Supporting Information

Nolte et al. 10.1073/pnas.1214786110

SI Materials and Methods

Data. ***Protected areas.*** We considered all protected areas included in the World Database of Protected Areas (WDPA) (1) situated in the Brazilian Legal Amazon. We used spatial data from the 2010 version of the WDPA as it included the original boundaries of protected areas that had recently been subject to downsizing as a result of their failure to stem deforestation (2). For example, the National Forest Bom Futuro had been significantly downsized in 2010 to exclude deforestation that had occurred between 2000 and 2010. We used 2012 data from the National Cadaster of Protected Areas (CNUC) of the Brazilian Ministry of the Environment to ensure that our pool of potential controls (unprotected forest parcels) did not any contain parcels situated in recently established protected areas or protected areas with expanded boundaries. We excluded from the pool of potential controls all unprotected forest parcels situated within 10-km buffers around any protected area (both the WDPA and CNUC) to reduce the vulnerability of our results to potential local spillover effects (3).

Deforestation. We used two different deforestation datasets to draw on their respective strengths in detecting tropical deforestation. The fine-grained PROgrama de Cálculo do DESflorestamento na Amazonia (PRODES) dataset published by the Brazilian Institute for Space Research (Instituto Nacional de Pesquisas Espaciais) is based on ∼30-m resolution LandSat imagery and thus capable of detecting deforestation in relatively small patches of forests (4). However, the low temporal resolution of LandSat imagery (biweekly images) hampers the detection of deforestation due to frequent cloud cover. PRODES' particularly high rate of error in early years (up to 2000) prompted us to use only 2001–2005 data for our first period of analysis. Our second deforestation measure, the Gross Forest Cover Loss (GFCL) published by South Dakota State University (5), is based on data from the Moderate Resolution Imaging Spectroradiometer (MODIS). With daily return rates, MODIS satellites are more likely to encounter cloud-free conditions. However, the lower resolution of their sensors (∼250 m) reduces their ability to detect small-scale deforestation patches (6). We ran separate analyses with both datasets and contrasted their respective results throughout.

Covariates. Probabilities of deforestation pressure and protection are influenced by a number of location-specific characteristics, most notably the suitability of a given plot for agriculture, ease of access, and distance to markets (3, 7, 8). We use the following covariates to control for differences in deforestation pressure:

- Agricultural suitability: Elevation and slope influence a forest parcel's suitability for agriculture (7). Similarly, the occurrence of seasonal flooding has been shown to influence agricultural suitability and the probability of forest conversion (9). We extracted average slope and average elevation from data provided by the International Institute for Applied Systems Analysis (10) and identified seasonally flooded areas using the GlobCover 2005 dataset based on the European Space Agency's Envisat platform (11).
- Forest cover: At ∼1-km resolution, low average tree cover on a forest parcel can indicate existing forest fragmentation and deforestation. Furthermore, the probabilities of forest conversion detected by GFCL are a function of baseline tree cover (12). We used tree cover estimates provided by the MODIS-based Vegetation Continuous Fields (VCF) dataset (collection 3) to control for this covariate (13).
- Distance to forest edge: Strongly influencing physical accessibility, the distance to the forest edge has been shown to be strongly associated with deforestation (3). We computed distance to forest edge as the shortest Euclidian distance of a given forest parcel to (i) parcels with less than 25% forest cover (VCF), (ii) rivers (ESRI hydropolygons), and (iii) major roads (14).
- Travel time to major cities: Accessibility to markets is an important predictor of deforestation patterns (7). We used the algorithm, datasets, and assumptions of an existing travel time dataset from the European Union's Joint Research Center (15) to compute our own travel time estimates using (i) improved and more detailed Brazilian road data (14) and (ii) a land cover map that reflected baseline land cover conditions in the year 2000 (MODIS Land) (16).
- State: Brazil's federal states can exercise considerable autonomy in devising state-level policies that can influence deforestation pressure and its spatial distribution. We used state boundaries provided by the Global Administrative Areas database (www.gadm.org) to control for this covariate.

We did not include distance to roads as a covariate in our analysis. Roads facilitate physical access to forest parcels and the transport of timber and agricultural products to markets. However, in the Brazilian Amazon, roads are only one element of transport infrastructure, with river travel being the main means of travel and transport in remote areas of the basin. We argue that (*i*) our estimates of travel time to major cities capture such interactions between road and river travel better than an estimate of distance to roads and that (*ii*) our estimates of distance to forest edge, with forest edge including major roads and rivers, capture the remainder of local-level variation in physical accessibility.

