

Do Payments for Environmental Services Affect Forest Cover? A Farm-Level Evaluation from Costa Rica

Rodrigo A. Arriagada, Paul J. Ferraro, Erin O. Sills, Subhrendu K. Pattanayak, and Silvia Cordero-Sancho

ABSTRACT. *Payments for environmental services (PES) are popular despite little empirical evidence of their effectiveness. We estimate the impact of PES on forest cover in a region known for exemplary implementation of one of the best-known and longest-lived PES programs. Our evaluation design combines sampling that incorporates prematching, data from remote sensing and household surveys, and empirical methods that include partial identification with weak assumptions, difference-in-differences matching estimators, and tests of sensitivity to unobservable heterogeneity. PES in our study site increased participating farm forest cover by about 11% to 17% of the mean area under PES contract over eight years. (JEL Q57, Q58)*

I. INTRODUCTION

Economic theory views the problem of ecosystem degradation as a market failure, which can be resolved through transfers between the beneficiaries and the providers of ecosystem services, whether through government Pigouvian subsidies or Coasian contracting (Pattanayak, Wunder, and Ferraro 2010). Over the last 15 years, theory has been put into practice in the form of programs using payments for environmental services (PES). PES is an incentive-based conservation approach involving financial transfers to suppliers conditional on the supply of ecosystem services or on actions that are believed to generate ecosystem services (Wunder 2007; Engel, Pagiola, and Wunder 2008). PES programs now number in the hundreds globally (Porras, Grieg-Gran, and Neves 2008;

Ferraro 2009; Huang et al. 2009; Southgate and Wunder 2009).

Economic theory also suggests, however, that PES may have limited effectiveness in practice (Ferraro 2008; Wünscher, Engel, and Wunder 2008). PES effectiveness depends on where and to whom the payments go and the degree of compliance. For example, adverse self-selection and poor administrative targeting can direct payments to lands that are not threatened or of little environmental value.¹ Moreover, in developing nations, where PES has seen the most growth, there are myriad institutional design and governance challenges (Mueller and Albers 2004). Thus, despite theoretical arguments that PES programs can be cost-effective (Ferraro and Simpson 2002), their effectiveness is ultimately an empirical question (Ferraro and Pattanayak 2006).

¹ Poor targeting of conservation investments has long been a concern in the burgeoning conservation planning literature, which focuses on spatially, and sometimes temporally, optimizing limited conservation resources (e.g., Margules and Pressey 2000; Polasky, Camm, and Garber-Yonts 2001; Costello and Polasky 2004; Naidoo et al. 2006). Note that we focus on economic issues that can reduce PES effectiveness rather than ecological issues, such as the uncertainty of environmental measures (see, e.g., Kleijn et al. 2006).

The authors are, respectively, assistant professor, Department of Agricultural Economics, Pontificia Universidad Católica de Chile, Santiago; professor, Department of Economics, Andrew Young School of Policy Studies, Georgia State University, Atlanta; associate professor, Department of Forestry and Environmental Resources, North Carolina State University, Raleigh; associate professor, Sanford School of Public Policy and Nicholas School of the Environment, Duke University, Durham, North Carolina; Ph.D. candidate, School of Natural Resources and Environment, University of Michigan, Ann Arbor.

A recent review of the empirical evidence base on PES effectiveness finds few studies with credible empirical designs, that is, designs that can plausibly identify causal impacts (Pattanayak, Wunder, and Ferraro 2010). Most of these studies examine the longest-running and best-known PES program: Costa Rica's Programa de Pagos por Servicios Ambientales (PSA). Initiated in 1997, this national PSA program is financed by an earmarked gasoline tax, international donors, and environmental service buyers (Blackman and Woodward 2010). In its first two years, it allocated forest protection performance payment contracts on about 137,000 ha of forest. By the end of 2008, it had allocated contracts on about 600,000 ha.

Despite the substantial scientific and popular media attention directed at the PSA (Dirzo and Loreau 2005; Pearce 2008), the empirical results to date do not support claims that the PSA has had a large impact on land use.² Sierra and Russman (2006) use farm-level remotely sensed forest cover and simple ordinary least squares (OLS) regressions with a randomly selected group of PSA and non-PSA participants in the Osa Peninsula who had been surveyed in a livestock study. They conclude that PSA had no impact on total forest cover by 2003. Yet they observed a positive association between the PSA and the area of *charral* (regrowth) and a negative association between time since contract and the area in agricultural production. Sánchez-Azofeifa et al. (2007) use 535 km grid cells, remotely sensed forest cover, and OLS regressions to conclude that PSA contracts signed between 1997 and 2000 did not reduce deforestation rates or total deforestation in Costa Rica over that same time period. Two recent working papers (Pfaff, Robalino, and Sánchez-Azofeifa 2008; Robalino et al. 2008) use pixels,

remotely sensed forest cover, and a matching approach to select a control group from the set of all non-PSA pixels. They conclude that PSA contracts signed between 1997 and 2005 prevented deforestation on well under 1% of the land under contract.

One could interpret the results from these four studies as indicating that the first phase of this famous PES program was a failure. However, the study designs may suffer from one or more potential biases. For example, it is well known that the PSA was initially poorly targeted in much of the nation because of inadequate attention to costs and benefits (Hartshorn, Ferraro, and Spergel 2005; Wünscher, Engel, and Wunder 2008; Pfaff, Robalino, and Sánchez-Azofeifa 2008). Given potential heterogeneous impacts, detecting impacts may be difficult in national-scale studies or in areas where targeting or implementation was poor. Furthermore, we believe that previous PSA studies fail to adequately control for farm-level confounding factors that jointly determine PSA participation and forest-use decisions. Previous studies of PSA participation (Ortiz, Sage, and Borge 2003; Miranda, Porras, and Moreno 2003; Zbinden and Lee 2005; Arriagada et al. 2009; Morse et al. 2009) have consistently found that PSA participants differ from nonparticipants in important farm-level characteristics that directly affect land use.³

To contribute to this emerging literature, we make several advances. Guided by the insights from the conservation planning and mechanism design literatures, we select a study location where PSA impacts have the best chance of being observable: a region with active targeting. Second, we conduct our analysis at the farm level, which allows us to incorporate within-farm spillovers, both

² Although most claims in popular media and public presentations are informal, there are at least two formal analyses that argue the PSA has had a substantial impact. First, Ortiz, Sage, and Borge (2003) use cash-flow accounting to estimate that 22% of all forests under contract would have been deforested or degraded in the absence of the PSA. Second, an unpublished study claims that about 40% of the contracted area would have been deforested in the absence of the PSA (F. Tattenbach, personal communication, 2005; Pagiola 2006).

³ Sierra and Russman have a few proxies for farm-level attributes from secondary data, but lack a clear strategy for using these proxies to identify the PSA's causal impact. Pfaff, Robalino, Sánchez-Azofeifa (2008) have a clear identification strategy and rich secondary spatial covariate data, but lack some key farm-level attributes, such as previous participation in forestry programs, which have been shown to affect PSA eligibility and participation and likely affect land-use trends. They also lack geo-located farm boundaries and thus consider only pixels within contracted forest polygons.

negative and positive. Third, our study period spans eight years to allow time for treatment effects to become detectable. Fourth, to select PSA and similar non-PSA participants, we follow a subject-search protocol that is based on detailed understanding of the program assignment process. Fifth, to further control for confounding factors associated with assignment and selection into treatment, we complement the subject-search protocol with a difference-in-differences (DID) matching design and test the sensitivity of our estimates to unobservable heterogeneity.

In contrast to previous analyses, we find that the PSA had a moderately large mean treatment effect on PSA participant farms in our study site: additional forest cover equal to about 11% to 17% of the mean contracted forest area.

