What are the social impacts of land use restrictions on local communities? Empirical evidence from Costa Rica

Kwaw S. Andam[1]
Paul J. Ferraro[2]
Margaret B. Holland[3]

Contributed Paper prepared for presentation at the International Association of Agricultural Economists' 2009 Conference, Beijing, China, August 16-22, 2009.

[1] International Food Policy Research Institute, 2033 K Street, NW, Washington, DC 20006-1002, USA; Phone: +251.11.617.2507; Fax: +251.11.646.2927; Email: k.andam@cgiar.org (Presenting Author)

[2] Department of Economics, Andrew Young School of Policy Studies, Georgia State University, PO Box 3992, Atlanta, GA 30302-3992, USA; Phone: +1.404.413.0201; +1.404.413.0205 (fax); Email: pferraro@gsu.edu

[3] Nelson Institute for Environmental Studies University of Wisconsin-Madison 550 N. Park Street, Room 122 Madison, WI 53706; Phone: +829.962.2659; Email: mvbuck@wisc.edu

Acknowledgements: This work was supported by the Resources for the Future's Fisher Dissertation Award and the Global Environment Facility (Andam), and the World Wildlife Fund's Kathryn Fuller Science for Nature Fund (Ferraro). The authors are grateful to Kate Sims and David Spielman for their helpful comments and acknowledge research assistance from Merlin Hanauer.

Abstract

Global efforts to reduce deforestation rely heavily on protected areas and land use restrictions. The effect of these restrictions on local communities is currently the subject of heated debate among conservation and development experts. Measuring the social impacts of protected areas is difficult because the effects cannot be isolated from other factors, given the nonrandom placement of protection. We address this problem by applying a quasi-experimental approach to establish the counterfactual ("what would have been the socioeconomic outcome if a protected area had not been established?"). We use matching methods to measure the impacts of pre-1980 protected areas in Costa Rica on socioeconomic outcomes in 2000. In 2000, neighboring communities near protected areas were substantially poorer than average. However, after controlling for pre-protection characteristics associated with both protection and economic growth, the results indicate that poverty declined as a result of protection. Although the statistical significance of this decline is moderately sensitive to potential hidden bias, the results emphatically do not support a hypothesis that ecosystem protection, on average, exacerbates poverty. In contrast, conventional empirical methods implied erroneously that protection had negative social impacts, suggesting that failure to control for confounding factors or baselines can lead to substantially inaccurate estimates.

Keywords: forest conservation, social impacts, quasi-experimental methods

JEL Codes: C14, O13, Q23

Introduction

Global efforts to reduce tropical deforestation rely heavily on the establishment of protected areas (MA, 2005). Protected areas are the most widely used conservation tool in developing countries, with more than $6.5 billion in annual global expenditures (Emerton et al., 2006). One of the most controversial debates in conservation policy centers on the effect of these protected areas on local communities. This debate is particularly contentious with regard to developing countries, where protected area networks have rapidly expanded since the 1970s and where alleviating widespread rural poverty is a paramount but sometimes conflicting concern. The debate has intensified recently as policymakers seek to design interventions to reduce emissions from deforestation and degradation (REDD) to mitigate climate change.

Although most studies of protected areas focus on the environmental impacts of protection, conservation experts and policymakers now recognize that socio-economic impacts must also be considered (CBD, 2002; WPC, 2004; Adams et al. 2004; Balmford et al., 2005). Protected areas may have negative impacts on neighboring communities by restricting land use, or they may have positive impacts by creating economic opportunities for local communities (e.g. ecotourism). A credible study of the net effects of protected areas on the welfare of neighboring communities would include the following four elements: 1) objectively measurable indicators of human welfare at an appropriate scale of analysis (e.g., households, census tracts, villages, or regions); 2) observations of these indicators before and after the establishment of the protected area, or if no baseline observations are available, some other control for the initial state and trend of the indicators; 3) observations of these indicators from both treated units (i.e., areas known to be potentially affected by protected areas) and control units (i.e., areas similar to treated units in economic potential but known to be unaffected, or less affected, by protected areas); and 4) observations of baseline characteristics that affect both where protected areas are located *and* social welfare. Such baseline characteristics can bias the estimate of the protected areas' impacts; for example, if protected areas are located on less productive lands, a simple comparison of economic growth between communities near and far from protected areas may erroneously suggest protection is detrimental to growth when, in fact, growth differences arise from inherent land productivity differences. To date, no study with all of these elements has been published.

Most studies that attempt to estimate the net impact of protected areas (see Ferraro, 2002) focus on a single protected area and are based on attitudinal surveys, case study narratives, ex ante predictions based on historical use patterns and author assumptions, or ex post analyses that often prove little more than that rural people near protected areas are poor (Scherl et al., 2004; Agrawal and Redford, 2006; Wilkie et al., 2006). In general, these studies do not directly measure the impact of protected areas on poverty or social welfare, nor do they use data from before and after a protected area has been established or allow for sufficient time after establishment to observe an effect (Coad et al., 2008). Furthermore, with the exception of two county-level regional analyses in the United States (Duffy-Denno, 1998; Lewis et al., 2002),[4] previous analyses are unable to isolate the effects of protected areas from confounding factors that co-vary with protected area establishment.

We conduct a rigorous, controlled study to estimate the social impact of protected areas Costa Rica, where more than 1 million hectares of land had been protected by 2000. We address the question, "How different would socio-economic outcomes have been in neighboring

[4] These two studies find no effect of protected areas on wage or employment indicators, but they also lack some data on pre-establishment conditions.

communities in the absence of these protected areas?" by combining quantitative indicators of community welfare, pre-protection and post-protection data, and matching methods that allow us to select control communities that are observationally similar to communities near protected areas.

