Measuring the effectiveness of protected area networks in reducing deforestation

Kwaw S. Andam*†, Paul J. Ferraro†‡§, Alexander Pfaff¶, G. Arturo Sanchez-Azofeifa‖, and Juan A. Robalino**

*International Food Policy Research Institute, 2033 K Street, NW, Washington, DC 20006-1002; ‡Department of Economics, Andrew Young School of Policy Studies, Georgia State University, P.O. Box 3992, Atlanta, GA 30302-3992; ¶Terry Sanford Institute of Public Policy, Duke University, Durham, NC 27708; ‖Department of Earth and Atmospheric Sciences, University of Alberta, Edmonton, AB, Canada T6G 2E1; and **Centro Agronómico Tropical de Investigación y Enseñanza, Apdo 19, 7170 Centro Agronómico Tropical de Investigación y Enseñanza, Turrialba, Costa Rica

Edited by William C. Clark, Harvard University, Cambridge, MA, and approved August 22, 2008 (received for review January 15, 2008)

Global efforts to reduce tropical deforestation rely heavily on the establishment of protected areas. Measuring the effectiveness of these areas is difficult because the amount of deforestation that would have occurred in the absence of legal protection cannot be directly observed. Conventional methods of evaluating the effectiveness of protected areas can be biased because protection is not randomly assigned and because protection can induce deforestation spillovers (displacement) to neighboring forests. We demonstrate that estimates of effectiveness can be substantially improved by controlling for biases along dimensions that are observable, measuring spatial spillovers, and testing the sensitivity of estimates to potential hidden biases. We apply matching methods to evaluate the impact on deforestation of Costa Rica's renowned protected-area system between 1960 and 1997. We find that protection reduced deforestation: approximately 10% of the protected forests would have been deforested had they not been protected. Conventional approaches to evaluating conservation impact, which fail to control for observable covariates correlated with both protection and deforestation, substantially overestimate avoided deforestation (by over 65%, based on our estimates). We also find that deforestation spillovers from protected to unprotected forests are negligible. Our conclusions are robust to potential hidden bias, as well as to changes in modeling assumptions. Our results show that, with appropriate empirical methods, conservation scientists and policy makers can better understand the relationships between human and natural systems and can use this to guide their attempts to protect critical ecosystem services.

avoided deforestation | conservation policy | empirical evaluation | spatial spillovers

Conservation practitioners and policymakers need credible information on how policies affect ecosystems (1, 2). Despite the importance of such information, we know little about the impact of conservation policies. After a review of the evidence base, the *Millennium Ecosystem Assessment* (3) listed the following as one of its "Main Messages" (p 122): "Few well-designed empirical analyses assess even the most common biodiversity conservation measures."

One of the most common biodiversity conservation measures is the use of protected areas to reduce deforestation (3). Reducing deforestation has also become central to climate mitigation strategies (4–6). Given that protected areas now cover more than 11% of global land surface (7), an important question to ask is "How effective are protected areas in reducing deforestation?" Answering this question is complicated because "reduced deforestation" is not directly measurable. Most evaluations rely on indirect estimates based on comparisons between protected and unprotected areas. Such methods can easily be biased when protection is not randomly assigned but rather is determined by characteristics that also affect deforestation (e.g., land productivity, accessibility). Moreover, humans can respond to protection in one location by changing land uses in neighboring locations (8), and these spillovers can further bias estimates of protection's impacts.

Any analysis of a program designed to protect ecosystems and their concomitant services should include at least the following three elements: (*i*) control for bias that arises when observable biophysical and socioeconomic factors affect both which ecosystems are protected and which are most threatened; (*ii*) measurement of spatial spillovers; and (*iii*) assessment of the sensitivity of results to possible hidden bias caused by unobservable factors that affect both which ecosystems are protected and which are most threatened. By combining these three elements, we make a methodological contribution to the conservation-science literature, as well as illustrate the potential pitfalls of conventional approaches to measuring conservation impact.

Our study examines the measurement of avoided deforestation from protected areas in Costa Rica. We chose Costa Rica because it has one of the most widely lauded protected-area systems (9) and is a leader in the debate to have "avoided deforestation credits" recognized by international climate-change conventions. It also had one of the top deforestation rates during the 1960s and 1970s (10), driven mainly by the expansion of cattle grazing and coffee and banana production (11). In 1960, Costa Rica had ≈3 million hectares of forest. By 1997, more than one million hectares had been cleared and ≈900,000 hectares assigned to legal protection. We address the question, "How much more forest would have been cleared in the absence of these protected areas?"

In a review of 49 protected-area assessments (12), 13 assessments examine deforestation only in the protected areas. The other assessments compare deforestation inside and outside protected areas, and all but four find lower deforestation rates inside protected areas. Other studies use similar methods and report similar results (e.g., 13, 14). For example, Oliveira *et al.* (13) assess protected-area effectiveness by comparing deforestation rates within 20 km of roads inside and outside of protected areas. They find lower rates inside the protected areas and conclude that protected areas are effective.