II. PAYMENTS FOR ENVIRONMENTAL SERVICES IN COSTA RICA

The Costa Rican government is well known for its conservation policies, including its large protected area network, support for private reserves, and efforts to control illegal logging (Snider et al. 2003). The PSA evolved from a long history of financial incentives for forest management, dating from at least 1969 when timber plantation expenditures became tax-deductible (Ortiz 2002; Rodríguez 2002). In contrast to previous conservation initiatives, however, the PSA offered direct payments to private landowners for forest preservation.⁴

The PSA was established by Forestry Law 7575 of 1996, which recognizes four environmental services: (1) mitigation of greenhouse gas emissions, (2) water protection, (3) protection of biodiversity, and (4) provision of scenic beauty. The law provided the regulatory basis for paying forest landowners for these services. Although PSA payments are tied to the area of forest under contract, rather than directly to service flows, lawmakers be-

lieved that framing the PSA in terms of the provision of environmental services made the importance of forest conservation more obvious and relevant to stakeholders (cf. Mainka, McNeely, and Jackson 2008). The law also established the National Fund for Forest Financing (*Fondo Nacional de Financiamiento Forestal* [FONAFIFO]), which is the agency in charge of administering the PSA program.

The first PSA contracts were signed in 1997. Our analysis focuses on contracts that were signed in 1997 or 1998 and were still in force in 2005. We describe the essential features of the PSA program during this period, but note that details affecting the enrollment of new contracts changed over time (see Arriagada 2008; Arriagada et al. 2009). Prior to 2000, the enrollment policy was essentially "first come, first served." In order to participate in the program, landowners had to produce an official cadastral map from the National Land Registry (*Oficina Catastro Nacional*), a copy of a cartographic map to indicate the location of the forest parcel to be considered for a PSA payment, and proof of land ownership. A professional topographer determined the size of the offered contract area, which could be less than the farm's total forested area. A professional forester prepared a forest management plan for approval by the National System of Conservation Areas (*Sistema Nacional de Areas de Conservación* [SINAC]). In some areas, local nongovernmental organizations act as intermediaries to facilitate the application process for landowners.

If a landowner's application were accepted, a contract would be signed and the government would make annual payments for five years with the potential for renewal. To lower transaction costs, FONAFIFO offered a uniform annual payment per hectare on all contracts (Sánchez-Azofeifa et al. 2007). In addition to the payment, PSA participants may benefit from greater tenure security against potential squatters (Miranda, Porras, and Moreno 2003; Porras and Hope 2005; Arriagada et al. 2009), as well as technical assistance from intermediary organizations. Landowners are required to protect the contracted area from deforestation or degradation

⁴ Our analysis focuses on forest preservation contracts. During the study period, the PSA also made payments for reforestation and, in some years, forest (timber) management. These payments are a small proportion of the contracted area (~11%).

(e.g., preventing fires, excluding livestock and refusing access to hunters). FONAFIFO, SINAC, and intermediaries may visit contracted parcels to ensure compliance. Decree No. 30761-MINAE charged SINAC with monitoring contract compliance and FONAFIFO with managing applications and payments.

III. STUDY REGION

The region of Sarapiquí, in the province of Heredia bordering Nicaragua and the San Juan River Basin, covers 2,934 km² and had a population of 222,467 in 2000. It comprises the cantons of Sarapiquí, Guacimo, Pococí, and Oreamuno and includes two of SINAC's management districts: Cordillera Volcánica Central (ACCVC) and Tortuguero (ACTO). The great climatic and topographic variation of the ACCVC results in high ecosystem diversity, with 9 out of 12 of Costa Rica's Holdridge life zones. The ACTO has more humid tropical forest than any other district in Costa Rica and is also known for its great diversity of bird species.⁵

The PSA program in Sarapiquí has been promoted and facilitated by a nongovernmental intermediary called the Foundation for the Development of the Central Volcanic Range (*Fundación para el Desarrollo de la Cordillera Volcánica Central* [FUNDECOR]). FUNDECOR recruits landowners, provides technical assistance, and maintains excellent records.⁶ In the first phase of the PSA program, FUNDECOR focused its efforts on establishing contracts in areas that it believed to be under greater deforestation threat (Arriagada et al. 2009). Interviews with FONAFIFO and SINAC employees indicated that FUNDECOR was one of the best-organized PSA intermediaries nationally and the only intermediary that explicitly focused on perceived deforestation threat.

We selected Sarapiquí as our study region for four reasons: (1) conducting a national-level analysis at the level of the farm would

be cost prohibitive and thus we wished to confine ourselves to one region; (2) FUNDECOR had a pool of 1997/98 contracts large enough to statistically evaluate PSA impacts on forest cover; (3) FUNDECOR had excellent records, which facilitated survey sampling; and (4) the combination of agricultural activity in the region and FUNDECOR's commitment to target and monitor the PSA contracts made the area one of the most likely places in Costa Rica to find an impact. In the words of a key informant, if an impact were not observed in Sarapiquí, it would be hard to imagine where an impact would be observed in Costa Rica. Thus, we make no claims that our Sarapiquí sample is representative of the rest of the country. Detecting an impact in Sarapiquí would not imply that impacts exist elsewhere, but rather it would simply imply that PES impacts are possible given similar conditions elsewhere. On the other hand, failure to find an impact would make claims of large impacts throughout the nation as a result of PSA during the study period harder to support.

IV. SAMPLING AND DATA COLLECTION

To develop a detailed understanding of PSA administration in Sarapiquí, we first conducted qualitative interviews and reviewed records at FUNDECOR, FONAFIFO, and SINAC (Arriagada et al. 2009). Proceeding in an iterative field research framework, we gathered information through semistructured interviews with government officials and forestry professionals, and through case studies of participant and nonparticipant forest landowners, based on in-depth interviews, field visits, and review of records. These case studies and semistructured interviews helped identify the determinants of participation in the first phase of the PSA, informing the sampling frame, questionnaire design, and empirical modeling.

For our household surveys and remote sensing, we focus on the FUNDECOR-mediated forest conservation contracts that were originally signed in 1997 or 1998 and were still in force when we conducted our survey in 2005: 70 contracts out of a total of 123 contracts that were signed in 1997–1998. We

⁵ Information on ACCVC and ACTO is available at www.sinac.go.cr.

⁶ More information about FUNDECOR can be found at www.fundecor.org.

focus on renewed contracts for three reasons: (1) shorter periods, like those used in other empirical analyses of PES, may not provide enough time for detectable impacts to arise; (2) records on contracts that had been renewed were better (including dates of initial application and renewal, ownership type, hectares owned, biophysical characteristics, and digital maps); (3) some contracts were not renewed because landowners could not meet new program requirements (e.g., legal title rather than just proof of possession was later required), and thus by excluding terminated contracts, we obtain a sample more representative of current PSA contracts. From the population of 70 contracts, 50 participants were randomly selected for household surveys and identification of farm land cover from remote sensing. These contracts, which form our treated group, were distributed across 13 districts in the four cantons of the Sarapiquí region.