Data

Data Sources. We use socioeconomic data from the population and housing census conducted by the Instituto Nacional de Estadistica y Censos (INEC) in 1973 and 2000; GIS data layers for census segments for 1973 and 2000 were digitized and provided by the Cartography Department at INEC; GIS data layers for forest cover, protected areas, and the locations of major cities provided by the Earth Observation Systems Laboratory, University of Alberta, Canada; GIS layers of land use capacity from the Instituto Tecnologico de Costa Rica (ITCR, 2004); and GIS layers for roads digitized from hard copy maps for 1969 (map source: Instituto Geográfico Nacional (IGN) of the Ministerio Obras Publicas y Transporte (MOPT) of Costa Rica).

Unit of observation. The unit of analysis is the census segment (*segmento censal*), which is the smallest level of aggregation for which we have comparable census data in 1973 and 2000. Each census segment represents between forty to sixty households, depending on its location in a rural vs. urban area. Due to the increase in population and number of households between each census, the relative size and number of segments shifted considerably between both census years of interest. Therefore, we faced the challenge of reconciling segment geography from the two periods. We overcome this limitation by using a simple areal interpolation technique, known as areal weighting (Reibel, 2007). The main assumption, and criticism, of areal interpolation is that it assumes homogeneity within the unit of analysis. While other techniques, such as dasymetric mapping, attempt to address the heterogeneity inherent in administrative units such as those in this analysis, the ancillary data required were not available. Areal interpolation was therefore the simplest and clearest choice of methods to address this problem. Using the TwoThemes extension developed for ArcView®, we aggregated census data from 2000 to the segments from 1973, and disaggregated census data from 1973 to the segments from 2000. As noted above, areal interpolation assumes spatial homogeneity within the unit of analysis. We believe it is more accurate to assume spatial homogeneity for the census segments of 2000 than to do so for the census segments of 1973, since the 2000 segments are smaller in size. Thus we select as our main dataset the census segments of 1973, with all census data from 2000 aggregated to the census segment georgraphy of these 1973 segments. This dataset comprises 4,691 census segments, after excluding two segments for which there were no census data in 1973[5]. However, we test the robustness of our estimates by repeating our analyses with the 1973 census data disaggregated to match the 2000 census segments. Finally, we overlay the segments GIS data layers with the layers for biophysical and infrastructure variables to obtain our full dataset.

Outcomes. We analyze the effects of protection on three socioeconomic indicators: a poverty index and the population density (number of persons per square kilometer) both measured in 2000, and the population growth 1973-2000 (change in population as a fraction of the baseline 1973 population). Poverty is multidimensional and varies according to sociopolitical and biophysical realities. The index constructed for this analysis is not intended to capture the full meaning of who is poor and where they live in the study region, and the measure is limited to

[5] We have anecdotal information that these two segments were not surveyed in 1973 because there are no residents within those segments at that time.

variables available in the national population and housing censuses for both years. The poverty index for this analysis was obtained by using principal components analysis (PCA), a type of factor analysis. The first principal component, that which captures the most variance among the combination of factors, is used to construct the index: factor scores from the first component are used as weights for each variable, which are then combined into a single index score. Cavatassi et al. (2004) used PCA in developing a time-variant poverty index for Costa Rica at the 3rd administrative, or district, level. They selected PCA because it can focus solely on census data, is flexible for constructing an index of change over time, is relatively inexpensive and easy to calculate once the data are compiled, and has been used in several countries with results comparable to those of consumption-based welfare indicators (Filmer and Pritchett, 1998; Skoufias et al., 2001).

The 17 variables included in the population index are described in Table 2. As noted by Cavatassi et al., these variables have been found in other studies to be associated with poverty in Costa Rica. To make the indexes comparable over time, we follow Cavatassi et al. by pooling the data for 1973 and 2000 before applying the PCA to generate weights for estimating the poverty index.

Treatment. The treatment is defined as “more than 10% of segment protected before 1980”. The 10% threshold has some policy relevance: it reflects the call by the 4th World Congress on National Parks and Protected Areas to protect 10% of each of the world’s major biomes by 2000, as well as the call by the Conference of Parties to the Convention on Biological Diversity to conserve 10% of each of the world’s ecoregions. We focus on terrestrial protected areas established before 1980 in order to allow 20 years or more for the impacts of the protected area to be experienced by local residents. Our treatment group comprises 230 segments. To further reduce bias, we trim the sample by excluding all potential control segments that received any protection before or after 1980 (a total of 376 segments). Thus, we also exclude control segments with any part of their segments protected before 1980 (there were 316 such control segments with protection below the 10% threshold level before 1980, of which 21 also received additional protection after 1980). The number of available controls (unprotected segments) therefore comprises 4085 segments. We show that our results are not sensitive to the 10% threshold for defining treatment, the 1980 cut-off date for protection to affect outcomes, or the exclusion of controls with some protection.

The pre-1980 protected areas which we overlap in the GIS dataset to select the treated segments comprise Biological Reserves (Cordillera Volcanica Central, Golfo Dulce, Grecia, Los Santos, Rio Macho, Taboga), National Monuments (Guayabo), Forest Reserves (Pacuare-Matina, Zona de Emergencia Volcan Arenal), National Parks (Barra Honda, Braulio Carrillo, Cahuita, Chirripo, Corcovado, Juan Castro Blanco, Manuel Antonio, Palo Verde, Rincon De La Vieja, Santa Rosa, Tortuguero, Volcan Iraza, Volcan Poas, Volcan Tenorio, Volcan Turrialba), Protected Zones (Arenal-Monterverde, Caraigres, Cerro Atenas, Cerros de Escazu, Ceros de la Carpintera, El Rodeo, Miravalles, Rio Grande, Tenorio) and Wildlife Refuges (Corredor Fronterizo). Five protected areas established before 1980 are not represented in our sample because they are islands that overlap with neither the segment layers nor the 1960 forest cover layer.