Such assessments are valid only if protection were randomly assigned across the landscape. However, the *Millennium Ecosystem Assessment* (3) (p 130) reports that "many protected areas were specifically chosen because they were not suitable for human use." Empirical studies from various countries support this assertion (e.g., 9, 15–20). Thus protected and unprotected lands differ, on average, in characteristics that also affect deforestation.

A few assessments have formally controlled for such differences (19, 21–24), but they either use a small set of covariates, which can exacerbate bias when other relevant covariates are not included, or

SUSTAINABILITY SCIENCE

Author contributions: K.S.A., P.J.F., and A.P. designed research; K.S.A., P.J.F., and G.A.S.-A. performed research; K.S.A., P.J.F., and J.A.R. analyzed data; and K.S.A., P.J.F., A.P., and G.A.S.-A. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Freely available online through the PNAS open access option.

†K.S.A. and P.J.F. contributed equally to this work.

§To whom correspondence should be addressed. E-mail: pferraro@gsu.edu.

This article contains supporting information online at www.pnas.org/cgi/content/full/0800437105/DCSupplemental.

highly parametric, regression-based methods. Moreover, no published analysis has tested the sensitivity of results to hidden bias that may not have been removed by conditioning on covariates. As explained below, we use a rich data set and matching methods to control for observable sources of bias and then test the sensitivity of our results to hidden bias.

Furthermore, protected areas can influence human use on unregulated lands. Such spillovers ("leakage") can be negative (e.g., displacement of agricultural pressures, exploitation to meet tourism demand, or preemptive clearing by nearby landowners to prevent regulation of their lands). They can also be positive (e.g., enhanced law enforcement on private lands or establishment of private reserves nearby).[††] If spillovers exist, the use of surrounding unprotected lands as controls biases estimates of protection's effect. Moreover, any estimate of avoided deforestation from protection must assess the protection's impacts outside protected areas.

Results

Since the 1960s, more than 150 protected areas have been designated in Costa Rica. We estimated the effect of protection on deforestation between 1960 and 1997. Protection comprises national parks, biological reserves, forest reserves, protected zones, and wildlife reserves. We have comparable forest-cover data for 1960, 1986, and 1997 (see *Data*). To permit more accurate estimates of protection's impact using these observations, which span decades, we broke this thirty-seven-year time period into two cohorts of protected areas: (*i*) established before 1979, for which we took land that was forested in 1960 and compared deforestation in protected and unprotected forests and (*ii*) established between 1981 and 1996, for which we took land that was forested in 1986 and compared deforestation in protected and unprotected forests. No protected areas were established in 1979 or 1980 and, further, only 4% of the forest protected between 1981 and 1996 was protected in 1981 [see *Methods* and supporting information (SI) *Text* for more details].

Controlling for Overt Bias. We wished to control for differences among protected and unprotected plots across characteristics that affect both deforestation and protection decisions. Based on our knowledge of the history of Costa Rica's protected areas and the literature on tropical deforestation (11, 25–28), we controlled for a core set of variables consistently found in studies to affect deforestation: land use productivity (based on climate, soil and slope), distance to forest edge, distance to roads, and distance to nearest major city. We also controlled for an extended set of variables that includes factors whose causal effects are less clear: distance to railroads and rivers, population density, immigrants, education, poverty and size of the administrative district (see *Data*). We focus on the core covariate set here (results are similar using the extended covariate set; Tables S1–S8).

In Table 1, we assess the differences between protected and unprotected plots, before and after matching, for the pre-1979 cohort. All plots were forested in 1960 (see *Data*). The second column of Table 1 presents mean covariate values for protected plots and the third column presents mean covariate values for unprotected plots. The fourth column shows the difference in these means. Clearly, looking at the sample before matching, the inherent productivity of protected plots is much lower than that of unprotected plots; whereas >90% of unprotected plots are on high- or medium-productivity lands, only 10% of protected plots comprise such lands. Protected plots were also farther from national roads and the forest frontier than unprotected plots in the 1960s. Such characteristics tend to lower the likelihood of deforestation. Protected plots were, however, a little closer to major cities, which may decrease the likelihood of deforestation (more law enforcement) or increase it (higher market demand). A probit model that regresses a binary variable for protection on the covariates indicates that these covariates indeed influence the probability of protection, with the land-use productivity classes being the most influential. Similar covariate patterns were found for protected plots after 1981 (Table S9). Analysis with the extended covariate set indicates that protection is positively related to district size; population density; and the proportion of poor, immigrants, and educated citizens.

