Supporting Information

Chhatre and Agrawal 10.1073/pnas.0905308106

SI Materials and Methods

Data Description. Our data were collected as part of the data collection effort of the International Forest Resources and Institutions (IFRI) Program (www.umich.edu/~ifri). Founded in 1992, the program collaborates with 10 research centers in nine countries. All IFRI data are collected through 10 research instruments focusing on different aspects of forests, user groups, and institutions in a given location. The data help identify connections between social and ecological processes in forested landscapes. Over the past 15 years IFRI researchers have collected data in >200 settlements in 12 countries. In the absence of lists of forest commons in each country in which data has been collected, IFRI researchers draw on their deep knowledge of the country and the region to select study sites so as to ensure variation on the hypothesized causal variables. Useful introductions to relevant sets of causal variables used by IFRI researchers can be found in refs. 1 and 2. An annual research and training seminar provides researchers a common understanding of data collection strategies and relevant concepts.

Selection Bias. The 80 forest commons in our analysis do not constitute a random sample. Therefore, care should be taken before generalizing the results. Given the lack of records and documentation regarding the nature, occurrence, spread, and extent of forest commons in developing countries, it is impossible to draw a fully random sample from the universe of cases. Care has been taken to ensure that the sample is not skewed on relevant dimensions, including those not included in the statistical model. Variable descriptions and summary statistics are provided in Table S1 and Table S2. The possibility of selection bias may mean that the statistical findings are not generalizable beyond the sample. Rather than reaching such an extreme conclusion, we suggest that because the cases were selected without a deliberate focus on outcomes, the analysis is generalizable for the range of values of the independent variables in our data.

Dependent Variable. We constructed our dependent variable by dichotomizing two continuous variables at their means—carbon storage and livelihood benefits—losing some information in the process. There are 34 forests in the sample of 80 with above average carbon storage and 41 forests with above average livelihood benefits. The choice of the average for cut-off is arbitrary but meaningful, given our research aims.

There are some cases that are close to the average on both dimensions, but are classified into different categories given the choice of average for cut-off. We recognize that this affects our results by disproportionately weighting the cases close to the average. But, because we are interested in outcomes on two dimensions, our approach allows us to analyze the joint outcomes simultaneously. As part of our examination of the data, we also checked whether there is any correlation between the two dependent variables within the four categories for the full set of cases in the sample and for the cases that fall within the four categories of high–high, low–low, high–low, and low–high combinations of carbon storage and livelihood contributions. The outcomes were uncorrelated for within the four categories as they were for the full dataset.

We considered and tried other analytical approaches to address the joint nature of carbon storage and livelihood outcomes before choosing to work with a categorical classification. Seemingly unrelated regressions (SURs), a nested class of structural equations models provides a potential approach to explore the joint distribution of errors across two equations. However, SURs only model relationships of independent variables with one dependent variable per equation. Thus, for example, even though distance is significant in our analysis and in both equations of the SURs model we estimated (but not reported), SUR does not allow us to explore the relationship of distance to both outcomes jointly in different parts of the response-factor space. We did not pursue simultaneous equations models because there is no relationship between livelihoods and carbon storage (i.e., the two outcomes are not endogenous to each other in the sample; Spearman's $\rho = -0.017$, Prob $>$ t $= 0.8781$, $n = 80$).

Ownership and Local Autonomy. Because ownership and local autonomy in rule making play an important role in our analysis, we provide an extended explanation of how these variables were constructed from the IFRI database. In the database, the variable F_OWNLAND is a nominal variable with seven categories indexing different ownership categories. Two of these categories pertain to ownership by governments at different levels, whereas the others capture variations in the form of communal ownership across different countries for which data has been collected. We created a new variable "OWNERSHIP" = 1 where governments own the land on which the forest exists, and OWNERSHIP = 0 for the remaining categories. It necessarily reduces the variation in types of ownership to a dichotomy and loses on the diversity. However, we are interested in the applicability of our findings with respect to ongoing decentralization reforms around the world. Given that such reforms typically transfer ownership from governments down toward communities/municipalities, this classification allows us to model the impact of relatively clear (even if gross) differences in ownership on carbon storage and livelihood benefits.