Selecting nonparticipants at random from the Sarapiquí landscape to form a control group would not be a cost-effective way to find a valid counterfactual for PSA participants. Instead, we used three sampling methods that we believed were more likely to identify farms with baseline trends and characteristics similar to those of our treated group (i.e., a prematching screen): (1) a sample of immediate neighbors of program participants;⁷ (2) a random sample, stratified by district of PSA participants, using the National Land Registry as the sampling frame;⁸ and (3) a random sample, stratified by a defined buffer ring around each PSA farm, using the National Land Registry as the sampling frame. This buffer sample is randomly drawn from properties that are between 1,920 and 3,840 m from the centroids of PSA properties. The inner distance is set to avoid selecting any neighboring properties as controls, and the outer distance ensures controls are close to PSA farms and thus more likely to have simi-

lar observable and unobservable characteristics. We wanted to ensure that some of our control farms are not contiguous with PSA farms, should the interviews have revealed local spillovers from PSA farms to neighboring non-PSA farms (e.g., transmission of conservation messages). We found no evidence of such contamination and thus pool all three samples in the analysis. Because properties are highly irregular shapes, the inner limit was set equal to the 75th percentile of the diameters of the PSA properties in our sample (i.e., maximum straight-line distance across the properties), while the outer limit is 2.5 times the 75th percentile of the diameters. The buffering process created donut-shaped buffers, which in many cases overlapped. We therefore excluded any areas less than 1,920 m (the inner radius) from the centroid of any PSA property before randomly drawing a sample of properties located in these buffers.

The National Land Registry provides the most complete sampling frame of farms eligible for PSA contracts in 1997 and 1998. We excluded properties on the registry that were smaller than 5 ha (based on program rules and the minimum contract size among our sample of participants), properties listed in FONAFIFO's records as having PSA protection contracts sometime between 1997 and 2005, and properties owned by the state and large companies. For each stratification scheme (district and buffer), three landowners were selected at random for each participant. If the interviewer failed to find the first landowner after three documented attempts or if the landowner was ineligible for a PSA contract, the next landowner on the list would be sought. Before beginning an interview with a potential control landowner, the interviewer verified that the landowner had (1) owned or managed the farm since 1996, (2) had some forest cover on that farm in 1996; and (3) had never held a PSA forest protection contract.

Only one landowner in our selected control sample refused to be interviewed. However, because of ineligibility, remote locations, and lack of on-farm telecommunications (which placed practical constraints on the effectiveness of attempted contacts), we did not interview equal numbers of landowners in each sampling frame. To achieve our target of 150

⁷ If the selected landowner was ineligible to receive a PSA forest protection contract (e.g., no forest on property in 1996) or already had such a contract, or if the interviewer failed to find the neighbor after three documented attempts, another neighbor was selected (process that defines "documented attempts" available from authors).

⁸ One landowner selected through this search process had also been selected in the immediate neighbor search.

control farms, we increased the number of landowners sampled in the buffer sampling frame to compensate for lower numbers from the district sampling frame. Our control group comprises (1) 51 immediate neighbors, (2) 43 landowners in the sample stratified by districts, and (3) 58 landowners with properties located in buffers around each PSA property. Our sample size is similar to samples in recent applications of quasi-experimental evaluation approaches in resource and environmental policy (Greenstone 2004; Somanathan, Prabhakar, and Mehta 2009; Jumble and Angelsen 2006).

For each treated and control farm, we conducted a household survey to collect data on farm and farmer characteristics (Arriagada et al. 2009). Global positioning system (GPS) readings were taken on the farms and linked to maps from the National Land Registry to create a geographic information system (GIS) layer with property polygons. We interpreted aerial photographs to determine farm-level land cover changes, which serve as our outcome variable. We used 1992 aerial photos to establish baseline forest cover because there was excessive cloud cover in the 1997 photos. Areas obscured by cloud cover in the 1992 aerial photographs were classified using a Landsat 5 satellite image. At the Instituto Tecnológico de Costa Rica, aerial photographs were orthorectified and interpreted separately for 1986, 1992, and 2005 to obtain forest cover area in hectares (“forest” includes mature native forest and natural regeneration but not plantations).

Our outcome variable is the change in forest cover on the farm between 1992 and 2005. We use changes in farm-level forest cover, rather than in contracted parcel forest cover, for two reasons: (1) there is no clear analogue for “contracted area” among the control farms, and (2) the potential for within-farm displacement (“leakage”) of deforestation pressure from PES is a concern (Ferraro 2008). Thus the relevant unit for empirical analysis is the farm.⁹ We also measured the

change in farm forest cover between 1986 and 1992.

FUNDECOR records include geo-referenced property boundaries, and thus each PSA farm could be identified and land-use classified from aerial photographs. The variable quality of maps in the National Land Registry and some uncertainty matching immediate neighbors selected on a geographic basis to the correct property map prevented us from classifying land cover on all farms in the control group. We have classifications for 32 farms in buffers, 38 farms from districts, and 16 immediate neighbors. To address the missing data on the other control farms, we use a multiple multivariate imputation process (Rubin 1987; van Buuren, Boshuizen, and Knook 1999; Royston 2004). Imputation was done in STATA (v9) using the “switching regression” method of multiple multivariate imputation. As recommended in the literature, the imputed variables are averages of five imputed copies of the complete data set. In Section VI, we estimate treatment effects using the full data set with imputed values as well as a reduced set that includes only farms with observed values.

V. EMPIRICAL STRATEGY

We wish to estimate the average treatment effect on the treated: the difference between the expected potential change in forest cover on PSA farms with PSA contracts and the counterfactual expected potential change in forest cover on PSA farms without PSA contracts. We start with the partial identification approach of Manksi (1995, 2003), which uses the observed data on PSA and non-PSA farms to provide information about plausible ranges of the treatment effect using only weak assumptions.

We then apply stronger assumptions to further narrow the range of plausible treatment effect estimates. We use a simple DID estimator (called the before-after-control-impact estimator, in the ecology literature), which controls for time-invariant unobservable characteristics. When estimating the average treatment effect on the treated with the simple DID estimator, the key identification assumption is that the expected trend in forest cover of the

⁹ Given the small area of PSA contracts relative to the area of Sarapiquí, we are less concerned about general equilibrium spillover effects (positive or negative) across farms in the region.

TABLE 1
Covariate Balance (Full Sample with Imputed Values)

Variable	Sample ^a	Mean Value PSA	Mean Value Non-PSA ^b	Diff. Mean Value	Raw eQQ Diff. ^c
1992 forest cover (ha)	Unmatched	86.13	38.21	47.92	20.09
	Matched	86.13	68.99	17.13	4.96
Farm size (ha)	Unmatched	165.11	71.43	93.68	40.00
	Matched	165.11	108.46	56.65	17.00
Participation in past forestry programs	Unmatched	0.32	0.15	0.17	0.17
	Matched	0.32	0.28	0.03	0.03
Distance to extension office (km)	Unmatched	30.10	24.60	5.50	5.00
	Matched	30.10	27.78	2.33	1.00
% Farm on steep slopes	Unmatched	38.40	25.14	13.26	10.00
	Matched	38.40	38.11	0.29	0.00
Forest cover change 1986–1992	Unmatched	–11.35	–10.83	–0.52	5.89
	Matched	–11.35	–8.33	–3.02	2.00

Note: PSA, Costa Rica's Programa de Pagos por Servicios Ambientales.

^a $N = 202$; 50 PSA farms.

^b Weighted means for matched controls.

^c 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.

control units is equal to the expected trend in forest cover of the PSA units in the absence of the PSA program. To make this assumption plausible, recall that we first screened all control farms in our sample for program eligibility and then selected farms based on geographic rules expected to make treated and control farms more similar at baseline (a form of “prematching”). The mean DID for forest cover change on PSA and non-PSA farms between 1986 and 1992 (i.e., the trends *prior* to our baseline year) was 0.52 ha (std. err. 4.33; $p = 0.90$). This similarity in mean *pretreatment* trends gives us some faith in the plausibility of the identifying assumption of our DID design; that is, the expected *posttreatment* trends in the absence of the PSA program would be the same on treated and control farms.