Covariates. We control for covariates that could potentially confound the estimation of the effects of protection. See Table 1 for summary statistics. We confirm the narrative and empirical evidence that these variables also affect the designation of protected areas by modeling the selection process directly using our data and a probit model (regressing a dummy variable for treatment on the covariates).

We control for the following covariates in the matching analysis:

Proportion of segment under forest cover in 1960 area: This is the earliest measure of forest cover prior to the establishment of protected areas. Forest area is likely to be highly correlated with the likelihood of protected area location. It is also likely to affect socioeconomic outcomes. For example, segments with more forest cover may offer more opportunities for exploiting forest products.

"Road-less volume": Road-less volume is a metric developed by Watts et al. (2007) to measure accessibility to transportation infrastructure. Road-less volume provides a better way of capturing this effect than measures such as road density or the distance from each segment to the nearest road, because such measures only reflect accessibility at the larger segment scale. In contrast, road-less volume measures the accessibility of each plot of land and aggregates this measure to the segment level. Furthermore, road-less volume simultaneously measures the extent to which roads have penetrated a segment as well as the extent to which roads have penetrated adjacent segments. First, we calculate the road-less volume for each square of length 100m across the country (road-less volume = distance from center of the square to nearest major road in 1969 * area of the square). We then add the road-less volumes for all squares within a segment to obtain the total road-less volume for the segment. Road-less volume may have opposing effects on the likelihood of protection. On the one hand, remote lands may be considered less threatened by deforestation and therefore may be more likely candidates for protection. Thus, segments with larger road-less volume may be more likely to be protected. On the other hand,

protected areas that are created for ecotourism may be located near roads to make those parks more accessible, implying that segments with smaller road-less volumes would be protected. Road-less volume also affects socioeconomic outcomes by affecting access to forest, agricultural lands, and markets.

Land use capacity: We use Costa Rica's *land use capacity classes*, which are determined by slope, soil characteristics, life zones (Holdridge 1967), risk of flooding, dry period, fog, and wind influences. The classes are defined in Table 1. We define classes I-III as "high productivity land," class IV as "medium productivity land," classes V-VII as "medium-low productivity land," and classes VIII and IX as "low productivity land" (the last is the omitted category).
Distance to nearest major city: This variable is a measure of proximity to agricultural markets. Following Pfaff and Sanchez (2004), we identify three major cities: Limon, Puntarenas, and San Jose. This variable measures the distance from the centroid of the segment to the closest of these cities..
Baseline poverty index in 1973: See description of the poverty index variable under the section on Outcomes above. Note that we control for this variable only in the post-matching regression for effects of protection on poverty index in 2000.

Baseline population density in 1973: The number of persons per square kilometer (in the post-matching regression for effects of protection on population density in 2000 only).

As a robustness check on the sensitivity of the estimates to the selected covariates, in the second type of post-matching regressions, we modify this set of covariates described above. Instead of controlling for just one baseline socioeconomic indicator we control for both baseline indicators in each regression (that is, instead of controlling for just baseline poverty index in measuring the effect of protection on the poverty index, we control for baseline population density as well, and similarly for the population density regression). And instead of controlling for the proportion of segment under forest in 1960 we control for the area under forest in 1960, and instead of the proportion of the segment under each land use class we control for the area of the segment under the respective land use classes. Instead of controlling for the road-less volume, we control for the *distance to nearest road* (the distance from the centroid of the segment to a road in 1969). We also control for the segment area (measured in square kilometers).

Matching methods

The key challenge in this study is to control for biophysical and socio-economic covariates that affect both changes in social welfare *and* the location of protected areas. We establish this control through matching methods, which are increasingly used as a way to test cause-effect relationships using non-experimental data (Imbens, 2004). Matching works by identifying, *ex post*, a comparison group that is "very similar" to the treatment group with only one key difference: the comparison group did not participate in the program (Rubin, 1980; Rosenbaum and Rubin, 1983; Imbens, 2004). Matching identifies a control group with similar covariate distributions to the treated group (called covariate balancing), thereby removing observable sources of bias. If the researcher can select observable characteristics so that any two land units with the same value for these characteristics will display homogenous responses to the treatment (i.e., protection is independent of social impacts for similar land units), then the treatment effect can be measured without bias.[6]

---

[6] Mathematically, the key assumption is: $E[Y(0)\,|\,X,T=1]=E[Y(0)\,|\,X,T=0]=E[Y(0)\,|\,X]$ and $E[Y(1)\,|\,X,T=1]=E[Y(1)\,|\,X,T=0]=E[Y(1)\,|\,X]$, where $Y_i(1)$ is the socioeconomic outcome when census segment i

We tried a variety of matching methods and selected the one that gave us the best covariate balance (Ho et al., 2007): covariate matching that uses the Mahalanobis distance metric (with and without calipers[7]) to identify matches that are similar to the protected segments. Matching was done in R (Sekhon, 2007).

Results

The analysis focuses on the effect of protection in 230 segments with more than 10% of the segment protected before 1980. In 2000, the mean poverty index in protected segments is almost five points higher (4.92) than in unprotected segments, which is a substantial difference of slightly greater than one standard deviation (4.82).