Table 1. Covariate balance

Variable	Mean value protect plots	Mean value control plots*	Diff mean value	Avg. raw eQQ†
High productivity land,‡ proportion				
Unmatched	0.006	0.204	−0.199	0.199
Matched	0.006	0.006	0.000	0.000
Medium productivity land, proportion				
Unmatched	0.021	0.203	−0.182	0.182
Matched	0.021	0.021	0.000	0.000
Medium-low productivity land, proportion				
Unmatched	0.073	0.507	−0.434	0.434
Matched	0.073	0.073	0.000	0.000
Distance to forest edge in 1960, km				
Unmatched	2.916	2.026	0.890	1.029
Matched	2.916	2.731	0.203	0.174
Distance to road in 1969, km				
Unmatched	17.041	15.461	1.580	1.820
Matched	17.041	16.134	0.907	0.428
Distance to city, km				
Unmatched	77.525	80.542	−3.017	16.807
Matched	77.525	77.603	−0.078	2.450

*Weighted means for matched controls.

†Mean (for categorical covariate) or median (for continuous covariate) difference in the empirical quantile-quantile plot of treatment and control groups on the scale in which the covariate is measured.

‡Dummy variables. Low productivity land is the omitted category.

Given that protection is influenced by observable characteristics that also affect deforestation, we used matching methods to estimate avoided deforestation. Matching methods are being increasingly applied as one way to establish cause–effect relationships with nonexperimental data (29). Matching works by comparing outcomes on protected and unprotected forest plots that were "very similar" in terms of the observed baseline covariates. The goal of matching is to make the covariate distributions of protected and unprotected plots similar (called covariate balancing). Matching can be viewed as a way to make the protected and unprotected covariate distributions look similar by reweighting the sample observations (e.g., unprotected plots that are poor matches receive a weight of zero). Thus, matching mimics random assignment through the *ex post* construction of a control group.

The fourth and fifth columns of Table 1 present two measures of the differences in the covariate distributions between protected and unprotected plots: the difference in means and the average distance between the two empirical quantile functions (values >0 indicate deviations between the groups in some part of the empirical distribution; Table S10 presents other balance measures). If matching is effective, both of these measures should move dramatically toward zero (30). Given the central role of agriculture in deforestation in Costa Rica, we particularly wanted good balance on land-productivity classes.

††We focused on local spillovers rather than more distant spillover effects, such as those related to changes in global market prices, which are most appropriately studied in a computable general equilibrium model.

Table 2. Estimated avoided deforestation as a proportion of forest protected

Approaches	Protected before 1979 (control: never protected and forested in 1960)	Protected after 1981 (control: never protected and forested in 1986)
Matching approaches*		
Covariate matching	−0.111 (0.029)	−0.027† (0.022)
[N matched controls]	[933]	[681]
Covariate matching with calipers	−0.124 (0.019)	−0.053 (0.010)
[N outside calipers]	[411]	[916]
{N matched controls with calipers}	{924}	{642}
Conventional conservation science approaches		
Difference in means (DIM)‡	−0.438	−0.083
DIM: controls within 10 km of protected area	−0.375	−0.131
[N available controls]	[3866]	[302]
DIM: controls within 10 km of PA, include plots deforested pre-protection	−0.497	−0.518
{N protected plots}	{1996}	{1494}
[N available controls]	[4956]	[603]
Baseline reference estimate	−0.392	−0.224
N protected plots	2711	2022
N available controls	(10371)	(4724)

*Standard errors for post-matching estimates, using variance formula in ref. 35, are in parenthesis.
†$P > 0.10$; all other estimates significant at $P < 0.01$.
‡A Chi-squared test is used to evaluate the difference in means.

The last two columns of Table 1 indicate that matching substantially improves covariate balance. Each protected plot is matched with two unprotected plots (*Methods*). The matched unprotected plots have the same distribution of productivity classes and have very similar (but not equal) distributions of the accessibility covariates. The matched unprotected plots remain slightly closer to roads and forest edges. Similar patterns of prematching covariate imbalance and postmatching balance are also observed for the post-1981 cohort (Table S9).

Instead of using matching methods, one could control for observable sources of bias by using a parametric regression analysis. We prefer matching followed by a simple test of mean forest cover change between matched protected and unprotected plots for three reasons: (*i*) we wish to make as few parametric assumptions as possible about the underlying structural model that relates protection to deforestation, and regression analysis risks a specification bias (it assumes linearity in the response surface); (*ii*) regression analysis uses observations off the common support; and (*iii*) simple postmatching comparisons of means allows us to contrast our results directly with conventional methods in the literature that depend on "inside-outside" comparisons of means. Successful matching makes treatment-effect estimates less dependent on the specific postmatching statistical model (30). To confirm that our postmatching avoided deforestation estimates are not model dependent, we also ran postmatching regressions.

Avoided Deforestation Estimates. Table 2 presents estimates of avoided deforestation as a proportion of forest protected. Estimates based on matching methods are compared with estimates based on more conventional methods in the conservation-science literature. Plots are the minimum mappable unit (3 ha each, chosen at random). Thus our outcome variable is binary: a plot is either forested or deforested (deforested, <80% canopy cover). The outcome variable is the difference between the change in forest cover on protected plots (Y = 1 if deforested) and the change in forest cover on matched unprotected plots in the same period (1960–1997 for pre-1979 protected areas; 1986–1997 for post-1981 protected areas). Thus, a negative sign indicates that protection resulted in avoided deforestation.