Methods. ***Estimating deforestation pressure.*** Matching is a quasi-experimental method that seeks to mimic random assignment of treatment by identifying artificial control groups of untreated units that differ from treated units in all relevant aspects but the treatment itself. Matching estimators rely on the assumption that treatment selection is on observables, i.e., that the observable covariates used in the matching procedure account for all differences between treatment and control units that are associated with both the probability of treatment (protection type) and the outcome (deforestation). Given the absence of randomly controlled trials of the assignment of protection to forest parcels, an explicit test of the validity of this assumption is not possible. Assessments of the validity of matching estimators therefore have to rely on (*i*) a sound theoretical and empirical argument for the choice of covariates and an (*ii*) assessment of the extent to which matching was able to balance covariates between control and treatment groups.

Choice of covariates. In the section *Covariates* above, we list the covariates included in our matching estimator, together with an empirical and theoretical rationale for the inclusion of each. Controlling for baseline forest cover, political boundaries, agricultural suitability, accessibility, and distance to markets has been considered both necessary and sufficient by a large number of matching studies that assess the impact of protection on deforestation and/or forest fires (3, 7, 17–19). One study from Costa Rica tests the sensitivity of matching estimates to using an extended set of covariates, including poverty, population density, and immigration, and finds results to be similar (3). Although we cannot explicitly test the extent to which matching successfully mimics random assignment, we consider the existing theoretical and empirical support for our choice of covariates sufficient to trust in the extent to which our estimator successfully controls for the most relevant joint bias in treatment assignment and deforestation outcomes.

Covariate balance. Matching relies on the existence of a pool of control units whose covariates are sufficiently similar to the pool of treatment units to qualify as matches (statistical support). Whether matching has been successful can be assessed by comparing covariate distributions between treated units and control units both before and after matching. A commonly used indicator to assess such similarity is the mean difference of empirical quantile-quantile (eQQ) plots of covariates in the treatment and control group (3). To obtain an aggregate balance indicator for each of the 292 protected areas, we averaged the standardized mean difference of eQQ plots across 30 repetitions and our six continuous covariates (matching was exact for categorical covariates). We then examined the distributions of the 292 estimates using Kernel density estimators, weighting each balance indicator by the number of matched forest parcel pairs (*Density Estimation*). We also examined distributions for each protection type separately.

Our results indicate that matching dramatically improved covariate balance for all protected areas in our sample (Fig. S7). Matching reduced the mean of our 292 balance estimates from 6.13 to 0.07. Furthermore, matching achieved similar improvements in covariate balance for all protection types (Fig. S8), suggesting that the remaining differences in covariates were not biased toward either protection type. We therefore consider our matching estimator to have successfully controlled for differences in observable covariates between forest parcels in control and treatment groups.

Dropped forest parcels. Causal inference through matching relies on the existence of control units that are sufficiently comparable to the pool of treated units to qualify as observations of counterfactual outcomes (statistical support). We followed earlier matching studies in removing protected forest parcels if no control parcels could be found within 1 SD of each covariate (calipers). Calipers retained 91.5% of forest parcels from the treated sample, distributed roughly equally among protection types (strict protection: 91.7%; sustainable use: 92.6%; indigenous lands: 90.9%). Visual inspection of the results suggests that protected areas with a high rate of dropped forest parcels are situated in both high- and low-pressure areas for all three protection types. The counterfactual outcome (deforestation pressure) cannot be observed for these dropped parcels. However, the large percentage of retained pixels and their distribution among protection types suggests that our results are likely to hold for the full sample of forest parcels.

Leakage. Leakage occurs when treatment influences the outcomes on untreated units. If protection of a given set of parcels leads to increased (or decreased) deforestation in unprotected parcels, a comparison of protected and unprotected units will overestimate (or underestimate) the effects of protection. A recent study did not find evidence for leakage occurring as the result of the creation of protected areas in the Brazilian Amazon (20). Nevertheless, we limited the risk of an influence of differences in local leakage on our findings by excluding from our pool of potential control parcels a 10-km buffer around all protected areas and military areas that had been created up to 2010. Although protection types may differ in the extent to which they engender leakage, the fact that our pool of control parcels covers a vast region reduces the probability that controls of different protection types may be differently affected by the leakage problem. Although we cannot rule out the possibility that leakage is occurring, we do not consider its possible existence to alter our findings about the differential impacts of protection types.