This difference, however, does not necessarily reflect a causal relationship between protection and poverty. Segments surrounding protected areas established before 1980 were, at baseline, among the poorest segments. The odds of a segment having more than 10% of its area protected before 1980 are more than 20 times higher for segments with above-average baseline poverty. This baseline difference is important because poverty at baseline and in 2000 are highly correlated (0.83). Protected areas are clearly not randomly assigned across the landscape. An

---

is protected, $Y_i(0)$ is the outcome when segment i is unprotected, T is treatment (T=1 if protected), and X is the set of pretreatment characteristics on which segments are matched. This assumption, called the conditional independence assumption (CIA), implies that participation in the project depends solely on a set of observable characteristics (X), and that we can observe the variables which simultaneously affect both participation and outcomes. For identification purposes, we also need one other assumption: $c < P(T = 1 | X = x) < 1 - c$ for c > 0. In other words, if all segments with a given vector of covariates were protected, there would be no observations on similar unprotected segments, and therefore, no suitable comparison group.

[7] Calipers define a tolerance level for judging the quality of the matches: if a treated segment does not have a match within the caliper (i.e., available controls are not good matches), it is eliminated from the sample. We set a caliper of one standard deviation of each control variable.

analysis must control for confounding baseline characteristics that affect both changes in poverty and where protected areas are established.

The third column of Table 3 presents mean values for covariates among protected segments and the fourth column presents mean covariate values among unprotected segments, both before and after matching. The first row reports data that show the pre-existing differences. Without matching, the 1973 poverty index values are much higher among protected segments than unprotected segments (14.67 vs. 5.17). The inherent productivity of protected segments (rows 3-5) is also much worse among the protected segments: whereas more than 75% of unprotected segments comprise high or medium productivity lands, just over 25% of protected segments comprise such lands. Protected segments also have much greater percent of forest cover in 1960, greater roadless volume, and are much farther from major cities. All of these characteristics have negative impacts on a segment's potential for economic growth.

The fifth and sixth columns of Table 3 present two measures of the differences in the covariate distributions between protected and unprotected segments: the difference in means and the average distance between the two empirical quantile functions (values greater than 0 indicate deviations between the groups in some part of the empirical distribution). If matching is effective, both of these measures should move dramatically towards zero (Ho et al.). The measures in the fifth and sixth columns indeed move dramatically towards zero after matching (we use the matching method that yields the best covariate balance). In Table 4, we present balance metrics for the matching with calipers, which are similar to the ones for Table 3.

The first column in the first and second rows of Table 5 present the impact estimates from the matching approach. It reports mean differences in 2000 poverty index values between protected and matched unprotected segments. A negative sign thus indicates that protection alleviates poverty. The first row (matching without calipers) indicates that the mean poverty index value among protected segments was 1.95 points lower because of the presence of protected areas ($p<0.01$). Recall that this result is opposite in sign from the estimate based simply on mean differences between unmatched protected and unprotected segments, which is presented in the first row (*Conventional Estimates*). The second row presents an estimate based on matching that uses calipers to improve covariate balance (balancing results in Table 4). Calipers define a tolerance level for judging the quality of the matches: if a treated segment does not have a match within the caliper (i.e., available controls are not good matches), it is eliminated from the sample. Twenty-two protected segments are eliminated, but the estimated impact on poverty is similar to the estimate without calipers.

Although matching substantially improves the covariate balance between protected and unprotected segments, some imbalance remains: protected segments have lower proportion of lands in the most productive land use classes and have less accessibility to roads compared to their matched counterparts. A post-matching, linear regression that adjusts for small remaining imbalances in the matched sample yields similar estimates to those in Table 5. To test model dependence (Ho et al.), we use a variety of post-matching regression specifications with an extended covariate set that includes, among other variables, alternative measures of the underlying profitability of land conversion. We find the poverty impact estimates differ little from those in Table 5 (less than one-half point).

Thus although a simple comparison of mean differences in post-protection poverty suggests that protection exacerbated poverty, there is no evidence of such an impact conditional on baseline conditions. In fact, the evidence suggests the opposite: protection helped to alleviate poverty. The estimated effect sizes from matching with and without calipers are -0.29 and -0.33, respectively (calculated by dividing the average treatment effect on the treated estimate by standard deviation of the matched control segments).

Other Robustness Checks. The conclusions are also robust to changes in the sample composition, the matching specifications, and the scale at which the analysis is conducted. We confirm that the estimated treatment effects are robust to these variations in the analysis. In all these robustness checks, the matching estimates of the effect of protection on the poverty index always lie between -3.047 and -0.811 and are all significantly different from zero ($p<0.01$). Moreover, while the matching estimates always have a negative sign suggesting that protection alleviated poverty, their corresponding estimates obtained using the simple comparison of the difference in means always suggest dramatically different effects of protection: those estimates are always positive and significantly different from zero ($p<0.01$), lying between 4.376 and 5.969, suggesting that protection exacerbates poverty. For the population density, the estimates always lie between -0.711 and 0.067 and are never significantly different from zero ($p>0.1$), with the exception of the matching estimates with the 20% threshold which are significantly different from zero (but only at the 10% level of significance). Note that while these estimates are significantly different from zero, the estimate is not economically significant. That is, the estimate indicates that there are only 10 more people in protected segments compared to the unprotected segments. Again, while these matching estimates suggest little or no impact of protection on the population density, the simple comparison of means suggests that protection reduced population density, with those estimates lying between -4.724 and -3.986 and always significantly different from zero ($p<0.01$).