The first column presents results for protection before 1979. The first row presents the avoided deforestation estimates from the matching approach. It implies that 11.1% of protected plots would have been deforested by 1997 in the absence of protection ($P < 0.01$). The second row presents an estimate based on matching that uses calipers to improve covariate balance (*Methods*) (Table S11). Calipers define a tolerance level for judging the quality of the matches; if a protected plot does not have a match within the caliper (i.e., available controls are not good matches), it is eliminated from the sample. Four hundred and eleven protected plots were eliminated. They tended to be very remote plots on poor lands. Calipers reduce bias, but at the cost of estimating avoided deforestation on a subsample that may not be representative of the population of protected plots. Yet the avoided deforestation estimate of 12.4% ($P < 0.01$) is not much different from the estimate without calipers.[‡‡]

In contrast, the avoided deforestation estimates generated by conventional methods used in the conservation-science literature are much larger. The third row in Table 2 (DIM) replicates the kind of analysis done in the majority of protected area evaluations: deforestation on protected plots is compared with deforestation on unprotected plots, without controlling for other covariates. This method implies that 44% of the protected plots would have been deforested by 1997 had they not been protected before 1979.

Some of the conventional inside-outside analyses restrict the control group to an unprotected zone around each protected area (e.g., 14). Using a 10-km zone, the fourth row replicates this type of analysis and generates a slightly smaller estimate of 38%. Note that some analyses of this type (e.g., 14) do not, as we did, exclude lands already deforested at the baseline. Because protection is much less likely to be assigned to deforested plots, such methods suffer from an additional source of bias. As indicated in the fifth row of Table 2, this approach implies that 50% of protected plots would have been deforested had they not been protected.

The final row represents an estimate derived from a baseline reference, which is the most commonly suggested way of measuring avoided deforestation in climate-change negotiations. This method models deforestation in a given period as a function of observable

‡‡We also developed an approach that allows the use of plots protected after 1981 as controls by directly adjusting their observed outcomes based on the post-1981 analysis. Such an adjustment is useful when excluding such plots substantially worsens covariate balancing, which is not the case in our context (*SI Text*).

SUSTAINABILITY SCIENCE

covariates (*SI Text*). The estimated equation, based on the same core covariate set used in the matching approach, is used to predict the expected deforestation probability for each forested plot in the next period. The difference between the predicted and the actual deforestation rates is the estimated avoided deforestation. Using this method implies that 39% of protected plots would have been deforested in the absence of protection.

The dramatic differences between the estimates based on matching (first two rows) and the estimates based on methods conventionally used to evaluate protected-area effectiveness (rows three through six) suggest that the conventional methods can lead to substantially inaccurate estimates. To put Table 2's estimates into perspective, consider that 483,339 ha of forest were protected between 1960 and 1980. Thus, conventional methods imply 181,252–240,220 ha of avoided deforestation. In contrast, the matching methods imply 53,651–59,934 ha of avoided deforestation.

In column 2, the analysis for the post-1981 cohort indicates similar patterns. Matching without a caliper (first row) suggests that protection post-1981 has little effect on avoided deforestation; fewer than 3% of the protected plots would have been deforested in the absence of protection ($P > 0.10$). However, one-third of the forest protected between 1981 and 1996 was in a single protected area (La Amistad), located in a remote region near the Panamanian border. Good matches do not exist for many of the plots from this protected area. When we apply calipers, many of these remote plots are dropped and the estimated avoided deforestation rate rises to 5.3% (dropping all plots protected before 1985 increases it to 6%).[§§] These estimates imply 11,342–22,264 ha of avoided deforestation from protection after 1981. The conventional estimates are again higher than the matching estimates (1.6 to 19 times higher).

Although matching substantially improves the covariate balance between protected and unprotected plots, some imbalance remains: protected plots are slightly farther from the forest edge and from transportation infrastructure than their matched counterparts. A postmatching (weighted) regression that adjusts for any small remaining covariate imbalances yields identical estimates to those in Table 2. To test model dependence (30), we used a variety of postmatching regression specifications with the extended covariate set (i.e., match on core set but regress on variables in the extended set). We found the avoided deforestation estimates differ little from those in Table 2 (fewer than two percentage points) (Table S12).

Fig. 1 visually presents the avoided deforestation estimates for the full period 1960–1997 based on the different methods in Table 2. Matching methods suggest that between 64,993 and 82,198 ha of forest protected between 1960 and 1996 would have been deforested by 1997 in the absence of protection (7–9% of the protected area system). Conventional methods, however, imply between 236,282 and 457,818 ha of avoided deforestation (26–51%).