Nevertheless, as seen in Table 1, the treated and control farms differ on baseline forest cover and other key baseline covariates that could affect both program participation and changes in forest cover between 1992 and 2005. The PSA farms, at baseline, tend to be larger and farther from forestry extension offices, with more forest cover, greater participation in previous forestry incentive programs, and on more steeply sloped land. Although the mean forest cover change from 1986 to 1992 is statistically identical for PSA

and non-PSA farms, the distributions are not ($p < 0.01$ based on a Kolmogorov-Smirnov test with bootstrapped p -values) (Sekhon 2007b). Thus one might worry that, despite our prematching effort and the similar trends in mean forest cover change before the PSA program, the mean change in forest cover among PSA farms from 1992 to 2005 in the absence of the PSA may not be well represented by the mean change in forest cover among the control farms during the same period.

To make the DID identification assumption more tenable, we use matching methods to preprocess the data and remove observable sources of bias (Ho et al. 2007). Smith and Todd (2005) found that the DID matching estimator performs best among other matching estimators, and Imbens and Wooldridge (2009) recommend combining methods in this way because results are more robust to misspecification, a common problem in parametric models. In other words, successful matching makes treatment effect estimates less dependent on the specific postmatching statistical model (Ho et al. 2007).

The goal of matching is to make the covariate distributions of PSA and non-PSA farms similar (called covariate balancing). To determine which variables to include in the matching algorithm, we use our knowledge of

the program and the way in which it was implemented in the Sarapiquí region, which we obtain from interviews and case studies with program staff and participants (Arriagada 2008; Arriagada et al. 2009). We also draw on previous descriptive work of PSA participation in Costa Rica (Zbinden and Lee 2005; Ortiz 2002) and the rich literature on tropical deforestation.

First and foremost, we believe that we should achieve balance on the baseline forest cover given that changes in forest cover are influenced by the initial forest area and that farms with greater areas of forest are more likely to participate in the PSA. Similarly, we want to control for baseline farm size, which affects the degree to which forest cover can change on a farm and also captures characteristics of the landowner that influence land-use decisions. Third, we know that during the early years of the PSA, the government did not actively publicize or promote the program, and thus most applications were from landowners who were familiar with MINAE regional offices and previous forestry programs. Thus we match on two other baseline conditions: (1) previous participation in forestry incentive programs and (2) distance to forestry extension offices. Controlling for previous program participation also reduces the likelihood that our estimated treatment effects arise from defunct forestry incentive programs rather than the PSA. The distance to extension offices also captures distance from forest law enforcement nodes. Moreover, given that forestry extension offices are often located in market towns, distance to a forestry extension office also represents distance to markets, which is an important factor in the deforestation literature.

Fourth, the deforestation literature reflects the Ricardian model of land conversion by emphasizing the biophysical capacity of the land. In the case of the PSA, Ortiz, Sage, and Borge (2003) predicted only marginal lands would be enrolled in PSA given the payment level. Biophysical capacity is often summarized in the deforestation literature by slope (Joppa and Pfaff 2010; Kaimowitz and Angelsen 1998), which has been found to be positively correlated with PSA participation (Zbinden and Lee 2005). Thus we match on

the percentage of steeply sloped land on the farm, as reported by the farmer.¹⁰ Finally, we want to ensure the baseline forest cover trends stays similar among treated and matched control units, so we match on forest cover change between 1986 and 1992.¹¹

Based on an assessment of covariate balance quality across a variety of matching methods (Ho et al. 2007; Sekhon 2007b), we chose one-to-one, nearest-neighbor covariate matching with replacement using a generalized version of the Mahalanobis distance metric and genetic matching algorithm that maximizes covariate balance (Diamond and Sekhon 2006). Matching was done in R (Sekhon 2007a). We use a postmatching bias-correction procedure that asymptotically removes the conditional bias in finite samples (Abadie and Imbens 2006b). Bootstrapped standard errors are invalid with nonsmooth, nearest-neighbor matching with replacement (Abadie and Imbens 2006a), and thus we use Abadie and Imbens' (2006b) variance formula to conduct a *t*-test of the mean DID.

Table 1 shows some metrics of covariate balance before and after matching for the sample with imputed values (see appendix for more information on balancing). The fifth and sixth columns of Table 1 present two measures of the differences in the covariate distributions between PSA and non-PSA farms: 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). Appendix Tables A1 and A2 present other balance measures. If matching is effective, these measures should move toward zero (Ho et al. 2007), which is

¹⁰ We believe that a farmer's perception of his or her land quality is more likely to influence behavior than a remotely sensed indicator of quality. Interestingly, neither the landowner nor government official surveys indicated that environmental beliefs are a major factor affecting participation.

¹¹ Note that contrary to what one might expect, forest cover change in the 1986–1992 period is negatively correlated with forest cover change in the 1992–2005 period among the control units. In other words, the less deforestation or more growth experienced before 1992, the more deforestation or less growth experienced between 1992 and 2005.

what we observe. Prior to matching, six out of the six variables exhibited statistically significant differences at the 5% level in either means (*t*-test) or in the overall distributions (Kolmogorov-Smirnov test with bootstrapped *p*-values) (Sekhon 2007b). After matching, none showed a difference.

To further address concerns about potential bias, we also present estimates based on matching using calipers to improve covariate balance. Calipers define a tolerance level for judging the quality of the matches: if available controls are not good matches for a treated unit (i.e., there is no match within the caliper), the unit is eliminated from the sample. Calipers reduce potential bias, but at the cost of estimating a treatment effect for only a subset of the sample. In our study, we view calipers as a robustness check. We define the caliper as one standard deviation of each matching covariate.

As an alternative postmatching model to estimate treatment effects and control for imperfect covariate balance, we estimated postmatching, linear regressions with the matching covariates. Given our sample size, we match on a small set of covariates that we believed to be most important to support the DID identification strategy. However, to allay concerns that the six covariates on which we match might not be sufficient to ensure that the trend in forest cover change on control farms is a valid counterfactual, we also run postmatching regressions on an extended set of covariates that might plausibly affect both PSA participation and forest cover change, but which were not used in the matching: age, baseline (1996) absentee landowner status, baseline experience with forestry plantations, percentage of farm with poor soils, baseline experience with private forester-written forestry management plans, birthplace in San Jose (the national capital), and baseline adult labor force. Postmatching regressions adjust for any small remaining imbalances across observable characteristics in the matched sample.

As in any observational (nonexperimental) study, unobservable heterogeneity threatens our ability to draw causal inferences. In spite of our efforts to control for observable and unobservable sources of bias through our DID

matching design, PSA participation and forest cover change may exhibit correlation in the absence of an effect of the PSA because of failure to match on a relevant but *unobserved* covariate. In our analysis, the main concern is that, in the absence of the PSA, PSA farms may have more regrowth or less deforestation than their matched controls because of unobservable factors. Sensitivity analysis examines the degree to which uncertainty about hidden biases in the assignment of PSA contracts could alter our conclusions.

We use Rosenbaum's (2002) recommended sensitivity test based on the Wilcoxon test statistic. This test assumes that each farm has a fixed value of an unobserved covariate (or a composite of unobserved covariates). This strong unobserved confounder not only affects PSA participation decisions, but also determines whether forest cover growth is more likely (or deforestation less likely) for the PSA farms or their matched controls. Thus this sensitivity test is conservative. Matched non-PSA farms 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 estimated effect of PSA on forest cover change remains significantly different from zero, the more confident one can be in the conclusion that the estimate is a causal effect. In other words, one can be more confident that the estimated treatment effect is not arising simply because of unobservable differences between PSA farms and their matched controls.