The robustness checks are described briefly below:

Robustness Check #1: Changing the scale of the unit of observation: Instead of using the aggregated segments in 1973, we use the disaggregated segments surveyed in 2000. The difference in these two scales is described in the Data section;

Robustness Check #2: Including protected areas established between 1980-2000: To ensure that there is no substantial bias from using the 1980 threshold for protection, we estimate the treatment effects of protected areas established before 2000;

Robustness Check #3: Varying the threshold of protection: We test the robustness of the 10% threshold for protection by varying the threshold at 20%, 50%, and 75%. Estimates using the difference in means test are 4.841, 4.587, and 5.847 for the poverty index, and 4.374, 4.292, and 4.381 for the population density for thresholds 20%, 50%, and 75% respectively. All of these estimates are significantly different from zero ($p<0.01$);

Robustness Check #4: Including control segments with protection before or after 1980. We estimate the treatment effects without excluding the 376 control segments that were originally excluded from the analysis because they had some protection below the 10% threshold before 1980 and/or were protected after 1980.

Discussion

We find no evidence that Costa Rica's protected areas have had net negative social impacts. In fact, we find the opposite: Protected areas have had a net positive effect on indicators of local social welfare. In contrast, conventional methods that fail to account for the non-random assignment of protected areas suggest the opposite relationship: Protection has negative impacts on social welfare.

How do protected areas lead to beneficial socioeconomic outcomes? First, protection may lead to the growth of an ecotourism industry that creates better economic opportunities for neighboring communities. Second, since tourism is Costa Rica's main source of foreign exchange, the establishment of a protected area may have led to an increase in government provision of infrastructure services to promote ecotourism. Third, some conservation programs have sought to reduce the deforestation pressure on protected areas by investing in communities living in or near protected areas.[8] Our findings suggest that such interventions may have improved the livelihoods of local communities.

The absence of a net negative effect is remarkable given that recent studies show that protection in Costa Rica has indeed resulted in reduced deforestation (Andam et al. 2008). Thus there have been opportunity costs incurred from protection, but our findings suggest that the loss of access to forest resources is more than offset by other economic activities, such as tourism, or infrastructure investments associated with protected areas. Whether our results would hold for other nations is an open question, since Costa Rica may be very different from those other countries in terms of its institutions and protected area management.[9] This type of analysis should be repeated in other nations and on other forest governance regimes (e.g., indigenous reserves).

[8] For example, a project called the Amistad Conservation and Development Initiative (AMISCONDE), worked with local farmers around protected areas to improve agricultural practices from 1991-1997. This project was implemented by Conservation International and various partners.

[9] Note that our results do not call into question the widely held belief that many of the benefits of biodiversity protection are enjoyed by residents far from protected areas, while many of the costs are incurred by local people (Balmford and Whitten, 2003), and thus transfers from wealthy to developing countries are needed to achieve conservation goals.

Study limitations

First, although we use a variety of indicators that are correlated with local well-being, they do not capture all aspects of well-being (e.g., hard-to-measure aspects such as "feeling in control of one's life" or "ability to maintain cultural traditions"). Second, the data are only available at the census segment level and therefore we can only observe average outcomes at this aggregated level. Adverse effects on subgroups within the community may not be observable at this level. For example, if protected areas cause shifts in economic activities from agriculture to ecotourism, farmers may be adversely affected while the tourism industry experiences growth. Theoretical models indicate that protection leads to higher land rents and lower agricultural wages, causing changes in income distribution (Robalino, 2007). Furthermore, after protected areas are established, some displaced residents may relocate, and the effects of protection on these people cannot be detected without individual-level, panel data collected pre- and post-protection.

Although we cannot fully observe the distributional impacts of protection, we conduct some analyses to explore this issue. If we assume that migration tends to be local (adversely affected residents move to nearby unprotected segments), then we would expect these nearby segments to have worse socioeconomic outcomes, on average, than segments that are farther away from protected areas. However, we find no significant differences when we test for differences between protected and unprotected segments in terms of population density in 2000 or population growth 1973-2000 (Table 5). While this test does not completely rule out the possibility of adverse effects from protection on subgroups of the affected communities, it does suggest that our main findings (that protection had positive impacts in protected segments) cannot be explained by a significant local migration out of protected segments to nearby unprotected areas.

Conclusion

We apply a quasi-experimental approach to provide rigorous estimates of the social impacts of protected areas in Costa Rica. We address the question "what is the effect of this protected area on economic outcomes within neighboring communities?" by using matching methods to identify suitable comparisons for affected communities. We find that Costa Rican protected areas have had a net positive impact on socioeconomic outcomes within neighboring communities.

Our research approach represents a major advance in estimating the social impacts of protected areas and makes an important contribution to strengthening the evidence base for conservation policy (Ferraro and Pattanayak, 2006; Sutherland et al. 2004). Our study thus highlights the need for cooperation between groups collecting spatially explicit poverty data, protected area data, and land-use -land-cover data[10]. Future studies can use similar methods to explore how impacts vary conditional on observable covariates, for example, whether impacts vary with the degree of baseline poverty.

[10] For example, the UNEP-WCMC Vision 2020 project, which seeks to expand the World Database on Protected Areas (WDPA) to cover socio-economic issues as well as develop indicators related to protected areas and social impacts.

References

Adams, W. M., R. Aveling, D. Brockington, B. Dickson, J. Elliott, J. Hutton, D. Roe, B. Vira, and W. Wolmer (2004) Biodiversity Conservation and the Eradication of Poverty, *Science*, Vol. 306, pp. 1146–1149.

Agrawal, A. and K. Redford (2006) Poverty, Development, and Biodiversity Conservation: Shooting in the Dark? Wildlife Conservation Society, New York (WCS Working Paper 26).