Density Estimation. We used Kernel density estimators to assess the skewness of the protection-type specific distributions of estimated deforestation pressure and to examine the shift in these distributions that occurred between 2000 and 2005 as a result of newly designated areas in all categories. We used *R*'s *density* function with a Gaussian kernel and default bandwidth computation and weighted observations by the number of matched forest parcels. We estimated density for each protection type separately (Fig. S2).

Transformations. We found that distributions of original deforestation pressure estimates were strongly skewed toward low levels of deforestation pressure (Fig. S2, *Left*). As a result, a small number of high-pressure protected areas were able to drive the differences in the aggregate estimates of pressure and impact (see main text). We also observed a strongly skewed distribution of observed deforestation rates whose variance increased with higher estimated deforestation pressure (Fig. S3). To reduce such heteroskedasticity and to allow for an estimation of pressure-specific effectiveness of protection types that would take advantage of the full sample, we transformed both observed deforestation rates and estimates of deforestation pressure. We did not use a logarithmic transformation due to the existence of real zeros in both variables. We found that a double-square-root transformation resulted in less skewed distributions and was therefore more amenable to subsequent regressions (Fig. S2, *Right*).

Regressions. ***Nonparametric regressions.*** We used locally weighted scatterplot smoothers (LOESS, using *R*'s *loess* function, span = 1) to nonparametrically estimate observed deforestation rates as a function of deforestation pressure. We computed 95% confidence intervals based on the SEs of the LOESS prediction. We applied separate LOESS estimators for each protection type, time period (2000-05 vs. 2006–10), protected area sample (established in or before 2000 vs. in or before 2005), and deforestation dataset (PRODES vs. GFCL) and compared the resulting functions (Fig. 2 and Figs. S3–S6).

Linear regressions. We used linear regressions to test the strength of the differences in pressure-specific observed deforestation between protection types. We regressed observed deforestation rates on estimated deforestation pressure (both transformed) and included dummy variables for sustainable use areas and indigenous lands. We ran models with three distinct specifications for each dataset and time period: (*i*) without interactions between pressure and protection types, (*ii*) with interactions between pressure and protection types, and (*iii*) with interaction between pressure and indigenous lands only (Table S1). The latter corresponds to our nonparametric observation that deforestation rates in indigenous lands responded differently to deforestation pressure than deforestation rates in strictly protected and sustainable use areas (see main text).

1. UNEP-WCMC (2011) *World Database of Protected Areas* (United Nations Environmental Programme-World Conservation Monitoring Centre, Cambridge, UK) Available at www.protectedplanet.net.
2. Araújo E, Barreto P (2010) Formal threats to protected areas in the Amazon. *State of the Amazon* 16:1–10.
3. Andam KS, Ferraro PJ, Pfaff A, Sanchez-Azofeifa GA, Robalino JA (2008) Measuring the effectiveness of protected area networks in reducing deforestation. *Proc Natl Acad Sci USA* 105(42):16089–16094.
4. Câmara G, De Morisson Valeriano D, Vianei Soares J (2006) *Metodologia para o Cálculo da Taxa Anual de Desmatamento na Amazônia Legal (PRODES)* (Instituto Nacional de Pesquisas Espaciais, São José dos Campos, Brazil).
5. Hansen MC, Stehman SV, Potapov PV (2010) Quantification of global gross forest cover loss. *Proc Natl Acad Sci USA* 107(19):8650–8655.
6. Hansen MC, Shimabukuro Y, Potapov P, Pittman K (2008) Comparing annual MODIS and PRODES forest cover change data for advancing monitoring of Brazilian forest cover. *Remote Sens Environ* 112(10):3784–3793.