In the IFRI database the variable "FCONSERVE" codes for the level of strictness of conservation measures adopted in relation to the forest, as perceived by a wide cross-section of users. It has four categories, ranging from 1 = "too restrictive" to 4 = "nonexistent," with 2 = "about the right level of conservation." We created a variable "AUTONOMY" = 1 where "FCONSERVE" = 2, and 0 otherwise. This captures the idea that communities with sufficient autonomy will create rules on the basis of local knowledge of the resources that are appropriate for conservation vs. sustainable forest use. In this sense, the variable AUTONOMY in our analysis stands in for the larger theoretical argument that institutions mediate the translation of local knowledge into sustainable resource management. In our analysis, we assume that high local autonomy would result in conservation measures that are appropriate in light of local demands and the capacity of the resource system to supply.

Omitted Variables and Endogeneity. The forest commons in our sample are located in human-dominated landscapes. They are thus not comparable to pristine forests with low-to-negligible human pressures. We did not include variables measuring population-related factors in the final model because different indicators of population are strongly correlated with two theoretically important variables included in the final model—distance to settlement and distance to administrative center. These variables capture the relative cost of local monitoring and rule enforcement by state agencies. Population density (number of user households per hectare of forest) increases in our sample with the distance of human settlement to the forest commons

(Spearman's $\rho = 0.3849$, Prob $> |t| = 0.0004$), and declines with the distance of the forest commons to the nearest administrative center (Spearman's $\rho = -0.4117$, Prob $> |t| = 0.0001$). Inclusion of either population or population density variables in the model leads to very high collinearity and loss of power. We are not specifically interested in the effects of population-related factors, focusing instead on institutional variables. We believe our model estimates are not significantly biased by the omission of the population variables insofar as they are captured by the distance variables. To the extent that this correlation between distance variables and population factors holds, the coefficients for the distance variables should be overestimated slightly. However, there is no association between population variables and other independent variables of interest—ownership and autonomy (difference of means 2-sample t tests with unequal variances (Ho: diff = 0) for population density: ownership - Pr($|T| > |t|$) = 0.9698, n = 80; autonomy - Pr($|T| > |t|$) = 0.1765, n = 80). Therefore, the omission of population variables should not induce bias in the coefficient estimates of the institutional variables that are the focus of the analysis.

The model does not include variables for poverty or socioeconomic context. There are two reasons for this. One, our measure of livelihood benefits—dependence on forest products—indirectly measures poverty; the two have been shown to be highly spatially correlated (3, 4). Including an independent variable directly measuring poverty introduces a high degree of circularity into the analysis. Two, there is no relationship between poverty levels and our independent variables of interest—ownership and autonomy; there is no reason to believe that governments hand over ownership of forest commons to communities on the basis of poverty levels, or vice versa. In that case, leaving out poverty-related variables does not introduce significant bias in the model estimates.

We have also statistically checked whether community ownership is endogenous to forest size, distance to forest, and distance to administrative centers, such that governments transfer to communities only relatively small forests far from settlements and administrative centers. We report that there is no difference between community and government-owned forests for these three variables in our sample (difference of means 2-sample t tests with unequal variances (Ho: diff = 0): forest size, Pr($|T| > |t|$) = 0.7900; distance to forest, Pr($|T| > |t|$) = 0.6687; distance to administrative center, Pr($|T| > |t|$) = 0.7837; n = 80). The results are the same with respect to rule-making autonomy.

1. Agrawal A (2001) Common property institutions and sustainable governance of resources. *World Dev* 29:1649–1672.
2. Ostrom E (2001) A general framework for analyzing sustainability of social-ecological systems. *Science* 325:419–422.
3. Sunderlin WD, et al. (2005) Livelihoods, forests, and conservation in developing countries: An overview. *World Dev* 33:1383–1402.
4. Sunderlin WD, et al. (2008) Why forests are important for global poverty alleviation: A spatial explanation. *Ecol Soc* 13:24. Available at http://www.ecologyandsociety.org/vol13/iss2/art24/. Accessed July 22, 2009.