Controlling for Spatial Spillovers. To test for spillovers from protected lands onto nearby unprotected lands, we used matching methods to control for observable differences in unprotected lands. In this context, the treatment group comprises unprotected plots that are within a specified distance from the boundary of a protected area. The control group comprises unprotected plots that are beyond this distance.

The results suggest that the average spillover effect is small and, if it exists at all, is positive; in other words, protection reduces deforestation outside the protected area (Table S13). For the pre-1979 cohort, the postmatching estimates imply that, at most, 4.5% of the 1997 forest within two kilometers of protected areas can be classified as avoided deforestation. Neither estimate is significantly different from zero at the 1% level; only one is significant at the 5% level. We find no evidence of spillover effects beyond two kilometers or on forests near the post-1981 cohort of protected areas.

[§§]Lower rates of avoided deforestation in this period are partially because of declines in deforestation after the most productive lands had been developed, international beef prices dropped, the manufacturing and service sectors grew, and donor-imposed structural readjustment in the mid-1980s led to a decline in agricultural subsidies (25).

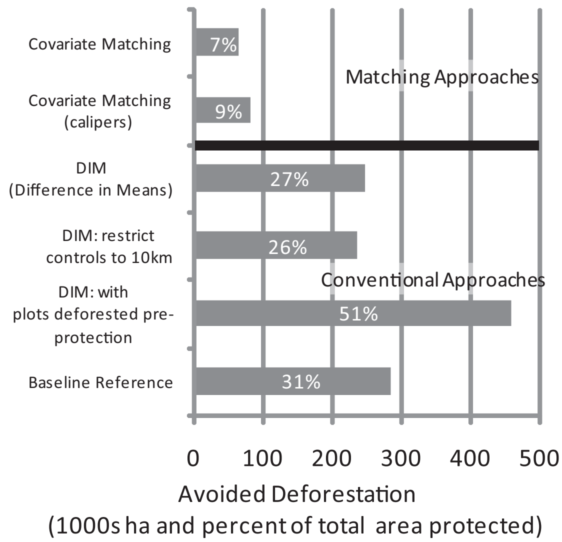

Fig. 1. Avoided deforestation estimates 1960–1997.

Given the weak evidence for spillover effects from Costa Rica's protected areas, we conclude that the matching estimates in Table 2 and Fig. 1 reflect the full effect of protected areas on deforestation within and outside protected areas between 1960 and 1997. Had we found strong evidence of spillovers, we would have controlled for potential bias from these spillovers in our estimates by excluding neighboring plots from the set of available controls or by directly adjusting the forest cover outcomes on neighboring plots based on our estimates of the extent and magnitude of the spillovers (see *SI Text*).

Sensitivity Test to Hidden Bias and Other Robustness Checks. Despite our efforts to control for observable sources of bias, protection and forest-cover change may exhibit correlation in the absence of an effect of protection because of failure to match on a relevant but unobserved covariate. In our analysis, the main concern is that protected plots may be unobservably less likely to be deforested than their matched controls. Sensitivity analysis examines the degree to which uncertainty about hidden biases in the assignment of protection could alter the conclusions of our study. We used Rosenbaum's recommended sensitivity test (31).

This test assumes that each plot has a fixed value of an unobserved covariate. The unobserved covariate not only affects protection decisions, but also determines whether deforestation is more likely for the protected plots or their matched controls. Thus, this sensitivity test is conservative. Matched forested plots differ in their odds of being protected by a factor of Γ as a result of this unobserved covariate ($\Gamma = 1$ in the absence of hidden bias). The higher the level of Γ to which the effect of protection on deforestation remains significantly different from zero, the less likely is the explanation that the avoided deforestation we detect is simply a result of matching protected plots with unprotected plots that are unobservably more likely to be deforested (details in *SI Text*).

Table 3 presents the analysis for the estimates from matching with calipers (recall the post-1981 without-calipers estimate is not statistically different from zero even in the absence of hidden bias). The second column in Table 3 indicates that our avoided deforestation estimate of 12.4% of the pre-1979 protected forest remains significantly different from zero even in the presence of moderate

Table 3. Sensitivity tests to hidden bias

	Critical p-values for treatment effects[†]		Lower bound of 99% confidence interval	
Γ	Protected pre-1979 (control: never protected and forested in 1960)	Protected post-1981 (control: never protected and forested in 1986)	Protected pre-1979 (control: never protected and forested in 1960)	Protected post-1981 (control: never protected and forested in 1986)
1.75	<0.001	<0.001		
2.00	<0.001	<0.001		
2.25	0.035	<0.001		
2.50	0.479	<0.001		
2.75	0.936	0.005		
3.00	0.998	0.020		
3.25	0.999	0.057		
1.5			−0.208	−0.074
2.0			−0.217	−0.081
3.0			−0.232	−0.098
4.0			−0.248	−0.104

[†]Test of the null of zero effect.

unobserved bias. If an unobserved covariate caused the odds ratio of protection to differ between protected and unprotected plots by a factor of as much as 2.15, the 99% confidence interval would still exclude zero (without-calipers estimate is slightly more robust). The third column indicates that the post-1981 estimate of 5.3% remains significantly different from zero ($P < 0.01$) in the presence of even greater hidden bias (up to $\Gamma = 2.85$). These results suggest that avoided deforestation is likely to be greater than zero unless there is relatively strong hidden bias.