VI. RESULTS

First we place bounds on plausible estimates of the average treatment effect on the treated. With a systematic sampling strategy (as we have), in the limit, the expected potential outcome of Sarapiquí PSA farms under the PSA program is equal to the mean change in forest cover for the observed sample of Sarapiquí PSA farms: 10.74 ha.¹² The expected

¹² For PES studies that use the forest area under contract as the unit of the observation, rather than the entire farm, this value will be constrained to be less than or equal to zero. Such studies would not measure a net gain in total forest

potential outcome of PSA farms in the *absence* of the PSA program can be no greater than 78.98 ha, which equals the mean forest cover growth if every PSA farm went from its 1992 forest cover to a farm completely covered in forest in 2005 (i.e., the difference between average farm area and average forest area in 1992 in Table 1). The expected potential outcome of PSA farms in the absence of the PSA program can be no smaller than -86.13 ha, which is the mean forest cover decline if every PSA farm went from whatever forest area it had in 1992 to no forest at all in 2005. Thus, the constraints implied by the observed data alone, the so-called no-assumptions bound, place the average treatment effect on the treated on the interval $[-68.24$ ha, 96.87 ha].

We next invoke a monotone treatment selection assumption. We assume that, in the presence or absence of PSA contracts, PSA farms would have expected forest cover change outcomes at least as high as the control farms. This assumption would hold if there was positive self-selection, which we would expect for a voluntary program like the PSA. This implies that the simple DID estimator in the first row and column of Table 2 would be an upper bound for the average treatment effect on the treated. With this weak assumption, the bound on the treatment effect narrows to $[-68.24$ ha, 12.72 ha].

We next invoke a monotone treatment response assumption. We assume that the farm-level average treatment effect on the treated cannot be negative. In other words, placing forest under a PSA contract cannot induce greater deforestation than what would have been observed in the absence of the PSA payment. Although in the presence of credit constraints, this assumption is not innocuous, it is plausible. Invoking it narrows the bound on the treatment effect to $[0$ ha, 12.72 ha]. As is

typical with partial identification approaches, the resulting bound still includes zero, but we have ruled out extreme causal claims using weak distributional assumptions. For example, a common statistic put forth to describe the impacts of PES programs, including Costa Rica's PSA, is the simple cross-sectional difference in posttreatment outcome variables between the participants and nonparticipants. In the Sarapiquí case, that number is 60.64 ha (std. err. 13.75), which is not credible under the single weak assumption of monotone treatment selection.

Table 2 presents the DID matching estimates, whose interpretation requires much stronger assumptions. In the second column are estimates based on the full sample. The fifth row presents the postmatching DID estimate: 12.1 ha. The sixth row presents an estimate from a postmatching linear regression that linearly adjusts for any remaining imbalances among the matching covariates: 9.7 ha. The seventh row presents an estimate from a postmatching linear regression that linearly adjusts for any remaining imbalances among the extended set of covariates: 8.5 ha.

The matching DID estimates are roughly similar in magnitude and suggest that the PSA program induced about 8 to 12 ha of additional forest cover per PSA farm in Sarapiquí. The eighth row presents estimates using calipers, in which 12 PSA farms are dropped from the analysis. These farms tend to be large with large forest areas. Calipers reduce potential bias, but at the cost of estimating forest cover change on a subsample that may not be representative of the population of PSA farms in Sarapiquí. The estimated effect is smaller, yet still significantly different from zero.

In the third column of Table 2, we present estimates based on the reduced sample that excludes observations with imputed values for either outcomes or covariates. The simple DID estimator yields 13.96 ha, and the matching DID without calipers yields estimates that range from 9.6 to 11.2 ha. The caliper estimate is 7 ha. Thus the estimated magnitude of the PSA treatment effect is robust to excluding observations with imputed values.

To give the reader a sense of the relative magnitude of these estimates, we present several alternative measures based on the range

cover. Understanding net gains, however, is important, particularly in countries in which the so-called forest transition is underway (Daniels et al. 2010). In this transition, forest cover begins to slowly increase after having substantially declined following economic development, industrialization, and urbanization (see Mather and Needle 1998; Mather 1990; Walker 1993; Kates et al. 2001; Rudel 1998; Rudel et al. 2005).

TABLE 2
Estimated Impacts of PSA on Change in Forest Cover (1992–2005)

Outcome	Aerial Photos (Full Sample) ^a	Aerial Photos (Reduced Sample)
Simple DID ^b	12.72*** (3.96)	13.96*** (4.92)
<i>N</i> treated	50	50
<i>N</i> controls	152	86
Matching estimates		
Matching DID ^c	12.09** (4.89)	11.24*** (3.25)
Postmatching OLS regression, ^d core covariate set	9.70*** (2.96)	11.19*** (2.50)
Postmatching OLS regression, ^d extended covariate set	8.48*** (2.80)	9.60*** (3.04)
Matching DID with calipers ^{c,e}	4.65*** (1.41)	7.03*** (2.40)
<i>N</i> treated dropped by calipers	13	12
Γ^f	2.1	3.1
<i>N</i> treated	50	46
<i>N</i> available controls	152	70

Note: DID, difference in difference; OLS, ordinary least squares; PSA, Costa Rica's Programa de Pagos por Servicios Ambientales.

^a Full sample includes units for which outcome or covariate values may be imputed. Reduced sample restricts sample to units for which all values are observed.

^b A two-sided *t*-test of the difference in means between treated and control segments. Standard errors in parentheses.

^c Robust standard errors in parentheses (Abadie and Imbens 2006b).

^d Huber-White robust standard errors in parentheses. Estimates postmatching with calipers are similar and not reported here.

^e Calipers restrict matches to units within one standard deviation of each covariate.

^f Γ is log odds of differential assignment to treatment due to unobservables. Value reported is Γ at which the critical *p*-value for the estimate implies the effect is insignificantly different from zero at *p* = 0.10.

, * Significance at 5% and 1%, respectively.

of estimates from the full sample: 8.5–12.7 ha. This range corresponds to 10% to 15% of the mean forest cover in 1992 on PSA farms (86.13 ha) and 11% to 17% of the mean contracted forest area on PSA farms (75.5 ha). Another way to indicate the relative magnitudes of the impacts is to normalize the results using effect sizes, which are calculated by dividing the average treatment effect on the treated estimate by the standard deviation of the matched control units. The estimated effect sizes range from 0.60 to 0.93, which would be considered a moderately large effect size in other social policy fields (McCartney and Rosenthal 2000). A final way to view the results is to consider that the DID matching estimate of about 12 ha implies that, instead of losing, on average, 1.35 ha of forest (the bias-adjusted imputed counterfactual forest cover change), PSA farms gained about 10.74 ha on average.

The tenth row of Table 2 presents results from sensitivity tests that assess the degree to which potential unobservable heterogeneity makes us uncertain about the inferences drawn in the previous rows. We present the results of the sensitivity analysis for the matching DID estimate (the simple DID esti-

mate is less sensitive, so we present the more conservative of the two). The result indicates that the estimate of 12.09 ha remains significantly different from zero even in the presence of moderate unobserved bias. If an unobserved covariate caused the odds ratio of protection to differ between PSA and the matched non-PSA farms by a factor of as much as 2.1, the 90% confidence interval would still exclude zero. The reduced sample estimate of 11.24 is more robust. These results suggest that the change in forest growth as a result of the PSA program is likely to be greater than zero unless there is relatively strong hidden bias.