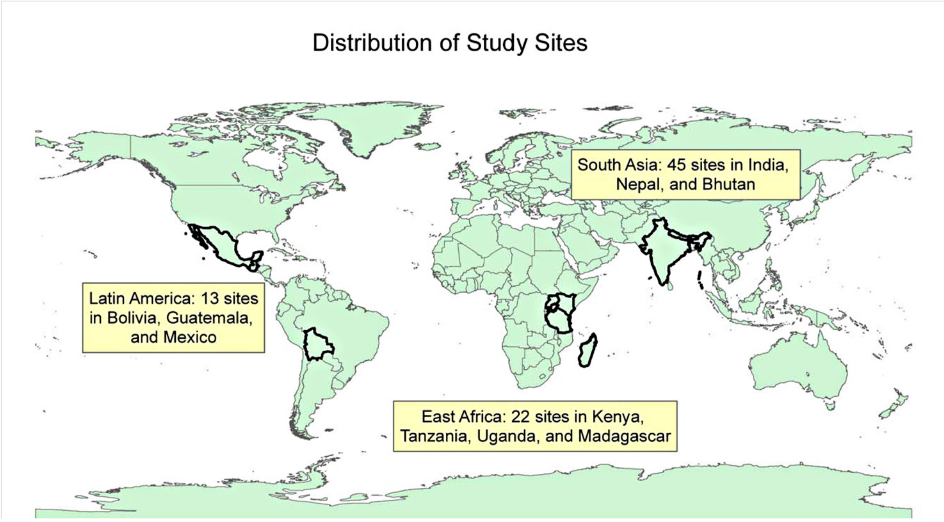

Fig. S1. The forest commons in our sample are located across Asia, Africa, and Latin America.

