differences? A synthesis of the recent econometrics literature. *Journal of Econometrics* **235**, 2218–2244 (2023).
15. Butts, K. *Difference-in-Differences Estimation with Spatial Spillovers.* http://arxiv.org/abs/2105.03737 (2023).
16. Hijmans, R. J., Bivand, R., Pebesma, E. & Sumner, M. D. *Terra Package - RDocumentation.* (2023).
17. Pebesma, E. *et al.* *Sf: Simple Features for r.* (2023).

18. Callaway, B. & Sant'Anna, P. H. C. *Did: Treatment Effects with Multiple Periods and Groups*. (2022).
19. Butts [aut, K., cre, Gardner, J., McDermott, G. & Berge, L. *Did2s: Two-Stage Difference-in-Differences Following Gardner (2021)*. (2023).