VII. DISCUSSION

As PES programs continue to proliferate globally, empirical estimates of their effectiveness are sorely needed. As the longest-lived and most widely studied PES system in the tropics, Costa Rica's PSA program continues to serve as a global leader in PES design (for a rare high-quality empirical study of a PES program outside of Costa Rica, see Alix-Garcia, Shapiro, and Sims 2010). Unlike previous studies that suggested Costa Rica's PSA

program had little to no effect on deforestation or forest cover, we estimate that in the Sarapiquí region, the PSA program had a moderate impact on forest cover.¹³ Rather than a net loss of forest cover, as implied by trends in matched control farms, there was a net increase in total forest cover on participating PSA farms. The PSA impact was equivalent to about 10% to 15% of the farm's pre-PSA forest cover. As with most analyses that use remotely sensed forest cover, we cannot identify the quality of the forest or quantity of ecosystem services gained, nor can we determine exactly how much of the PSA's impact in Sarapiquí comes from preventing the clearing of mature forest versus encouraging forest regrowth.

During the study period, FONAFIFO paid approximately \$43 per year per hectare under PSA contract (increased to ~\$62 per hectare in 2006). Ignoring administrative costs (estimated at about 5% of total payment budget [Ferraro and Kiss 2002]) and assuming that the treatment effect on forest cover was realized immediately upon contracting and sustained for the eight contracting years, we can generate a lower bound estimate of PSA cost-effectiveness (dollars per hectare gained per year over the study period). Our estimate of 8.5–12.7 ha of additional forest cover per farm for payments on about 75.5 ha of forest per farm implies that Costa Ricans (and donors) paid approximately \$255 to \$382 annually per hectare of additional forest induced by the PSA. The cost per hectare in earlier years could be higher if cumulative net gain in forest area over the study period were incremental (e.g., a couple of ha per year), or lower if the gains in early years were larger than the net gain observed at the end of the study period.

Of course, the PSA program continues to evolve and thus cost-effectiveness of more recent contracts may differ. Yet results from the first eight years of the program are still infor-

mative, especially when compared to existing results in the literature that show much smaller impacts. We cannot determine if the difference between our results and those from previous studies arises because of our choice of study region, the farm-level scale of analysis, our identification strategy, or all three. Nevertheless, our results highlight a potential problem with national-scale analyses of PES programs with poor targeting: one could erroneously determine that there were no impacts and thus undermine funding and political support for the program when in fact the program may have heterogeneous impacts (works in some places and not in others). Heterogeneity can arise because of heterogeneity in the quality of implementation or heterogeneous responses by different subgroups in the country. More research is needed to determine if our ability to detect an impact from the PSA in Sarapiquí arises from better targeting in the region, from Sarapiquí-specific characteristics that enhanced the behavioral response to the PSA, from the farm scale of analysis that allows for positive spillovers, or from some form of hidden bias in our analysis. Yet in light of the recent theoretical and empirical evidence that PES programs may have little or no impact on environmental outcomes, our results offer some hope that better designed and targeted PES can generate policy-relevant impacts (see also Alpízar, Blackman, and Pfaff 2007; Wünscher, Engel, and Wunder 2008; Robalino et al. 2008; Daniels et al. 2010). Future efforts to replicate our farm-level study in other regions of Costa Rica with more recent data can test the external validity of the results from Sarapiquí. In the absence of such replication, our Sarapiquí example suggests that PES can result in greater forest cover when appropriately designed and targeted.

PES and other incentive programs continue to be rolled out in many nations without any attempt to design them in ways that permit the evaluation of their effectiveness. Thus empirical designs that are effective for ex post evaluations are sorely needed. Our empirical strategy offers one design that, although expensive, can permit the careful evaluation of these globally popular approaches to environmental policy. However, we benefited from

¹³ Of course, comparisons across studies are complicated by, among other things, different units of analysis (pixels, grids, farms), different outcomes (deforestation, changes in forest cover, postprogram forest cover), and different baselines (Daniels et al. 2010).

the excellent records gathered by FUNDECOR, FONAFIFO, and the National Registry. Such record keeping is not present in all PES programs.

Program designers would do well to incorporate prospective ex ante evaluation designs into the program implementation itself (e.g., randomized experimental designs; program-designed instrumental variables). While such

evaluation designs may not have been politically feasible or desirable in the early years of the PSA program, the program is now well established and experiences excess demand for the available funds. This maturation would permit FONAFIFO and its partners, like FUNDECOR, to allocate contracts in ways that will facilitate impact evaluations in the future.

APPENDIX

TABLE A1
Covariate Balance (Full Sample with Imputed Values)

Variable	Sample ^a	Mean Value PSA	Mean Value Non-PSA ^b	Diff. Mean Value	Avg. Raw eQQ Diff. ^c	Mean eCDF Diff. ^d	p-Values for t-Test/K-S Test
1992 forest cover (ha)	Unmatched	86.13	38.21	47.92	20.09	0.14	0.03/0.04
	Matched	86.13	68.99	17.13	4.96	0.04	0.16/0.84
Farm size (ha)	Unmatched	165.11	71.43	93.68	40.00	0.19	0.06/<0.01
	Matched	165.11	108.46	56.65	17.00	0.08	0.17/0.22
Participation in forestry programs	Unmatched	0.32	0.15	0.17	0.17	0.14	0.02
	Matched	0.32	0.28	0.03	0.03	0.02	0.22
Distance to extension office (km)	Unmatched	30.10	24.60	5.50	5.00	0.10	0.02/0.08
	Matched	30.10	27.78	2.33	1.00	0.03	0.23/0.71
% Farm on steep slopes	Unmatched	38.40	25.14	13.26	10.00	0.13	<0.01/0.06
	Matched	38.40	38.11	0.29	0.00	0.04	0.83/0.87
Forest cover change 1986–1992	Unmatched	-11.35	-10.83	-3.02	5.89	0.14	0.94/<0.01
	Matched	-11.35	-8.33	-0.52	2.00	0.08	0.46/0.26

Note: PSA, Costa Rica's Programa de Pagos por Servicios Ambientales.

^a N = 202; 50 PSA farms.

^b Weighted means for matched controls.

^c 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.

^d Mean eCDF = mean differences in empirical cumulative distribution function.

TABLE A2
Covariate Balance (Reduced Sample with No Imputed Values)

Variable	Sample ^{a,i}	Mean Value PSA	Mean Value Non-PSA ^b	Diff: Mean Value	Avg. Raw eQQ Diff. ^c	Mean eCDF Diff. ^d	p-Values for t-Test/K-S Test
1992 forest cover (ha)	Unmatched	91.00	34.68	56.32	32.82	0.22	0.02/<0.01
	Matched	91.00	72.53	18.47	5.80	0.04	0.12/0.77
Farm size (ha)	Unmatched	161.71	80.98	80.73	30.00	0.16	0.14/<0.01
	Matched	161.71	123.17	38.53	7.50	0.04	0.40/0.59
Participation in forestry programs	Unmatched	0.33	0.14	0.18	0.20	0.09	0.03
	Matched	0.33	0.28	0.04	0.04	0.02	0.16
Distance to extension office (km)	Unmatched	29.15	24.96	4.20	50.00	0.08	0.14/0.17
	Matched	29.15	28.65	0.50	0.00	0.03	0.69/0.99
% Farm on steep slopes	Unmatched	38.15	22.49	15.67	0.14	0.14	<0.01/0.05
	Matched	38.15	32.74	5.41	0.08	0.09	0.15/0.21
Forest cover change 1986–1992	Unmatched	-3.92	-10.67	6.75	5.23	0.20	0.03/<0.01
	Matched	-3.92	-6.26	2.34	1.37	0.07	0.24/0.10

Note: PSA, Costa Rica's Programa de Pagos por Servicios Ambientales.

^a N = 116; 46 PSA farms.

^b Weighted means for matched controls.