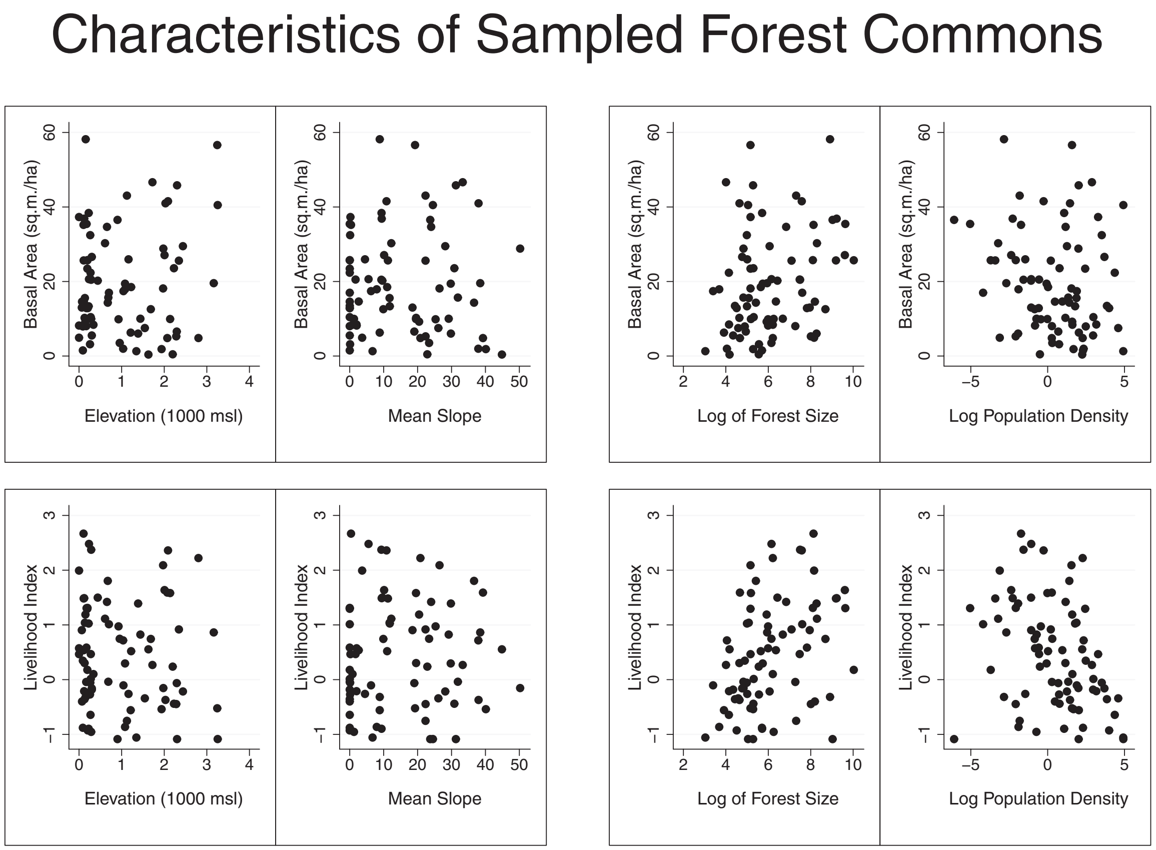

Fig. S2. Variation of carbon storage and livelihoods index: The sampled forest commons represent significant variation in their characteristics, like elevation (meters above mean sea level), slope (percentage), size (hectares), and population pressure (persons per hectare).