Assuming additive treatment effects and using a different matching method (kernel matching), we can use a similar test to examine the degree to which unobserved bias could cause us to underestimate the effect of protection (*SI Text*). As with the previous test, this test is conservative. We constructed 99% confidence intervals for our estimate under varying degrees of unobserved bias and considered the lower bound. The results are presented in the fourth and fifth columns of Table 3. Even if an unobserved covariate causes the odds ratio of protection to differ between protected and unprotected plots by a factor of 4 (a substantial amount of hidden bias), the 99% confidence interval would still exclude the conventional method estimates in Table 2.

In contrast, the pre-1979 spatial spillover estimate that was significantly different from zero is not robust to even a modicum of hidden bias: if an unobserved covariate caused the odds ratio of having a protected area located near the plot to differ between matched unprotected plots by a factor of only 1.15, the 90% confidence interval would include zero (Table S14). These results provide more evidence that spatial spillovers are negligible.

Finally, we conducted robustness checks (*SI Text*) that varied the sample composition (e.g., exclude plots), the matching specifications (e.g., vary number of matches), and the spatial scale at which the analysis is conducted (i.e., use administrative districts rather than pixels as the unit of analysis). In no case did our qualitative conclusions about avoided deforestation in Costa Rica or the difference in the estimates from matching and conventional approaches change.

Discussion

Our analysis illustrates how substantial improvements can be made to estimates of protected area effectiveness. Unlike previous studies, our analysis comprises three key components: (*i*) use of available data and matching methods to control for bias that arises when observable biophysical and socioeconomic characteristics affect both which forests are protected and which are deforested, (*ii*) measurement of local spatial spillovers that controls for these same observable characteristics; and (*iii*) assessment of the sensitivity of results to possible hidden bias because of unobservable characteristics that jointly affect which forests are protected and which are deforested. Taking into account local spillovers from protection (which we find are small), we show that between 64,993 and 82,198 ha of the 903,407 ha (7–9%) of Costa Rican forest protected between 1960 and 1996 would have been deforested by 1997 in the absence of protection. If our estimates are correct, conventional methods substantially overestimate the impact of protected areas by a factor of three or more: over two-thirds of the avoided deforestation claimed by these methods would be in error. These conclusions are robust to potential hidden bias from unobservable confounding variables, as well as to alternative modeling assumptions. Our methodology can guide future studies to measure the impact of conservation policies and programs on a variety of environmental and social outcomes.

Conventional methods overestimate avoided deforestation in Costa Rica because protection was not randomly distributed across the landscape. In comparison with unprotected forests, protected forests were located on lands that were, on average, less accessible and of lower agricultural productivity. Protected forests thus had a below-average probability of being deforested in the absence of protection. This pattern of protection is common globally (3, 9, 15–20, 22). Thus, although further empirical confirmation is needed, our analysis suggests that much of what is being described as protection's impacts may result from protected-area location rather than protection itself. This knowledge can inform conservation planning. For example, it implies that recent efforts in conservation planning to jointly consider benefits, costs, and measures of threat of conversion are warranted (32). Moreover, it suggests that protecting ecosystems and their services in the future may require investments that are substantially different, in size and in nature, from those made in the past.

Understanding how the spatial distribution of land-use regulations affects deforestation is also relevant when designing Reduced Emissions from Deforestation compensation policies. Such policies contribute to climate-change mitigation by allowing polluters to purchase emission offsets in the form of avoided deforestation credits. In such schemes, the incentives for sellers to erroneously posit high counterfactual deforestation rates, and thereby claim avoided deforestation from protective efforts, are strong. Program designers should attempt to mitigate such strategic incentives.

Costa Rica's deforestation processes may continue to change over time. Thus the future impacts of protected areas may differ from those found in our retrospective analysis. Our methodology can be used to validate predictive models of protection's future impact on deforestation (e.g., 33). Moreover, protected areas can be designated for reasons other than preventing deforestation (e.g., to promote tourism or reduce hunting). Clearly, more analyses of protected-area effectiveness in other regions of the world and on other outcomes are warranted. For future decision making, how-

SUSTAINABILITY SCIENCE

ever, our analysis points to the need for rigorous empirical assessments of the impacts of conservation investments.

Materials and Methods

For more details on data and methods see *SI Text* and Table S15.