Andam, K.S, P.J. Ferraro, A.S.P. Pfaff, A. Sanchez-Azofiefa, and J. Robalino (2008) Measuring the Effectiveness of Protected Area Networks in Reducing Deforestation, *Proceedings of the National Academy of Sciences of the United States of America* 105 (42): 16089-16094.

Balmford, A., L. Bennun, B. ten Brink, D. Cooper, I.M. Cote, P. Crane, A. Dobson, N. Dudley, I. Dutton, R.E. Green, R. Gregory, J. Harrison, E.T. Kennedy, C. Kremen, N. Leader-Williams, T. Lovejoy, G. Mace, R. May, P. Mayaux, J. Phillips, K. Redford, T.H. Ricketts, J.P. Rodriguez, M. Sanjayan, P. Schei, A. van Jaarsveld, and B. A. Walther (2005) The Convention on Biological Diversity's 2010 target. *Science* 307:212-213.

Balmford, A. and T. Whitten (2003) Who should pay for tropical conservation, and how could the costs be met? *Oryx* **37**: 238-250.

Cavatassi, R., B. Davis, and L. Lipper (2004) *Estimating poverty over time and space: Construction of a time-variant poverty index for Costa Rica* (No. ESA Working Paper No. 04-21): The Food and Agriculture Organization, Agricultural and Development Economics Division.

CBD (2002) The Convention on Biological Diversity [Electronic version] Retrieved September 2008 from http://www.cbd.int/convention/convention.shtml

Coad, L., A. Campbell, L. Miles, and K. Humphries (2008) *The Costs and Benefits of Protected Areas for Local Livelihoods: a review of the current literature*. Working Paper. UNEP World Conservation Monitoring Centre, Cambridge, U.K.

Cohen, J. (1988) Statistical power analysis for the behavioral sciences (2nd ed.). Hillsdale, NJ: Lawrence Earlbaum Associates.

Duffy-Deno, K. T. (1998) The effect of federal wilderness on county growth in the intermountain western United States. *Journal of Regional Science* 38(1): 109-36.

Emerton, L., J. Bishop, and L. Thomas (2006) Sustainable Financing of Protected Areas: A global review of challenges and options. IUCN, Gland, Switzerland and Cambridge, UK.

Ferraro, P. J. (2002) "The Local Costs of Establishing Protected Areas in Low Income Nations: Ranomafana National Park, Madagascar." *Ecological Economics* 43: 261-75.

Ferraro, P. J. and S. K. Pattanayak (2006). Money for Nothing? A Call for Empirical Evaluation of Biodiversity Conservation Investments. *PLoS Biol* 4(4).

Filmer D. and L. Pritchett (1998) *Estimating wealth effects without expenditure data -- or tears: An application to educational enrolments in States of India,* World Bank Policy Research Working Paper No. 1994.

Ho, D., K. Imai, G. King, and E. Stuart (2007) Matching as nonparametric preprocessing for reducing model dependence in parametric causal inference. *Political Analysis* 15:199-236
ITCR (Cartographer) (2004) *Atlas Digital de Costa Rica 2004*

Holdridge, L. (1967). *Lifezone Ecology* San Jose, Costa Rica: Tropical Science Center.

Imbens, G. (2004) Nonparametric estimation of average treatment effects under exogeneity: a review. *Review of Economics and Statistics* 86:4-29.
Kaimowitz, D. and A. Angelsen (1998). Economic models of deforestation: A review. Bogor, Indonesia: Center for International Forestry Research.

Lewis, D. J., G. L. Hunt, and A. J. Plantinga (2002). Public Conservation Land and Employment Growth in the Northern Forest Region. *Land Economics* 78(2): 245-259.

MA – Millennium Ecosystem Assessment (2005) Ecosystems and Human Well-being: Policy Responses (Vol. 3): Island Press.

Pfaff, A. and A. Sanchez (2004). Deforestation pressure and biological reserve planning: A conceptual approach & an illustrative application for Costa Rica *Resource & Energy Economics 26*, 237-254.

Reibel, M. (2007) Geographic information systems and spatial data processing in demography: a review. *Population Research and Policy Review* 26:601–618.

Robalino, J. (2007) Land conservation policies and income distribution: who bears the burden of our environmental efforts? *Environment and Development Economics 12*(4).

Rosenbaum, P.R. and D.B. Rubin (1983) The central role of the propensity score in observational studies for causal effects. *Biometrika* 70:41-55.

Rubin, D.B. (1980) Bias reduction using Mahalanobis-metric matching. *Biometrics* 36:293-298.

Scherl, L. M., A. Wilson, R. Wild, J. Blockhus (2004) Can Protected Areas Contribute to Poverty Reduction? Opportunities and Limitations. IUCN, Gland, Switzerland.

Sekhon, J.S. (2007) Multivariate and propensity score matching software with automated balance optimization: the matching package for R. *Journal of Statistical Software*.

Skoufias, E., Davis, B. & de la Vega, S. 2001. Targeting the poor in Mexico: an evaluation of the selection of households into PROGRESA. *World Development*, 29(10): 1 769–1 784.

Smith, J.A. and P.E. Todd (2005) Does Matching Overcome LaLonde's Critique of Nonexperimental Estimators? *Journal of Econometrics* 125:305-353.

Sutherland, W.J., A.S. Pullin, P.M. Dolman, and T.M. Knight (2004) The need for evidence-based conservation. *Trends in Ecology and Evolution* **19**, 305-308.

Watts, R. D., R. W. Compton, J. H. McCammon, C. L. Rich, S. M. Wright, T. Owens (2007) Roadless space of the conterminous United States. *Science,* 4 (5825), 736-738.