7. Nelson A, Chomitz KM (2011) Effectiveness of strict vs. multiple use protected areas in reducing tropical forest fires: A global analysis using matching methods. *PLoS ONE* 6(8):e22722.
8. Joppa L, Pfaff A (2010) Reassessing the forest impacts of protection: The challenge of nonrandom location and a corrective method. *Ann N Y Acad Sci* 1185:135–149.
9. Newton P, Endo W, Peres CA (2011) Determinants of livelihood strategy variation in two extractive reserves in Amazonian flooded and unflooded forests. *Environ Conserv* 39(2):97–110.
10. Fischer G, et al. (2007) *Global Agro-Ecological Zones Assessment for Agriculture* (International Institute for Applied Systems Analysis, Laxenburg, Austria).
11. Arino O, Perez JR, Kalogirou V, Defourny P, Achard F (2009) *Globcover 2009* (European Space Agency, Paris).
12. Carroll M, et al. (2011) MODIS Vegetative Cover Conversion and Vegetation Continuous Fields in *Land Remote Sensing and Global Environmental Change*, eds Ramachandran B, Justice CO, Abrams MJ (Springer, New York), pp 725–745.
13. Hansen MC, et al. (2003) Global percent tree cover at a spatial resolution of 500 meters: First results of the modis vegetation continuous fields algorithm. *Earth Interact* 7(10):1–15.
14. Soares-Filho BS, et al. (2006) Modelling conservation in the Amazon basin. *Nature* 440(7083):520–523.
15. Nelson A (2008) *Travel Time to Major Cities: A Global Map of Accessibility* (Global Environment Monitoring Unit, Joint Research Centre of the European Commission, Ispra, Italy).
16. Oak Ridge National Laboratory Distributed Active Archive Center (2011) *MODIS Subsetted Land Products* (Oak Ridge National Laboratory Distributed Active Archive Center, Oak Ridge, TN).
17. Joppa LN, Pfaff A (2011) Global protected area impacts. *Proc Biol Sci* 278(1712): 1633–1638.
18. Pfaff A, Robalino J, Lima E, Sandoval C, Herrera LD (2013) Governance, location & avoided deforestation from protected areas: Greater restrictions can have lower impact, due to differences in location. *World Dev*, in press.
19. World Bank (2013) *Crossing Paths: Role of Protected Areas and Road Investments in Brazilian Amazon Deforestation. Economic Sector Report* (World Bank, Washington, DC).
20. Soares-Filho B, et al. (2010) Role of Brazilian Amazon protected areas in climate change mitigation. *Proc Natl Acad Sci USA* 107(24):10821–10826.

