

Supplementary Materials

Supplementary Table S1: Descriptive Statistics of all variables

Statistic	Mean	St. Dev.	Median	Min	Max	N
Proportion of food crop deforestation in reserves (2000-2019)	0.387	0.266	0.320	0.013	0.969	76
Proportion of food crop deforestation outside reserves (2000-2019)	0.025	0.022	0.018	0.000	0.089	76
Proportion of forest loss in reserves (2022-2024)	0.097	0.091	0.077	0.001	0.454	76
Proportion of forest loss outside reserves (2022-2024)	0.013	0.014	0.008	0.000	0.096	76
Proportion of cocoa saturation outside reserves	0.765	0.162	0.795	0.176	0.979	76
Proportion of cocoa saturation inside reserves	0.101	0.156	0.042	0.000	0.788	76
Proportion of shade tree cover in cocoa areas	0.147	0.021	0.147	0.109	0.211	76
Proportion of remaining forest inside reserves (2000)	0.909	0.134	0.959	0.198	1.003	76
Proportion of remaining forest outside reserves (2000)	0.255	0.150	0.237	0.019	0.694	76
Proportion of remaining forest inside reserves (2021)	0.733	0.226	0.804	0.135	1.002	76
Proportion of remaining forest outside reserves (2021)	0.233	0.105	0.228	0.049	0.505	76
Number of cocoa buying company per village	5.066	1.907	5.000	1.000	11.00	76
Proportion of companies implementing sustainability initiatives within a village	0.356	0.227	0.333	0.000	1.000	76
Terrain Ruggedness Index inside reserves	3.739	1.635	2.975	1.720	7.650	76
Distance from village to nearest capital (km)	11.356	5.463	10.667	0.213	27.177	76
Distance to nearest main road (km)	4.701	4.858	3.376	0.001	15.184	76
Population Density (2010, per 1000)	1.750	1.393	1.384	0.268	7.028	76
Population Density (2020, per 1000)	1.877	1.598	1.373	0.299	6.941	76

Note: Statistics summarise village-level data for 76 villages in Ghana's High Forest Zone. Food crop deforestation proportions represent the share of village land area (inside and outside reserves) under food crop production. Forest loss proportions represent share of village land area (inside or outside

reserves) affected. Cocoa saturation is the proportion of off-reserve and inside-reserve land under cocoa. Shade tree cover is the percentage of cocoa area with canopy cover. Company variables reflect number and sustainability programme implementation. Terrain ruggedness index measures topographic variation. Distance variables measured in kilometres. Population density in persons per hectare (2020 estimates). N=76 villages for all variables.

Supplementary Table S2: Average Marginal Effects (AME) from Beta Regression Models for 2019 and 2022–2024 Periods

Variable	2000-2019	2022-2024 (AME, SE)	
	(AME, SE)	Model 1	Model 2
Proportion of cocoa saturation outside reserves	0.654*** (0.250)	0.151** (0.059)	0.144** (0.059)
Percentage of shade tree cover in cocoa areas	—	0.004 (0.004)	0.003 (0.004)
Proportion of companies implementing sustainability initiatives within a village	—	-0.094* (0.048)	-0.098** (0.043)
Number of cocoa sourcing companies in a village	—	-0.002 (0.005)	-0.002 (0.005)
Modified Taungya system (MTS)	—	0.030** (0.015)	0.034** (0.015)
Distance from village to nearest capital (km)	-0.011 ** (0.005)	-0.004* (0.002)	-0.005** (0.002)
Population density (2010, per 1000)	-0.040 (0.027)	-	-
Population Density (2020, per 1000)	-	0.011* (0.006)	0.010** (0.005)
Distance to nearest main road (km)	-0.001 (0.010)	-0.001 (0.002)	-0.001 (0.002)
Terrain ruggedness index inside reserves	-0.047* (0.026)	0.003 (0.007)	0.001 (0.007)
Proportion of undisturbed forest inside reserves (2021)	—	—	0.0034 (0.028)

Note: AMEs represent the average change in predicted forest loss (as proportion of village land) for a one-unit change in each predictor, averaged across all observations. These facilitate the interpretation of beta regression results on the original proportion scale. For 2000-2019, the dependent variable is food crop-driven forest loss in reserves. For 2022-2024, the dependent variable is total forest loss in reserves. Model 1 (2022-2024) excludes baseline forest condition; Model 2 includes it. '—' indicates variable not included in that period's model. Robust standard errors in parentheses. N=76 villages. *** p<.01, ** p<.05, * p<.1

Supplementary Table S3: Spatial Autocorrelation Test for both periods

Statistic	Value (2000-2019 period)	Value (2022-2024 period)
Moran I statistics	-0.0132	-0.0132
Expectation	-0.0132	-0.0132
Variance	5.13×10^{-17}	4.81×10^{-17}
Standard deviate	2.42×10^{-10}	-2.49×10^{-10}
p-value	0.5	0.5
Alternative hypothesis	Greater	Greater