Wilkie, D. S., G.A. Morelli, J. Demmer, M. Starkey, P. Telfer and M. Steil (2006) Parks and people: Assessing the human welfare effects of establishing protected areas for biodiversity conservation. *Conservation Biology* 20: 247-249.

WPC – World Parks Congress (2004) The Durban Action Plan.   Retrieved November 2005, from www.iucn.org/themes/wcpa/wpc2003/

Table 1. Descriptive Statistics

		Aggregated data (N=4691)			*Disaggregated data (N=17239)*		
Name	**Description**	**Mean**	**Stand-arid deviation**	**Min Max**	**Mean**	**Stand-arid deviation**	**Min Max**
Proportion of segment under forest in 1960	Total forest area in the segment in 1960 divided by total area of the segment	0.160	0.284	0.000 1.000	0.160	0.316	0.000 1.000
High Productivity Land (proportion)	Percent of segment area under the land classes I, II, and III, measured in km^2 Class I: Agricultural Production – annual crops; Class II: Suitable for agricultural production requiring special land and crop management practices such as water conservation, fertilization, irrigation, etc.; Class III: Suitable for agricultural production requiring special land and crop management practices such as water conservation, fertilization, irrigation, etc.	0.288	0.404	0.000 1.000	0.334	0.443	0.000 1.000
Medium Productivity Land (proportion)	Percent of segment area under the land class IV, measured in km^2 Class IV: Moderately suitable for agricultural production; permanent or semi-permanent crops such as fruit trees, sugar cane, coffee, ornamental plants, etc.	0.421	0.454	0.000 1.000	0.399	0.465	0.000 1.000
Medium-low Productivity Land (proportion)	Percent of segment area under the land classes V, VI, and VII, measured in km^2 Class V: Strong limitations for agriculture; forestry or pastureland Class VI: Strong limiting factors on agricultural production; land is only suitable for forest plantations or natural forest management Class VII: Strong limiting factors on agricultural production; land is only suitable for forest plantations or natural forest management	0.209	0.354	0.000 1.000	0.198	0.370	0.000 1.000
Low Productivity Land (proportion)	Percent of segment area under the land classes VIII and IX, measured in km^2 Class VIII: Land is suitable only for watershed protection Class IX: Land is suitable only for	0.078	0.219	0.000 1.000	0.065	0.219	0.000 1.000

		Aggregated data (N=4691)			***Disaggregated data (N=17239)***		
Name	**Description**	**Mean**	**Stand-arid deviation**	**Min Max**	**Mean**	**Stand-arid deviation**	**Min Max**
	protection						
Protected before 1980 (proportion)	Proportion of the segment area that was protected before 1980	0.023	0.113	0.000 1.000	0.015	0.100	0.000 1.000
Protected after 1980 (proportion)	Proportion of the segment area that was protected after 1980	0.017	0.096	0.000 0.999	0.010	0.081	0.000 1.000
Segment area (km^2)	Total land area of segment (in square km)	10713.088	32455.426	4.660 862683.313	2922.112	12576.472	1.154 736120.125
Roadless Volume (km^3)	The sum of the product of area and distance to nearest road (1969) for every square of length 100m within the segment	162.224	911.764	0.000 33970.148	44.692	348.149	0.000 25504.395
Distance to city (km)	Distance from centroid of the segment to closest major city (Limon, Puntarenas, or San Jose), measured in km	38.276	39.529	0.029 214.906	37.031	37.778	0.044 203.263
Distance to road	Distance from centroid of the segment to nearest road (1969) measured in km	4.520	8.328	0.000 62.656	4.949	9.410	0.000 63.556
Poverty index in 1973	Multidimensional index of poverty derived from a linear combination of a set of key socioeconomic variables (see Outcome section for detailed description and Table 2 for list of variables)	6.462	9.556	-16.133 28.855	6.987	7.829	-16.097 26.942
Poverty index in 2000	Multidimensional index of poverty derived from a linear combination of a set of key socioeconomic variables (see Outcome section for detailed description and Table 2 for list of variables)	-6.469	5.193	-16.390 23.290	-7.007	5.466	-17.265 31.233
Population in 1973	Total number of residents in segment in 1973	398.904	133.912	28 1527	108.315	117.949	0.000 1893
Population in 2000	Total number of residents in segment in 2000	811.291	1255.309	1.000 26169	220.913	86.026	1.000 2318.000
Population Density in 1973 (Persons per km^2)	Total number of residents in segment in 1973 divided by segment area	4.630	8.169	0.000 64.721	2.069	4.769	0.000 62.735
Population Density in 2000 (Persons per	Total number of residents in segment in 2000 divided by segment area	3.957	5.720	0.000 44.072	5.674	8.466	0.000 131.625

		Aggregated data (N=4691)			Disaggregated data (N=17239)		
Name	**Description**	**Mean**	**Stand-arid deviation**	**Min Max**	**Mean**	**Stand-arid deviation**	**Min Max**
km²)							
Population growth 1973-2000	Difference in population between 1973 and 2000 divided by population in 1973	1.098	3.765	-0.981 85.838	26.999	234.329	-1 26499

Table 2. Variables used to Calculate Poverty Indexes for 1973 and 2000

Variable	Description
Male	Percentage of men in total population
No toilet	Percentage of dwellings without toilet
No hot water	Percentage of dwellings without access to hot water
Use coal or wood	Percentage of households that cook with charcoal or wood
Dirt floor	Percentage of dwellings with dirt floor
Dependency ratio	Dependency ratio (children under 15 and people over 65 divided by rest of population)
House in bad conditions	Percentage of dwellings in bad condition
No washing machine	Percentage of dwellings without washing machine
No electricity	Percentage of dwellings without electricity
No telephone	Percentage of dwellings without telephone
No refrigerator	Percentage of dwellings without refrigerator
Employed	Percentage of people who are employed compared with the economically active population
Illiterate	Percentage of illiterate population aged 10+
No water system	Percentage of dwellings without connection to private or public water system
No sewage	Percentage of dwellings without indoor plumbing
Crowding	Percentage of dwellings with 3 or more occupants per bedroom
Adult primary or no education	Percentage of adult population (18+) with educational attainment of primary level or no formal education