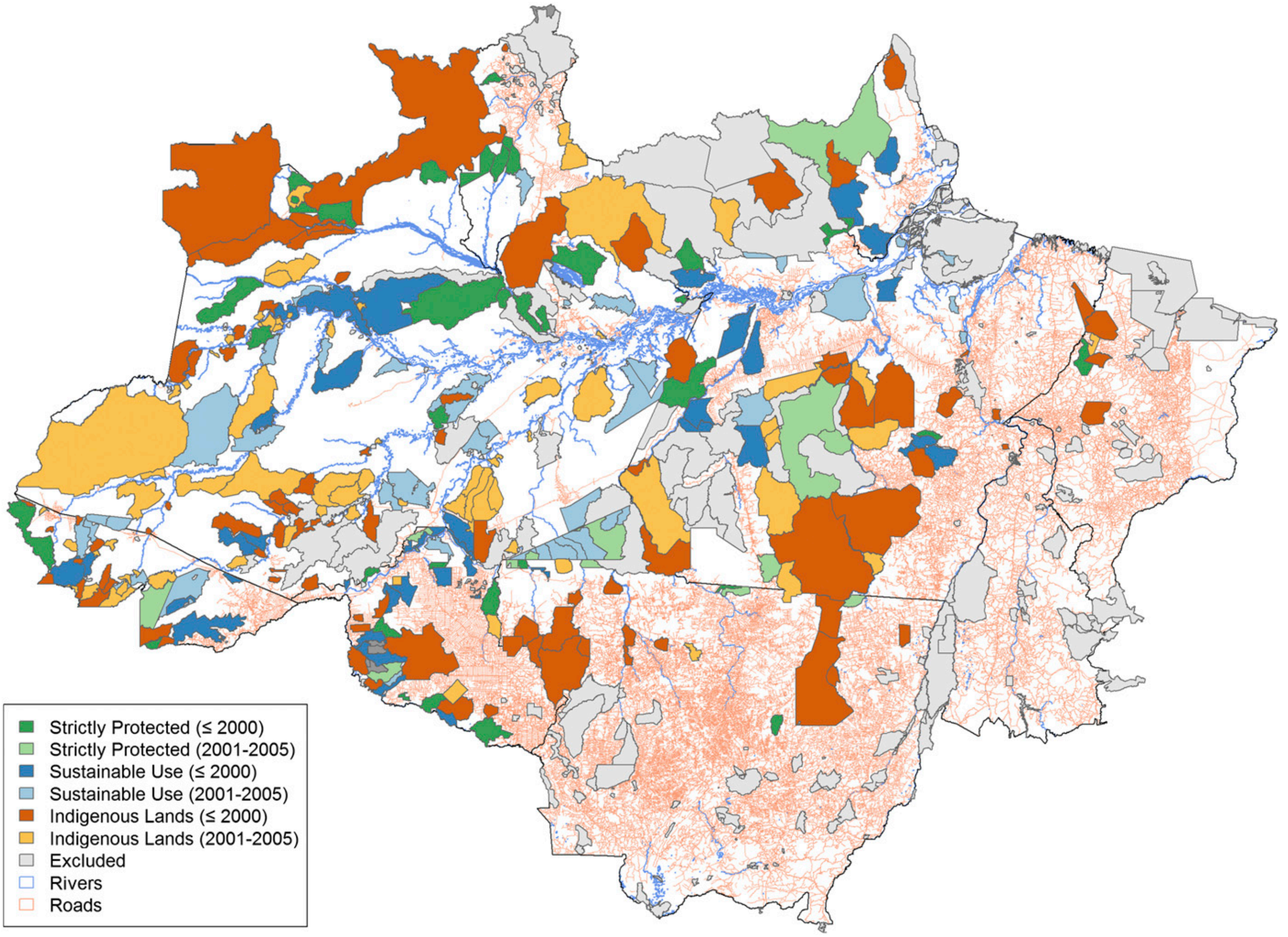

Fig. S1. Schematic map of the Brazilian Amazon protected areas included in this analysis. Excluded areas include protected areas established after 2005, Environmental Protection Areas, and protected areas outside the humid forest tropical biome with less than 50% tree cover or with fewer than 200 forest parcels in 2000. Data for this figure were provided by refs. 1 and 14.

