Supplementary Table S1. *OLS Regression Results: Predictors of Restoration Success Scores (n = 147 Projects).*

Variable	β	SE	t	p-value	VIF	95% CI
(Intercept)	8.42	3.18	2.65	0.009**	—	[2.13, 14.71]
Funding adequacy (%)	0.47	0.03	15.67	**<0.001*** **	1.82	[0.41, 0.53]
Multi-stakeholder governance[a]	14.23	2.41	5.90	**<0.001*** **	1.94	[9.47, 18.99]
Dual-agency governance[a]	7.86	2.67	2.94	0.004**	1.73	[2.58, 13.14]
Community-led governance[a]	10.51	2.89	3.64	**<0.001*** **	1.68	[4.80, 16.22]
Stakeholder engagement (no. of groups)	3.28	0.74	4.43	**<0.001*** **	2.15	[1.82, 4.74]
Technical capacity index (0–100)	0.18	0.04	4.50	**<0.001*** **	2.07	[0.10, 0.26]
Virginia[b]	−4.37	1.89	−2.31	0.022*	1.46	[−8.11, −0.63]
West Virginia[b]	−8.94	2.12	−4.22	**<0.001*** **	1.53	[−13.13, −4.75]
Model Diagnostics						
Observations (n)	147					
R^2	0.912					
Adjusted R^2	0.907					
F-statistic	$F(8, 138) = 178.4$, $p < 0.001$					
Breusch–Pagan test	$\chi^2 = 11.34$, $p = 0.183$ (homoscedasticity not rejected)					
Ramsey RESET test	$F = 1.87$, $p = 0.138$ (no specification error detected)					
Max VIF	2.15 (all VIF < 3.0; no multicollinearity concern)					
Durbin–Watson statistic	1.94 (no autocorrelation)					

Significance levels: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

[a] *Reference category for governance model: single-agency.*

[b] *Reference category for state: Maryland.*

Interpretation: The model explains 91.2% of variance in restoration success scores. Funding adequacy is the strongest predictor ($\beta = 0.47$, $p < 0.001$): each percentage-point increase in budget fulfillment corresponds to a 0.47-point increase in success scores. Multi-stakeholder governance yields a 14.23-point premium over single-agency models after controlling for other factors, consistent with the 2.3-fold improvement reported in bivariate analysis. The Virginia ($\beta = -4.37$, $p = 0.022$) and West Virginia ($\beta = -8.94$, $p < 0.001$) dummy variables reveal significant negative effects relative to Maryland, confirming that state-level institutional characteristics explain meaningful variation in outcomes even after controlling for project-level predictors. West Virginia's larger negative coefficient reflects documented constraints in institutional capacity, funding stability, and post-mining degradation severity. All VIF values remain below 3.0, indicating no multicollinearity concern. The non-significant Breusch–Pagan and Ramsey RESET tests confirm appropriate model specification and homoscedastic residuals.

Supplementary Table S2. *Restoration Success Scoring Rubric (0–100 Scale).*

Restoration success was assessed using a composite metric integrating four equally weighted domains (25 points each). Each domain comprises multiple indicators scored on objective criteria. Projects scoring ≥70/100 were classified as 'successful,' 50–69 as 'partially successful,' and <50 as 'unsuccessful.'

Domain	Indicator	Max Points	Scoring Criteria	Threshold / Benchmark
Ecological (25 pts)	Native species establishment	8	% of planted native species surviving and establishing at 3-year post-restoration assessment	≥65% = full points (8) 50–64% = 5 pts <50% = 2 pts
	Structural complexity index	7	Multi-strata canopy development index (0–1.0) measuring vertical and horizontal heterogeneity	≥0.6 = full points (7) 0.4–0.59 = 4 pts <0.4 = 2 pts
	Canopy closure	5	% canopy cover measured via hemispherical photography at 3-year assessment	≥40% = full points (5) 25–39% = 3 pts <25% = 1 pt
	Soil stabilization	5	Erosion reduction index based on Universal Soil Loss Equation (USLE) comparison pre/post restoration	≥50% reduction = full (5) 25–49% = 3 pts <25% = 1 pt
Social (25 pts)	Stakeholder satisfaction	10	Mean satisfaction score from structured surveys of local stakeholders (5-point Likert scale, converted to %)	≥70% satisfied = full (10) 50–69% = 6 pts <50% = 3 pts
	Employment generation	8	Full-time equivalent (FTE) jobs created per 100 acres of restoration area	≥5 FTE/100 ac = full (8) 3–4.9 FTE = 5 pts <3 FTE = 2 pts
	Community knowledge transfer	7	Documented training events, workshop participation, and educational outputs for local communities	≥3 events + >50 participants = full (7) 1–2 events = 4 pts None = 1 pt
Economic (25 pts)	Cost-effectiveness	10	Total restoration expenditure per acre relative to regional benchmark ($3,000/acre)	≤$3,000/ac = full (10) $3,001–$4,500 = 6 pts >$4,500 = 3 pts
	Sustainable revenue generation	8	Documented revenue streams (ecosystem service payments, sustainable harvest,	Revenue ≥20% of maintenance costs = full (8) 10–19% = 5 pts <10% = 2 pts

Domain	Indicator	Max Points	Scoring Criteria	Threshold / Benchmark
			recreation fees) generated post-restoration	
	Budget utilization	7	% of allocated budget effectively utilized within planned timeline	≥85% utilized = full (7) 70–84% = 4 pts <70% = 2 pts
Institutional (25 pts)	Adaptive management protocols	9	Documented adaptive management plan with triggers, decision criteria, and response protocols	Comprehensive plan with ≥3 trigger-response protocols = full (9) Basic plan = 5 pts No plan = 2 pts
	Monitoring systems	9	Operational ecological and social monitoring systems with defined frequency, indicators, and reporting	Active system with quarterly+ reporting = full (9) Annual reporting only = 5 pts No system = 2 pts
	Long-term stewardship	7	Documented stewardship commitments (conservation easements, maintenance agreements, institutional mandates) extending ≥5 years post-project	≥10-year commitment = full (7) 5–9 years = 4 pts <5 years = 2 pts

Classification: Total score ≥70/100 = 'Successful'; 50–69/100 = 'Partially Successful'; <50/100 = 'Unsuccessful.'