Table 3. Covariate Balance – Poverty Index Outcome, Matching without Calipers

Covariate	Sample	Mean Value Treated Segments	Mean Value Control Segments*	Diff in Mean Values	Mean eQQ Diff‡	Median eQQ Diff**	Max eQQ Diff**	Mean eCDF Diff^
Proportion of segment under forest in 1960	Unmatched	0.522	0.111	0.411	0.409	0.464	0.759	0.461
	Matched	0.522	0.478	0.044	0.044	0.039	0.139	0.044
High Productivity Land (proportion)	Unmatched	0.065	0.305	-0.240	0.241	0.000	0.896	0.221
	Matched	0.065	0.135	-0.070	0.070	0.016	0.578	0.129
Medium Productivity Land (proportion)	Unmatched	0.201	0.459	-0.258	0.260	0.094	0.824	0.225
	Matched	0.201	0.197	0.004	0.026	0.013	0.150	0.034
Medium-Low Productivity Land (proportion)¤	Unmatched	0.259	0.191	0.068	0.121	0.076	0.413	0.153
	Matched	0.259	0.253	0.006	0.034	0.018	0.129	0.049
Distance to City (km)	Unmatched	54.562	34.685	19.877	19.814	19.505	45.012	0.179
	Matched	54.562	55.181	-0.619	5.991	3.531	25.200	0.038
Roadless Volume (km^3)	Unmatched	872.500	59.607	812.893	789.510	118.560	15025	0.360
	Matched	872.500	669.840	202.660	220.890	39.611	15025	0.082
Poverty Index in 1973	Unmatched	14.670	5.169	9.501	9.521	9.508	19.258	0.301
	Matched	14.670	14.757	-0.087	1.939	2.062	4.289	0.079

¤ Low productivity land is the omitted category.
* Weighted means for matched controls.
‡ Mean raw eQQ is the mean difference in the empirical quantile-quantile plot of treatment and control groups on the scale in which the variable is measured.

Table 4. Covariate Balance – Poverty Index Outcome, Matching with Calipers

Covariate	Sample	Mean Value Treated Segments	Mean Value Control Segments*	Diff in Mean Values	Mean eQQ Diff‡	Median eQQ Diff**	Max eQQ Diff**	Mean eCDF Diff^
Proportion of segment under forest in 1960	Unmatched	0.509	0.136	0.373	0.372	0.379	0.863	0.380
	Matched	0.509	0.482	0.027	0.028	0.019	0.150	0.028
High Productivity Land (proportion)	Unmatched	0.052	0.348	0.296	0.296	0.000	0.999	0.281
	Matched	0.052	0.083	0.031	0.030	0.000	0.194	0.054
Medium Productivity Land (proportion)	Unmatched	0.162	0.415	0.253	0.253	0.000	0.925	0.231
	Matched	0.162	0.162	0.000	0.011	0.000	0.080	0.015
Medium-Low Productivity Land (proportion)¤	Unmatched	0.240	0.191	0.049	0.102	0.011	0.435	0.125
	Matched	0.240	0.226	0.014	0.026	0.001	0.150	0.034
Distance to City (km)	Unmatched	53.670	35.463	18.207	18.153	16.567	52.275	0.150
	Matched	53.670	54.514	-0.844	5.178	3.647	24.281	0.037
Roadless Volume (km^3)	Unmatched	396.790	22.831	373.959	368.750	46.291	18530	0.338
	Matched	396.790	265.250	131.540	132.820	12.829	18530	0.061
Poverty Index in 1973	Unmatched	13.243	6.517	6.726	6.752	6.229	17.013	0.264
	Matched	13.243	13.540	-0.297	0.530	0.497	1.963	0.025

¤ Low productivity land is the omitted category.
* Weighted means for matched controls.
‡ Mean raw eQQ is the mean difference in the empirical quantile-quantile plot of treatment and control groups on the scale in which the variable is measured.

Table 5. Estimated Impacts of Protected Areas on Poverty and Population in 2000

	1	2	3
Outcome	***Poverty Index^ (Costa Rica)***	***Population Density^ (Costa Rica)***	***Population Growth (Costa Rica)***
Difference in Means†	4.923***	-4.346***	-0.237
	(0.318)	(0.391)	(0.202)
Matching without calipers‡	-1.948***	-0.212	-0.667
	(0.547)	(0.136)	(0.566)
Matching with calipers‡¤	-2.011***	-0.219	0.307
[N treated dropped by calipers]	(0.486)	(0.135)	(0.240)
	[22]	[21]	[30]
N treated	230	230	230
(N available controls)	(4085)	(4085)	(4085)

^ Average treatment effect on the treated of more than 10% of the segment protected before 1980. Population density is measured in persons per square km (population density = total population / segment area in km)
‡ For Costa Rica, covariate matching on the Mahalanobis distance metric is used. Robust standard errors (Abadie-Imbens) are in parenthesis under estimate.
¤ For Costa Rica, calipers restrict matches to units within 1 standard deviation of each covariate.
† A t-test of the difference in means between treated and control segments.
***, **, * indicate significance at 1%, 5%, 10% respectively.