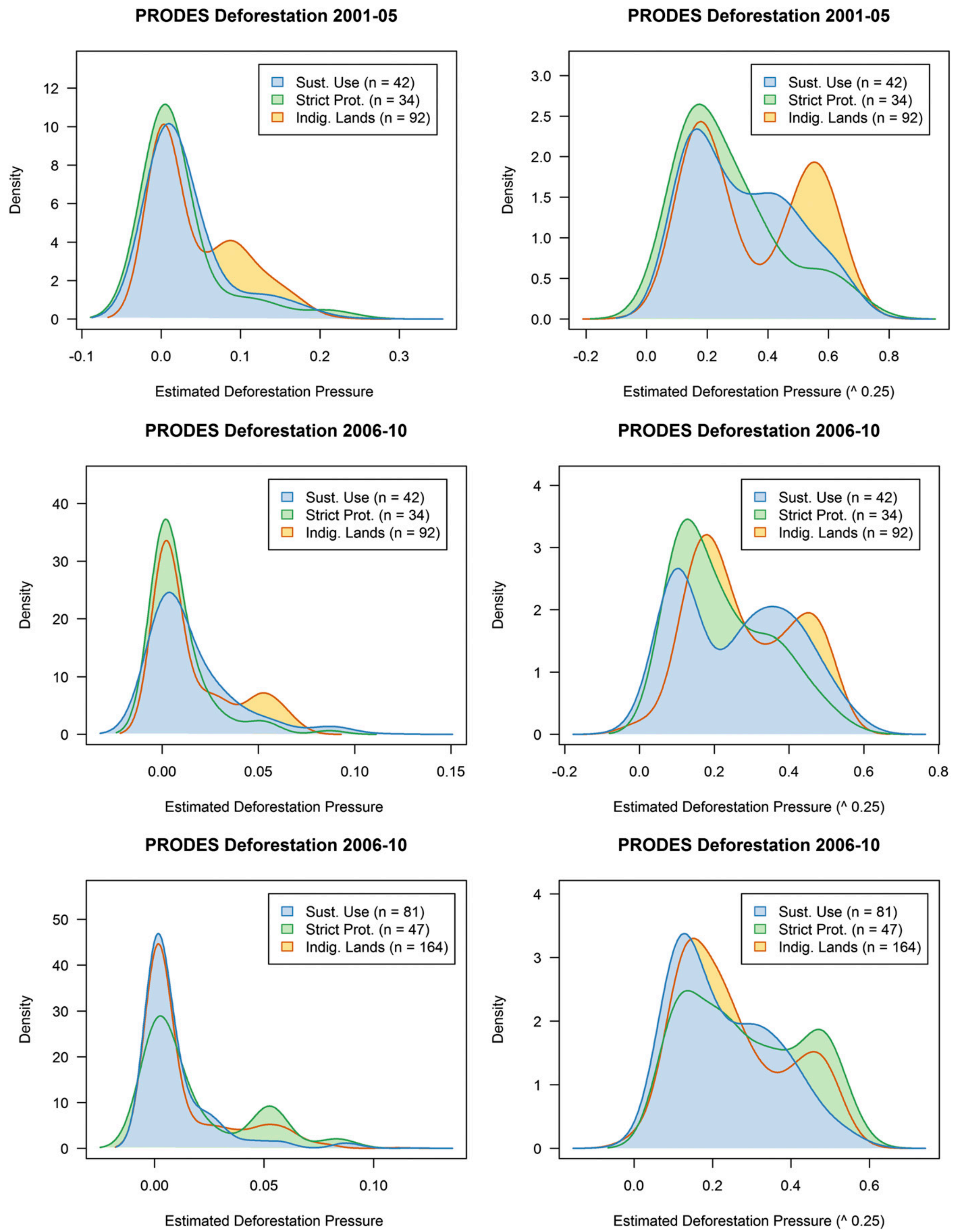

Fig. S2. Density distributions of original (*Left*) and transformed (*Right*) deforestation pressure estimates for protected areas established in or before 2000 (*Top* and *Middle*: 2001–2005 and 2006–2010 estimates, respectively) and 2005 (*Bottom*: 2006–2010 estimates). Observations were weighted by the number of matched forest parcels.

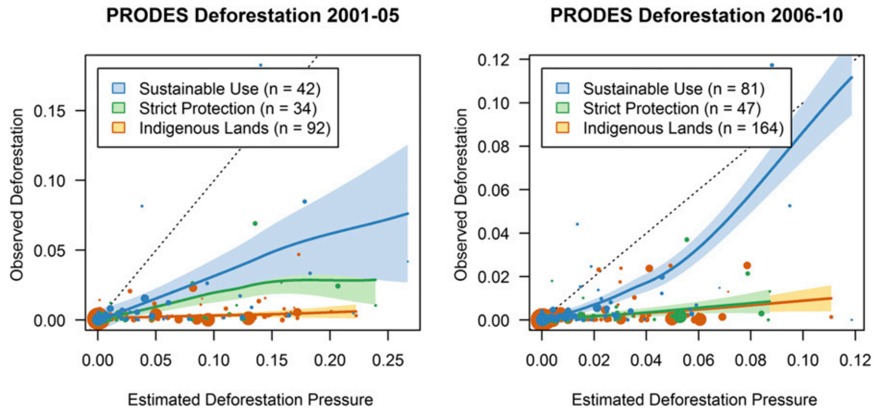

Fig. S3. As in Fig. 2, but based on original data (without transformation).

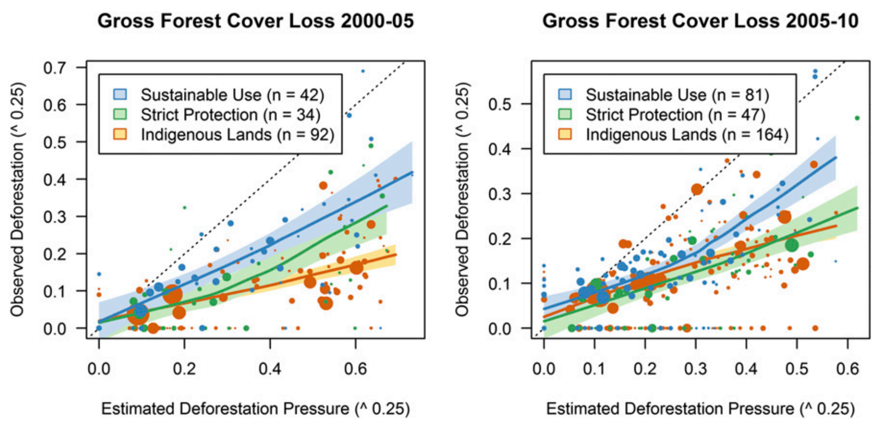

Fig. S4. As in Fig. 2, but using GFCL instead of PRODES.