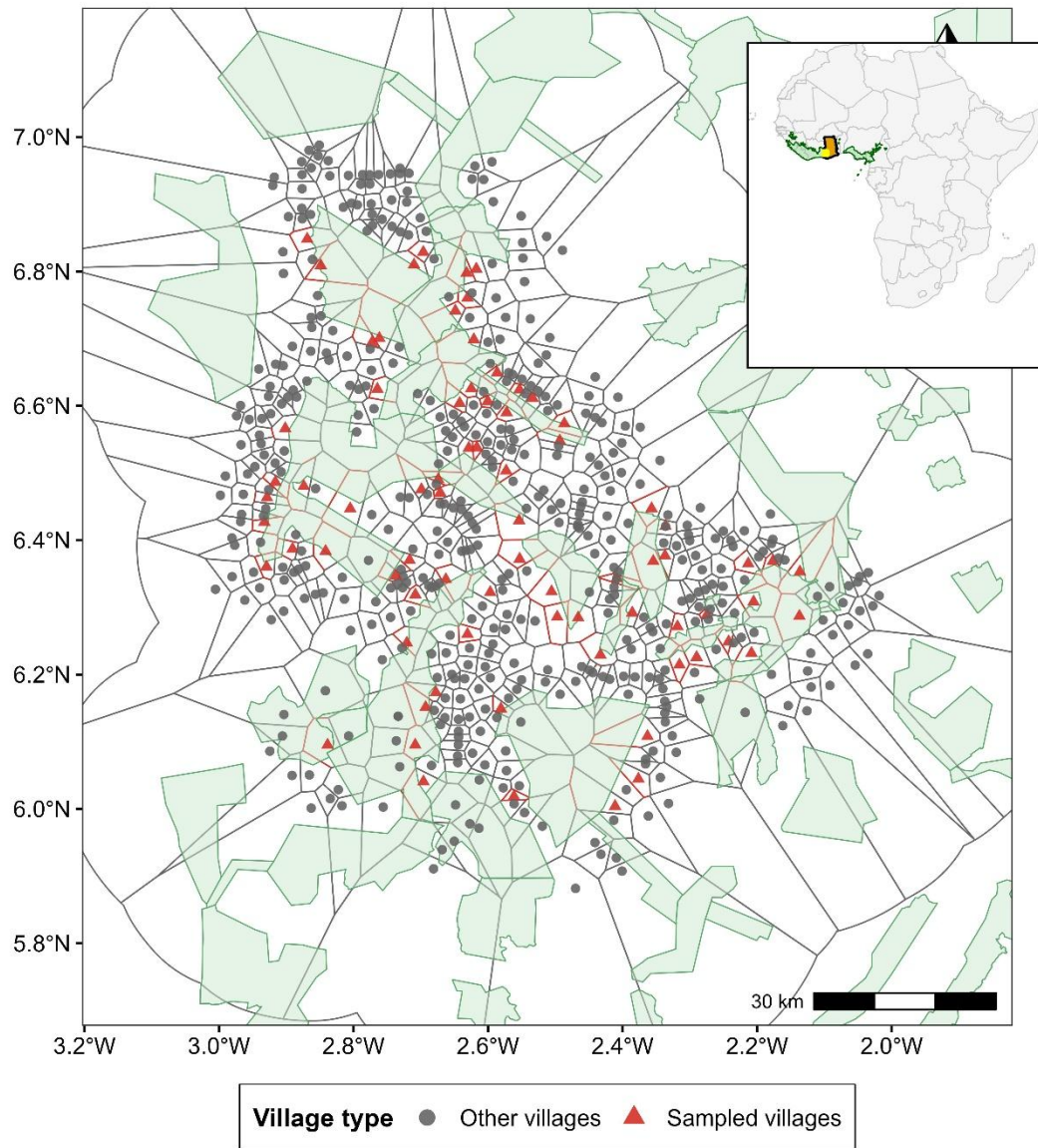
Note: Moran's I statistic tests whether forest loss values in neighbouring villages are more similar than expected by chance (spatial autocorrelation). Values near zero indicate no spatial autocorrelation; positive values indicate clustering of similar values. Standard deviation and p-values test the null hypothesis of no spatial autocorrelation. Results show no significant spatial autocorrelation in either period ($p=0.5$), indicating observations are spatially independent and standard regression approaches are appropriate. $N=76$ villages.

Supplementary Table S4: Balance tests comparing study villages ($N=76$) and full village sample ($N=298$)

Variable	Study villages (n =76)			Larger village sample (N = 298)				
	Mean	St. Dev	Median (Range)	Mean	St. Dev	Median (Range)	t test (p)	KS test (p)
Proportion of food crop forest loss inside reserves	0.08	0.09	0.05 (0; 0.42)	0.05	0.07	0.02 (0; 0.42)	0.03	0.028
Proportion of cocoa saturation outside reserves	0.76	0.17	0.79 (0.18; 0.98)	0.75	0.14	0.77 (0.18; 0.98)	0.645	0.603
Proportion of cocoa shade cover	0.15	0.02	0.15 (0.11; 0.21)	0.14	0.02	0.15 (0.09; 0.21)	0.439	0.937
Terrain ruggedness index	3.77	1.65	3.01 (1.72; 7.65)	3.58	1.5	2.95 (0.92; 8.8)	0.358	0.7
Average Distance to road (km)	4.68	4.83	3.35 (0; 15.18)	4.09	4.34	2.9 (0; 17.36)	0.33	0.492

Note: Balance tests comparing 76 study villages with the full 298-village population. T-tests compare means; Kolmogorov-Smirnov (KS) tests compare distributions. P-values >0.05 indicate no significant difference. Results confirm representativeness: no significant differences in cocoa saturation ($p=0.645$), shade cover ($p=0.439$), terrain, or accessibility (KS tests all $p>0.05$). Study villages show slightly higher food crop coverage inside reserves (t-test $p=0.030$), making them more relevant for studying

displacement. Elevation in meters; terrain ruggedness index measures topographic variation; food crop suitability is a composite biophysical measure.



Supplementary Fig S1: Thiessen polygons for all villages within 10 km of forest reserves. Polygons are coloured by village type: red triangles mark the 76 villages included in the statistical analysis, while grey circles mark other villages in the broader universe used to construct Thiessen boundaries. Forest reserves are shown in green. Including all nearby villages ensures valid polygon boundaries for adjacent settlements.