Note: *Scores within each indicator were assigned by trained assessors using standardized field protocols and verified documentation. Inter-rater reliability exceeded Cohen's kappa > 0.80 for all domains. Where field assessment was not feasible, scores were assigned based on project documentation review with conservative estimation.*

Supplementary Table S3. *Sustainability Score Criteria (0–100 Scale).*

Sustainability was assessed 3–5 years post-implementation using a composite index integrating five dimensions (0–20 points each). Each dimension comprises sub-indicators scored on objective criteria. Total scores range 0–100: High sustainability (≥70), Moderate (50–69), Low (<50).

Dimension (Weight)	Sub-Indicator	Scoring Criteria	Max Points	Point Allocation
Ecological Resilience (0–20 pts)	Native species persistence	% of planted native species persisting at 3–5 year assessment, accounting for natural recruitment	7	≥75% persistence = 7 60–74% = 5 <60% = 2
	Forest structure complexity	Multi-strata development: ground cover, understory, midstory, and canopy layer establishment	7	≥3 strata developed = 7 2 strata = 4 ≤1 stratum = 2
	Ecological connectivity	Contribution to landscape-level habitat connectivity measured through proximity to existing forest patches and corridor function	6	Connected to ≥2 patches = 6 Connected to 1 = 3 Isolated = 1
Community Benefits (0–20 pts)	Local employment retention	FTE jobs sustained ≥2 years post-project completion in restoration maintenance, monitoring, or derived livelihoods	7	≥3 FTE sustained = 7 1–2 FTE = 4 0 FTE = 1
	Educational outcomes	Documented educational programs, school partnerships, interpretive materials, or citizen science initiatives linked to restoration site	7	≥2 active programs = 7 1 program = 4 None = 1
	Equitable benefit distribution	Evidence that restoration benefits (jobs, access, ecosystem services) reach underserved or adjacent communities	6	Documented equitable access = 6 Partial = 3 No evidence = 1
Institutional Capacity (0–20 pts)	Monitoring system operation	Active ecological and social monitoring with regular data collection, analysis, and reporting cycles	7	Quarterly+ active monitoring = 7 Annual monitoring = 4 No monitoring = 1
	Adaptive management implementation	Documented evidence of management adjustments based on monitoring data (e.g., revised planting,	7	≥2 documented adaptations = 7 1 adaptation = 4 No evidence = 1

Dimension (Weight)	Sub-Indicator	Scoring Criteria	Max Points	Point Allocation
		invasive control, engagement strategies)		
	Institutional continuity	Sustained organizational commitment: dedicated staff, ongoing budget allocation, and formal stewardship agreements	6	Formal agreement + dedicated staff = 6 Partial commitment = 3 No continuity = 1
Financial Viability (0–20 pts)	Cost-effectiveness over time	Ratio of cumulative benefits (ecosystem services valuation, avoided costs) to total expenditure at 3–5 year assessment	7	Benefit-cost ratio ≥1.5 = 7 1.0–1.49 = 4 <1.0 = 2
	Long-term funding security	Documented multi-year funding commitments, endowments, trust funds, or self-sustaining revenue mechanisms	7	≥5-year secured funding = 7 2–4 years = 4 <2 years = 1
	Funding diversification	Number of distinct funding sources (federal, state, private, PES, green bonds) contributing to ongoing operations	6	≥3 sources = 6 2 sources = 3 1 source = 1
Stakeholder Engagement (0–20 pts)	Participation breadth	Number of distinct stakeholder groups (federal, state, NGO, academic, community, private landowners) actively participating in governance	7	≥5 groups = 7 3–4 groups = 4 ≤2 groups = 1
	Decision-making inclusivity	Evidence that diverse stakeholders hold meaningful decision-making authority (not merely consultative roles) in restoration planning and management	7	Formal co-governance = 7 Advisory board with input = 4 Consultation only = 1
	Sustained participation	Documented continued stakeholder involvement ≥2 years post-initial implementation, measured through meeting attendance, volunteer hours, and governance participation	6	Active sustained engagement = 6 Decreasing participation = 3 Disengaged = 1

Classification: Total score ≥70/100 = 'High Sustainability'; 50–69/100 = 'Moderate Sustainability'; <50/100 = 'Low Sustainability.'

Note: *Sustainability assessments were conducted 3–5 years post-implementation to allow sufficient time for ecological establishment, institutional consolidation, and community capacity development. Assessments integrated field observations, project documentation review, stakeholder interviews, and remote sensing data where available. Assessors received standardized training and calibration protocols; inter-rater reliability exceeded Cohen's kappa > 0.82 across all dimensions.*

Relationship to manuscript: *Sustainability scores reported in Section 3.3 (Stakeholder Engagement) and Figure 3B were calculated using this rubric. The five-dimension structure incorporates ecological persistence (30% conceptual weight in composite interpretation), institutional support (20%), community ownership (20%), financial mechanisms (15%), and adaptive management (15%), as described in the main text.*