Data. Forest cover is measured from a combination of aerial photographs acquired between 1955 and 1960 (called the 1960 dataset), and from 1986 and 1997 Landsat Thematic Mapper satellite images (Earth Observation Systems Laboratory, University of Alberta, Edmonton, AB) (9). We drew a random sample of 20,000 plots (3 ha) that were forested in 1960. After removing plots that were not comparable (e.g., indigenous reserves), the final dataset comprised 15,283 land plots, including 4,762 protected plots (2,711 pre-1979) covering all protected areas except four small ones and five on islands (*SI Text* lists names of protected areas). We combined forest-cover data with spatially explicit data on covariates believed to affect both protected area location and deforestation. Geographic Information System (GIS) data layers for forest cover, protected areas, and locations of major cities were provided by the Earth Observation Systems Laboratory. Other GIS data layers included a map of land-use capacity based on exogenous factors (soil, climate, topography) from the Instituto Tecnologico de Costa Rica (San José, Costa Rica) and socioeconomic data from the Instituto Nacional de Estadistica y Censos (Cartago, Costa Rica). GIS layers for transportation roads, railroads, and the river transportation network were digitized by M. Buck Holland (Madison, Wisconsin) from hardcopy maps of 1969 and 1991 road layers (map source: Instituto Geográfico Nacional of the Ministerio Obras Publicas y Transporte of Costa Rica, San José, Costa Rica). Data are summarized in Table S16.

Methods. Using cohorts with different years for the baseline forest reduces a potential bias that can arise when using a single baseline for all protected areas. The forest landscape facing a planner in the 1980s was different from the 1960 landscape we used as a baseline for the first cohort. The clearing that occurred over those two decades was likely to have been on the best lands for clearing, and the protection decisions taken later were made on the remaining forest land. Those decisions over time suggest the 1960 forest baseline may no longer resemble the conditions faced by a planner when new protection was established in the early 1980s. We controlled for differences in observed dimensions by using matching, but the greater the differences in unobservable dimensions because of these decisions, the more potential there is for hidden bias.

Splitting our sample into two cohorts reduced this potential for hidden bias. For example, protection decisions in 1990 were made in a forest landscape very similar to the 1986 baseline. Plots still forested in 1986 are thus much better comparisons for protected areas established in 1990 than plots forested in 1960. However, using two cohorts does not eliminate the potential for hidden bias. We addressed potential hidden biases like this one through sensitivity analysis.

Based on an assessment of covariate balance quality across a variety of matching methods, we chose nearest-neighbor covariate matching using the Mahalanobis distance metric. Matching is with replacement. The mean-variance tradeoff in the match quality is resolved by using two nearest neighbors: the counterfactual outcome is the average among these two (varying the number of neighbors from one to ten changes the estimates very little). Based on recent work that demonstrates that bootstrapping standard errors is invalid with nonsmooth, nearest-neighbor matching with replacement (34), we used Abadie and Imbens' variance formula, whose asymptotic properties are well understood (35). We used a postmatching bias-correction procedure that asymptotically removes the conditional bias in finite samples (35). For caliper matching, we defined the caliper as 0.5 standard deviations of each matching covariate. We used the same matching methods to measure spillovers, rather than highly parametric, conventional spatial statistical models (e.g., a probit with spatial lag), because the latter risk a specification bias. Moreover, generating a transparent estimate of the average spillover is not easily done through interpretation of the spatial lagged coefficient (see *SI Text*). We also tried a recently created matching approach that attempts to algorithmically maximize covariate balance via a genetic search algorithm (36). We saw little difference in the results (fewer than three percentage points). Rosenbaum bounds are calculated by using the McNemar test (31).

ACKNOWLEDGMENTS. We thank the editor, five anonymous reviewers, Amy Ando, Dave Lewis, Steve Newbold, and Kate Sims for their detailed and useful comments. This work was supported by the Resources for the Future's Fisher Dissertation Award and the Global Environment Facility (K.S.A.), Inter-American Institute for Global Change Research (G.A.S.-A.), and the World Wildlife Fund's Kathryn Fuller Science for Nature Fund (P.J.F.).