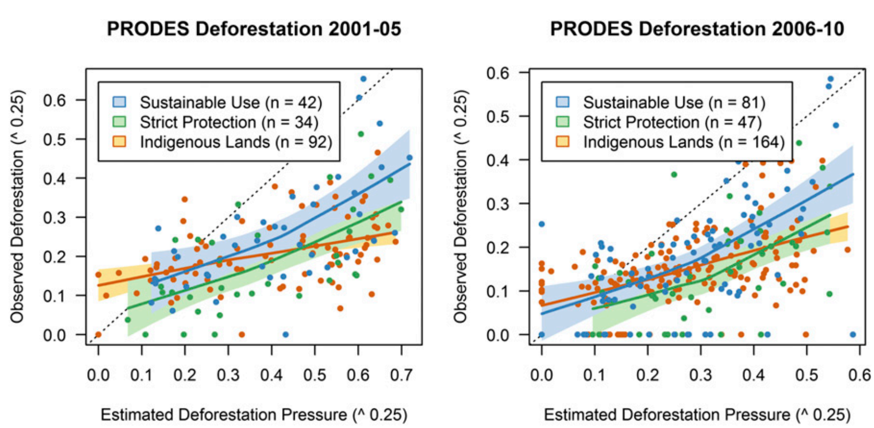

Fig. S5. As in Fig. 2, but without weighting protected areas by number of matched forest parcels.

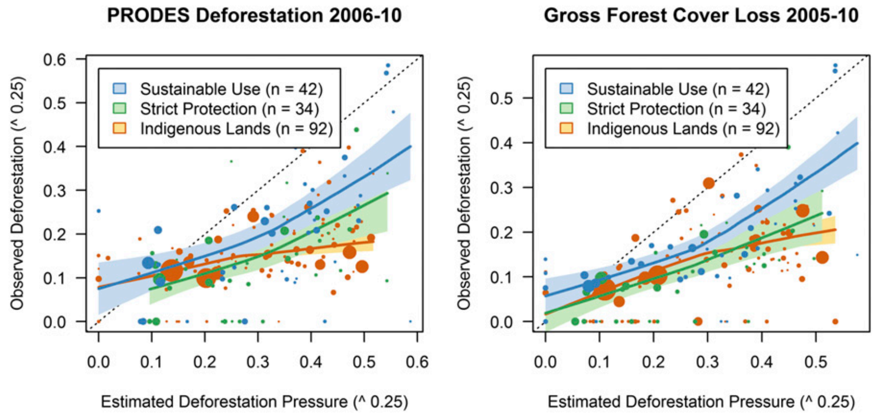

Fig. S6. As for Fig. 2 (*Right*) and Fig. S4 (*Right*), but excluding protected areas declared between 2000 and 2005 from the sample.

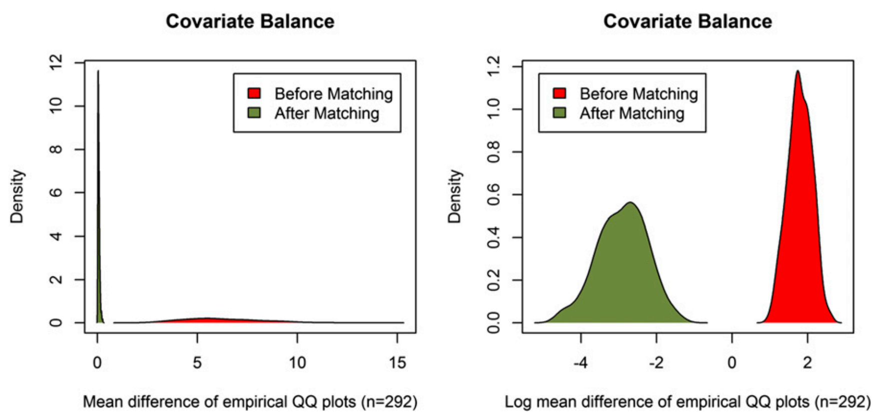

Fig. S7. Density distributions of mean standardized differences of eQQ plots (raw and log), averaged across 30 repetitions and six continuous covariates for each of the 292 protected areas considered in our analysis.