^c 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.

^d Mean eCDF = mean differences in empirical cumulative distribution function.

Acknowledgments

We gratefully acknowledge financial support from the National Science Foundation (SES-0519194) and FONAFIFO. Pattanayak acknowledges Conservation International for partial support. Edgar Ortiz, Katie Caldwell, Luis Carrasco, and Kwaw Andam contributed to study design, data collection, and data analysis and interpretation. We thank Francisco Alpizar, Allen Blackman, Stefano Pagiola, Juan Robalino, Kerry Smith, Sven Wunder, and two anonymous reviewers for helpful comments on earlier versions of this manuscript.

References

- Abadie, Alberto, and Guido Imbens. 2006a. "On the Failure of the Bootstrap for Matching Estimators." NBER Technical Working Paper 325. Cambridge, MA: National Bureau of Economic Research.
- . 2006b. "Large Sample Properties of Matching Estimators for Average Treatment Effects." *Econometrica* 74 (1): 235–67.
- Alix-Garcia, Jennifer, Elizabeth Shapiro, and Katharine Sims. 2010. "Forest Conservation and Slippage: Evidence from Mexico's National Payments for Ecosystem Services Program." Working paper, University of Wisconsin–Madison, Duke University, and Amherst College.
- Alpizar, Francisco, Allen Blackman, and Alexander Pfaff. 2007. "Payments for Ecosystem Services: Measurements with Impacts." *Resources* 165 (Spring): 22–24.
- Arriagada, Rodrigo. 2008. "Private Provision of Public Goods: Applying Matching Methods to Evaluate Payments for Ecosystem Services in Costa Rica." Ph.D. dissertation. Department of Forestry and Environmental Resources, North Carolina State University, Raleigh.
- Arriagada, Rodrigo, Erin O. Sills, Subhrendu K. Pattanayak, and Paul J. Ferraro. 2009. "Combining Qualitative and Quantitative Methods to Evaluate Participation in Costa Rica's Program of Payments for Environmental Services." *Journal of Sustainable Forestry* 28 (3): 343–67.
- Blackman, Allen, and Richard T. Woodward. 2010. "User Financing in a National Payments for Environmental Services Program: Costa Rican Hydropower." *Ecological Economics* 69 (8): 1626–38.
- Costello, Christopher, and Stephen Polasky. 2004. "Dynamic Reserve Selection." *Resource and Energy Economics* 26 (2): 157–74.
- Daniels, Amy E., Kenneth Bagstad, Valerie Esposito, Azur Moulart, and Carlos Manuel Rodriguez. 2010. "Understanding the Impacts of Costa Rica's PES: Are We Asking the Right Questions?" *Ecological Economics* 69 (11): 2116–26.
- Diamond, Alexis, and Jasjeet Sekhon. 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: University of California.
- Dirzo, Rodolfo, and Michel Loreau. 2005. "Biodiversity Science Evolves." *Science* 310 (5750): 943.
- Engel, Stefanie, Stefano Pagiola, and Sven Wunder. 2008. "Designing Payments for Environmental Services in Theory and Practice: An Overview of the Issues." *Ecological Economics* 65 (4): 663–74.
- Ferraro, Paul J. 2008. "Asymmetric Information and Contract Design for Payments for Environmental Services." *Ecological Economics* 65 (4): 810–21.
- . 2009. "Regional Review of Payments for Watershed Services: Sub-Saharan Africa." *Journal of Sustainable Forestry* 28 (3): 525–50.
- Ferraro, Paul J., and Agnes Kiss. 2002. "Direct Payments for Biodiversity Conservation." *Science* 298: 1718–19.
- Ferraro, Paul J., and Subhrendu K. Pattanayak. 2006. "Money for Nothing? A Call for Empirical Evaluation of Biodiversity Conservation Investments." *PLoS Biology* 4 (4): 482–88.
- Ferraro, Paul J., and R. David Simpson. 2002. "The Cost-effectiveness of Conservation Performance Payments." *Land Economics* 78 (3): 339–53.
- Greenstone, Michael. 2004. "Did the Clean Air Act Cause the Remarkable Decline in Sulfur Dioxide Concentrations?" *Journal of Environmental Economics and Management* 47 (3): 585–611.
- Hartshorn, Gary, Paul J. Ferraro, and Barry Spergel. 2005. *Evaluation of the World Bank: GEF Ecomarkets Project in Costa Rica*. Report prepared for the World Bank. Raleigh: North Carolina State University.
- Ho Daniel, Kosuke Imai, Gary King, and Elizabeth A. Stuart. 2007. "Matching as Nonparametric Pre-processing for Reducing Model Dependence in Parametric Causal Inference." *Political Analysis* 15 (3): 199–236.
- Huang, Marjorie, Shyam Upadhyaya, Rohit Jindal, and John Kerr. 2009. "Payment for Watershed Services in Asia: A Review of Current Initiatives." *Journal of Sustainable Forestry* 28 (3): 551–75.
- Imbens, Guido W., and Jeffrey M. Wooldridge. 2009. "Recent Developments in the Econometrics of Program Evaluation." *Journal of Economic Literature* 47 (1): 5–86.
- Joppa Lucas, and Alexander Pfaff. 2010. "Reassessing the Forest Impacts of Protection: The Challenge of Nonrandom Location and a Corrective