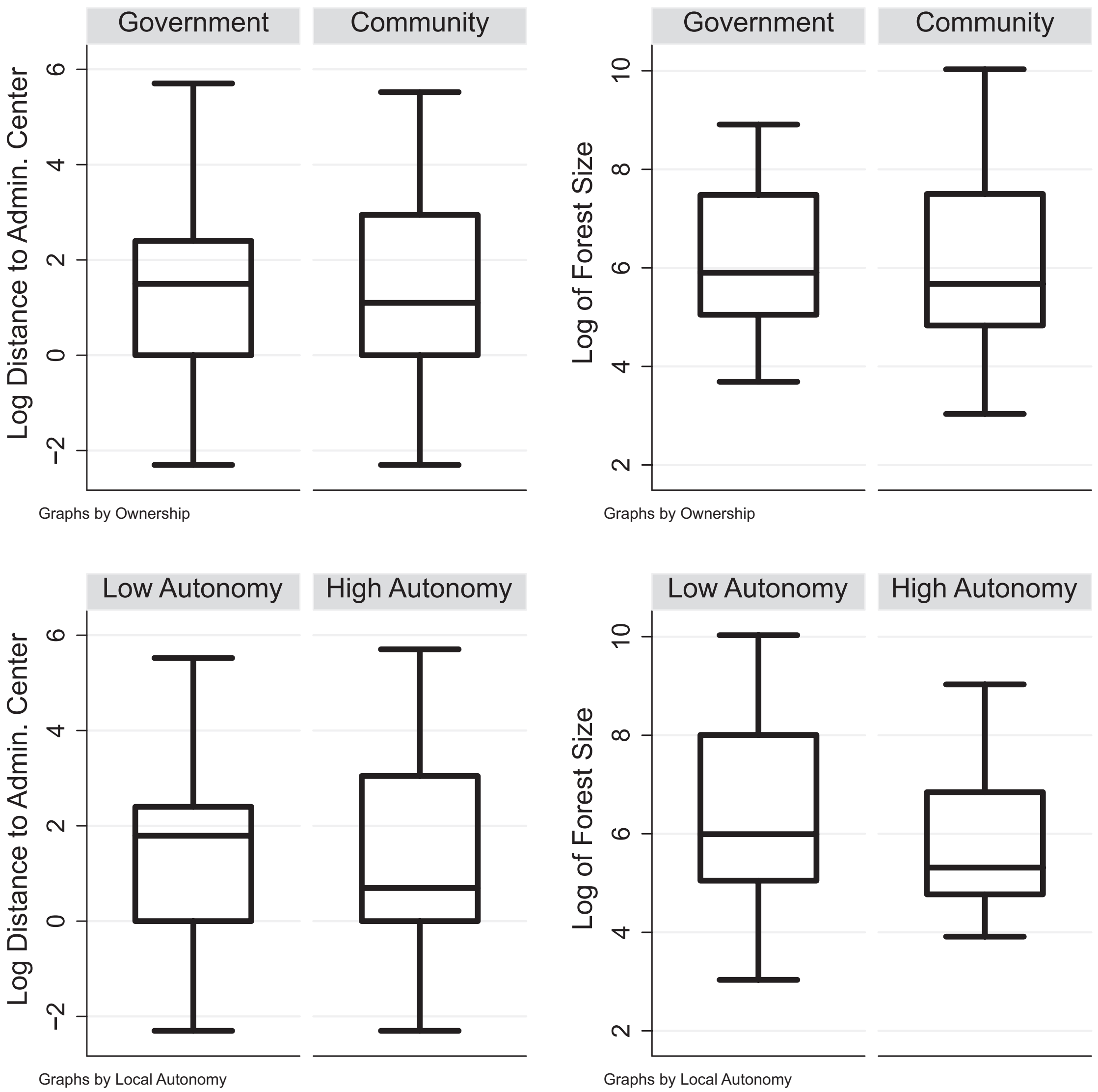

Fig. S3. The 80 forest commons in our sample vary greatly in their size and distance to administrative centers and are evenly distributed across ownership by governments or communities and local autonomy in making rules for forest management.

Table S1. Variable descriptions

Variable	Description and measurement
Carbon storage	Basal area (square meters per hectare), calculated by averaging across all trees >10 cm DBH (diameter at breast height) in >30 randomly selected 10-m radius plots.
Livelihoods index	Index extracted through factor analysis of the proportions of (*i*) food, (*ii*) firewood, (*iii*) fodder, and (*iv*) timber supplied from each forest (details below).
Ownership	Ownership of the forest commons Dichotomous: community owned = 0; government owned = 1
Autonomy	Level of autonomy to make rules at the local level. Dichotomous: low autonomy = 0; high autonomy = 1
Forest size	Size of forest commons in hectares
Distance	Average distance of users to forest commons 1 = <5 km; 2 = between 5 and 10 km; 3 = >10 km.
Admin.	Distance of forest commons to nearest administrative center, measured in kilometers.
Population density	Number of users per hectare of forest commons

Table S2. Descriptive statistics

Variable	Mean	SD	Median	Min	Max
Carbon storage	19.24	13.74	16.34	0.408	58.17
Livelihoods index	0.44	0.97	0.32	−1.085	2.66
Ownership	0.82	0.38	1	0	1
Autonomy	0.48	0.502	0	0	1
Forest size	1871.75	3856.71	301	20.8	22700
Log of forest size	6.10	1.66	5.707	3.034	10.03
Distance	1.65	0.59	2	1	3
Admin	21.6	53.04	3.5	0	300
Population density	11.45	26.31	2.05	0.0022	137.67

Table S3. Livelihoods index: Factor analysis

Factor	Eigen value	Difference	Proportion	Cumulative	Factor loadings (pattern matrix) and unique variances			
					Variable	Factor 1	Uniqueness	Scoring coefficients
Factor 1	2.13289	1.21497	0.5332	0.5332	Fodder	0.7506	0.4366	0.35193
Factor 2	0.91792	0.36992	0.2295	0.7627	Fuelwood	0.7815	0.3892	0.36642
Factor 3	0.54800	0.14680	0.1370	0.8997	Timber	0.7940	0.3696	0.37227
Factor 4	0.40120	—	0.1003	1.0000	Biomass	0.5729	0.6718	0.26861

Method: principal component factors; number of parameters = 4; retained factors = 1; rotation: unrotated; likelihood ratio test: independent vs. saturated; $\chi^2(6) = 218.18$ Prob. $> \chi^2 = 0.0000$.