1. Balmford A, *et al.* (2005) The convention on biological diversity's 2010 target. *Science* 307:212–213.
2. Ferraro P, Pattanayak S (2006) Money for nothing? A call for empirical evaluation of biodiversity conservation investments. *PLoS Biol* 4:e105.
3. Millennium Ecosystem Assessment (2005) *Ecosystems and Human Well-Being*: *Policy Responses* (Island Press, Washington, DC).
4. Gullison R *et al.* (2007) Tropical forests and climate policy. *Science* 316:985–986.
5. Righelato R, Spracklen D (2007) Carbon mitigation by biofuels or by saving and restoring forests? *Science* 317:902.
6. Schulze E *et al.* (2003) Making deforestation pay under the Kyoto Protocol. *Science* 299:1669.
7. Chape S, Blyth S, Fish L, Fox P, Spalding S (2003) *2003 United Nations List of Protected Areas.* (IUCN, Gland, and UNEP-WCMC, Cambridge, UK).
8. Armsworth P, Daily G, Kareiva P, Sanchirico J (2006) Land market feedbacks can undermine biodiversity conservation. *Proc Natl Acad Sci USA* 103:5403–5408.
9. Sanchez-Azofeifa G, Daily G, Pfaff A, Busch C (2003) Integrity and isolation of Costa Rica's national parks and biological reserves: Examining the dynamics of land-cover change. *Biological Conservation* 109:123–135.
10. FAO (1990) *Forest Resources Assessment 1990*: *Tropical Countries* (Food and Agriculture Organization of the United Nations Forestry Paper 112, Rome).
11. Sanchez-Azofeifa, G (2000) Land use/cover change in Costa Rica: A geographic perspective. In *Quantifying Sustainable Development*, eds Hall C, Leon-Perez C, Leclerc G (Academic, San Diego), pp 473–501.
12. Naughton-Treves L, Holland M, Brandon K (2005) The role of protected areas in conserving biodiversity and sustaining local livelihoods. *Annual Review of Environment and Resources* 30:219–252.
13. Oliveira P, *et al.* (2007) Land-use allocation protects the Peruvian Amazon. *Science* 317:1233–1236.
14. Bruner A, Gullison R, Rice R, da Fonseca G (2001) Effectiveness of parks in protecting tropical biodiversity. *Science* 291:125–128.
15. Green G, Sussman R (1990) Deforestation history of the eastern rain forests of Madagascar from satellite images. *Science* 248:212–215.
16. Hunter M, Yonzon P (1993) Altitudinal distributions of birds, mammals, people, forests, and parks in Nepal. *Conservation Biology* 7:420–423.
17. Pressey R (1995) Conservation reserves in NSW. Crown jewels or left overs? *Search* 26:47–51.
18. Brandon, K, Redford, K, Sanderson, S (1998) *Parks in Peril*: *People*, *Politics and Protected Areas* (Island Press, Washington DC).
19. Cornell, J (2000) Assessing the role of parks for protecting forest resources using GIS and spatial modeling. In *Quantifying Sustainable Development*: *The Future of Tropical Economies*, ed Hall, CAS (Academic, San Diego), pp 543–560.
20. Scott J, *et al.* (2001) Nature reserves: Do they capture the full range of America's biological diversity? *Ecological Applications* 11:999–1007.
21. Chomitz K, Gray D (1996) Roads, land use and deforestation: A spatial model applied to Belize. *World Bank Economic Review* 10:487–512.
22. Cropper M, Puri J, Griffiths C (2001) Predicting the location of deforestation: The role of roads and protected areas in North Thailand. *Land Economics* 77:172–186.
23. Deininger K, Minten B (2002) Determinants of deforestation and the economics of protection: An application to Mexico. *American Journal of Agricultural Economics* 84:943–960.
24. Mas J (2005) Assessing protected area effectiveness using surrounding (buffer) areas environmentally similar to the target area. *Environmental Monitoring and Assessment* 105:69–80.
25. de Camino R, Segura O, Arias L, Perez I (2000) *Costa Rica Forest Strategy and the Evolution of Land Use* (The World Bank, Washington DC).
26. Kaimowitz D, Angelsen A (1998) *Economic Models of Deforestation*: *A Review* (Center for International Forestry Research, Bogor, Indonesia).
27. Sanchez-Azofeifa G, Harriss R (2001) Deforestation in Costa Rica: A quantitative analysis using remote sensing imagery. *Biotropica* 33:378–384.
28. Pfaff A (1999) What drives deforestation in the Brazilian Amazon? *Journal of Environmental Economics and Management* 37:26–43.
29. Imbens G (2004) Nonparametric estimation of average treatment effects under exogeneity: A review. *Review of Economics and Statistics* 86:4–29.
30. Ho D, Imai K, King G, Stuart E (2007) Matching as nonparametric preprocessing for reducing model dependence in parametric causal inference. *Political Analysis* 15:199–236.
31. Rosenbaum, P (2002) *Observational Studies* (Springer-Verlag, New York) 2nd Ed.
32. Naidoo R, *et al.* (2006) Integrating economic costs into conservation planning. *Trends Ecol Evol* 21:681–687.
33. Soares-Filho B, *et al.* (2006) Modeling conservation in the Amazon basin. *Nature* 440:520–523.
34. Abadie A, Imbens G (2006) *On the Failure of the Bootstrap for Matching Estimators* (National Bureau of Economic Research Technical Working Paper 325, Cambridge, MA).
35. Abadie A, Imbens G (2006) Large sample properties of matching estimators for average treatment effects. *Econometrica* 74:235–267.
36. Diamond A, Sekhon J (2006) *Genetic matching for estimating causal effects*: *A general multivariate matching method for achieving balance in observational studies* (Institute of Governmental Studies Working Paper 2006-35, Berkeley, CA).