- Method." *Annals of the New York Academy of Sciences* 1185 (1): 135–39.
- Jumbe, Charles B. L., and Arild Angelsen. 2006. "Do the Poor Benefit from Devolution Policies? Evidence from Malawi's Forest Co-management Program." *Land Economics* 82 (4): 562–81.
- Kaimowitz, David, and Arild Angelsen. 1998. *Economic Models of Deforestation: A Review*. Bogor, Indonesia: Center for International Forestry Research.
- Kates, Robert, William Clark, Robert Corell, J. Michael Hall, Carlo C. Jaeger, Ian Lowe, James J. McCarthy, Hans Joachim Schellnhuber, Bert Bolin, Nancy Dickson, Sylvie Faucheux, Gilberto C. Gallopin, Arnulf Grübler, Brian Huntley, Jill Jäger, Narpal S. Jodha, Roger E. Kasperson, Akin Mabogunje, Pamela Matson, and Harold Mooney. 2001. "Sustainability Science." *Science* 292 (5517): 641–42.
- Kleijn, D., R. A. Baquero, Y. Clough, M. Díaz, J. De Esteban, F. Fernández, D. Gabriel, F. Herzog, A. Holzschuh, R. Jöhl, E. Knop, A. Kruess, E. J. P. Marshall, I. Steffan-Dewenter, T. Tscharntke, J. Verhulst, T. M. West, and J. L. Yela. 2006. "Mixed Biodiversity Benefits of Agri-environment Schemes in Five European Countries." *Ecology Letters* 9 (3): 243–54.
- Mainka, Susan, Jeffrey McNeely, and William Jackson. 2008. "Depending on Nature: Ecosystem Services for Human Livelihoods." *Environment* 50 (2): 42–55.
- Manski, Charles F. 1995. *Identification Problems in the Social Sciences*. Cambridge, MA: Harvard University Press.
- . 2003. *Partial Identification of Probability Distributions*. New York: Springer.
- Margules, Chris R., and Robert L. Pressey. 2000. "Systematic Conservation Planning." *Nature* 405 (11 May): 243–53.
- Mather, Alexander. 1990. *Global Forest Resources*. London: Bellhaven Press.
- Mather, Alexander, and Coby Needle. 1998. "The Forest Transition: A Theoretical Basis." *Area* 30 (2): 117–24.
- McCartney, Kathleen, and Robert Rosenthal. 2000. "Effect Size, Practical Importance, and Social Policy for Children." *Child Development* 71 (1): 173–80.
- Miranda, Miriam, Inas Porras, and Mary Luz Moreno. 2003. "The Social Impacts of Payments for Environmental Services in Costa Rica. A Quantitative Field Survey and Analysis of the Virilla Watershed." Unpublished paper, International Institute for Environment and Development, London.
- Morse, Wayne C., Jessica L. Schedlbauer, Steven E. Sesnie, Bryan Finegan, Celia A. Harvey, Steven J. Hollenhorst, Kathleen L. Kavanagh, Dietmar Stoian, and J. D. Wulffhorst. 2009. "Consequences of Environmental Service Payments for Forest Retention and Recruitment in a Costa Rican Biological Corridor." *Ecology and Society* 14 (1): 23. Available at www.ecologyandsociety.org/vol14/iss1/art23/.
- Mueller, Jeffrey, and Heidi J. Albers. 2004. "Enforcement, Payments, and Development Projects near Protected Areas: How the Market Setting Determines What Works Where." *Resource and Energy Economics* 26 (2): 185–204.
- Naidoo, Robin, Andrew Balmford, Paul J. Ferraro, Stephen Polasky, Taylor Ricketts, and Mathieu Rouget. 2006. "Integrating Economic Costs into Conservation Planning." *Trends in Ecology and Evolution* 21 (2): 681–87.
- Ortiz, Edgar. 2002. *Sistema de cobro y pago por servicios ambientales en Costa Rica*. Serie de Apoyo Académico No. 34. Cartago: Escuela de Ingeniería Forestal, Instituto Tecnológico de Costa Rica.
- Ortiz, Edgar, Luis Sage, and Carlos Borge. 2003. *Impacto del programa de pago de servicios ambientales en Costa Rica como medio de reducción de la pobreza en los medios rurales*. Serie de Publicaciones RUTA. San José, Costa Rica: Unidad Regional de Asistencia Técnica.
- Pagiola, Stefano. 2006. "Payments for Environmental Services in Costa Rica." MPRA Paper No. 2010, November 2006. Available at <http://mpra.ub.uni-muenchen.de/2010/>.
- Pattanayak, Subhrendu K., Sven Wunder, and Paul J. Ferraro. 2010. "Show Me the Money: Do Payments Supply Ecosystem Services in Developing Countries?" *Review of Environmental Economics and Policy* 4 (2): 254–74.
- Pearce, Fred. 2008. "Save the Climate by Saving the Forests." *New Scientist* 2008 (2648): 36–39.
- Pfaff, Alexander, Juan Robalino, and G. Arturo Sánchez-Azofeifa. 2008. "Payments for Environmental Services: Empirical Analysis for Costa Rica." Working Paper Series SAN08-05. Durham, NC: Terry Sanford Institute of Public Policy, Duke University.
- Polasky, Stephen, Jeffrey D. Camm, and Brian Garber-Yonts. 2001. "Selecting Biological Reserves Cost-effectively: An Application to Terrestrial Vertebrate Conservation in Oregon." *Land Economics* 77 (1): 68–78.
- Porras, Inas, Maryanne Grieg-Gran, and Nanete Neves. 2008. *All that Glitters: A Review of Payments for Watershed Services in Developing Countries*. London: International Institute for Environment and Development.
- Porras, Inas, and R. A. Hope. 2005. "Using Stated Choice Methods in the Design of Payments for Environmental Services Schemes." Unpublished

- paper, Center for Land Use and Water Resources Research (CLUWRR), Newcastle University, UK.
- Robalino, Juan, Alexander Pfaff, G. Arturo Sánchez-Azofeifa, Francisco Alpizar, Carlos León, and Carlos Manuel Rodríguez. 2008. "Deforestation Impacts of Environmental Services Payments: Costa Rica's PSA Program, 2000–2005." Environment for Development Discussion Paper 08-24. Washington, DC: Resources for the Future.
- Rodríguez, Jorge. 2002. *El éxito forestal de Costa Rica en cinco casos*. San José, Costa Rica: Ministerio de Ambiente y Energía.
- Rosenbaum, Paul. 2002. *Observational Studies*, 2nd ed. New York: Springer-Verlag.
- Royston, Patrick. 2004. "Multiple Imputation of Missing Values." *Stata Journal* 4 (3): 227–41.
- Rubin, Donald. 1987. *Multiple Imputation for Non-response in Surveys*. New York: John Wiley and Sons.
- Rudel, Thomas. 1998. "Is There a Forest Transition? Deforestation, Reforestation, and Development." *Rural Sociology* 63 (4): 533–52.
- Rudel, Thomas, Oliver Coomes, Emilio Moran, Fred-eric Achard, Arild Angelsen, Jianchu Xu, and Eric Lambin. 2005. "Forest Transitions: Towards a Global Understanding of Land Use Change." *Global Environmental Change* 15 (1): 23–31.
- Sánchez-Azofeifa, G. Arturo, Alexander Pfaff, Juan Robalino, and Judson Boomhower. 2007. "Costa Rica's Payment for Environmental Services. Program: Intention, Implementation, and Impact." *Conservation Biology* 21 (5): 1165–73.
- Sekhon, Jasjeet S. 2007a. "Multivariate and Propensity Score Matching Software with Automated Balance Optimization: The Matching Package for R." *Journal of Statistical Software* 42 (7). Available at <http://sekhon.berkeley.edu/papers/MatchingJSS.pdf> (accessed January 10, 2008).
- . 2007b. "Alternative Balance Metrics for Bias Reduction in Matching Methods for Causal Inference." Working paper, January 16, University of California–Berkeley.
- Sierra, Rodrigo, and Eric Russman. 2006. "On the Efficiency of Environmental Service Payments: A Forest Conservation Assessment in the Osa Peninsula, Costa Rica." *Ecological Economics* 59 (1): 131–41.
- Smith, Jeffrey, and Petra Todd. 2005. "Does Matching Overcome LaLonde's Critique of Nonexperimental Estimators?" *Journal of Econometrics* 125 (1–2): 305–53.
- Snider, Anthony G., Subhrendu K. Pattanayak, Erin O. Sills, and Jamie L. Schuler. 2003. "Policy Innovations for Private Forest Management and Conservation in Costa Rica." *Journal of Forestry* 101 (5):18–23.
- Somanathan, E., R. Prabhakar, and Bhupendra S. Mehta. 2009. "Decentralization for Cost-effective Conservation." *Proceedings of the National Academy of Sciences* 106 (11): 4143–47.
- Southgate, Douglas, and Sven Wunder. 2009. "Paying for Watershed Services in Latin America: A Review of Current Initiatives." *Journal of Sustainable Forestry* 28 (3): 497–524.
- van Buuren, S., H. C. Boshuizen, and D. L. Knook. 1999. "Multiple Imputation of Missing Blood Pressure Covariates in Survival Analysis." *Statistics in Medicine* 18 (6): 681–94.
- Walker, Robert. 1993. "Deforestation and Economic Development." *Canadian Journal of Regional Science* 16 (3): 481–97.
- Wunder, Sven. 2007. "The Efficiency of Payments for Environmental Services in Tropical Conservation." *Conservation Biology* 21 (1): 48–58.
- Wünscher, Tobias, Stefanie Engel, and Sven Wunder. 2008. "Spatial Targeting of Payments for Environmental Services: A Tool for Boosting Conservation Benefits." *Ecological Economics* 65 (4): 822–33.
- Zbinden, Simon, and David R. Lee. 2005. "Paying for Environmental Services: An Analysis of Participation in Costa Rica's PSA Program." *World Development* 33 (2): 255–72.