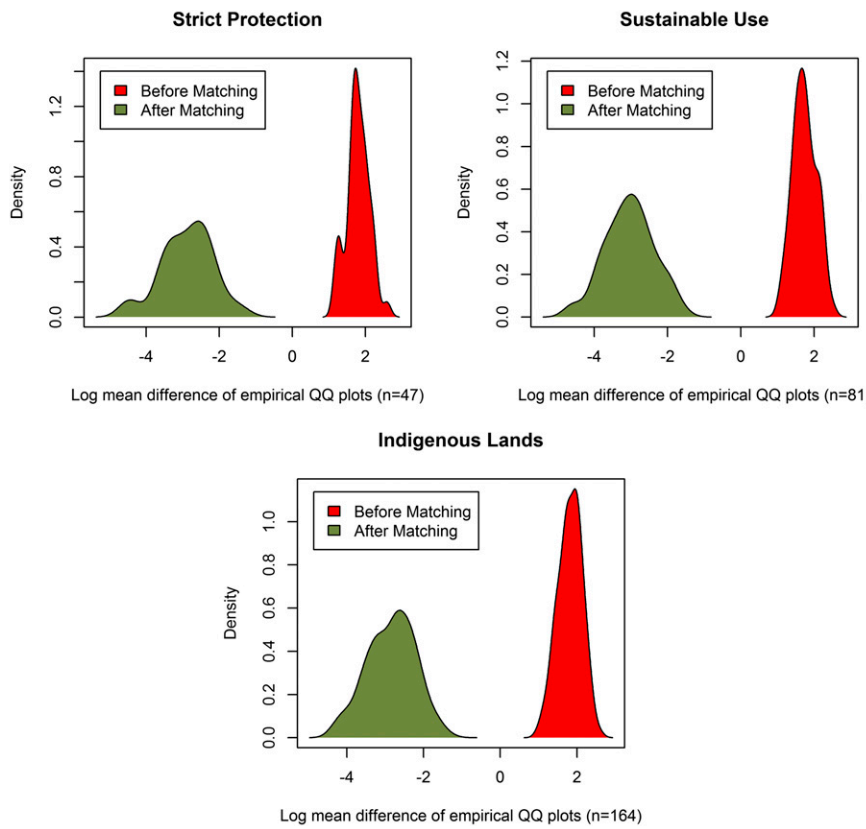

Fig. S8. Density distributions of mean standardized differences of eQQ plots (log), averaged across 30 repetitions and six continuous covariates, by protection type.

Table S1. Results of weighted regressions of observed deforestation rates on estimated deforestation pressure and protection types (transformed data)

Independent variables	Without interactions	With interactions	Interactions with indigenous lands only
PRODES deforestation 2001–2005, protected areas established in or before 2000			
Intercept	0.086***	0.014	0.010
Deforestation pressure (transformed)	0.233***	0.498***	0.515***
Sustainable use area[†]	0.056**	0.035	0.044*
Indigenous land	0.011	0.124***	0.128***
Sustainable use area × pressure		0.029	
Indigenous land × pressure		−0.382***	−0.399***
[Adjusted R^2]	[0.259]	[0.390]	[0.394]
PRODES deforestation 2006–2010, protected areas established in or before 2005			
Intercept	0.036**	0.021	−0.004
Deforestation pressure (transformed)	0.351***	0.403***	0.489***
Sustainable use area	0.043***	0.009	0.050***
Indigenous land	0.011	0.046+	0.071***
Sustainable use area × pressure		0.158+	
Indigenous land × pressure		−0.135+	−0.221***
[Adjusted R^2]	[0.351]	[0.386]	[0.381]
Gross Forest Cover Loss 2000–2005, protected areas established in or before 2000			
Intercept	0.018	−0.014	−0.023
Deforestation pressure (transformed)	0.327***	0.468***	0.510***
Sustainable use area	0.051**	0.023	0.042*
Indigenous land	−0.025	0.029	0.039+
Sustainable use area × pressure		0.075	
Indigenous land × pressure		−0.211*	−0.252***
[Adjusted R^2]	[0.478]	[0.525]	[0.526]
Gross Forest Cover Loss 2005–2010, protected areas established in or before 2005			
Intercept	0.011	0.009	−0.010
Deforestation pressure (transformed)	0.395***	0.402***	0.474***
Sustainable use area	0.037**	0.009	0.042**
Indigenous land	0.015	0.028	0.047**
Sustainable use area × pressure		0.143	
Indigenous land × pressure		−0.051	−0.124*
[Adjusted R^2]	[0.423]	[0.434]	[0.431]

***$P < 0.001$, **$P < 0.01$, *$P < 0.05$, +$P < 0.1$. Bracketed values indicate sample sizes.

[†]Protection types are dummy variables. The omitted protection type is strict protection.