

---

# Abordagens analíticas na avaliação de impactos reais de programas de conservação

ERIN O. SILLS<sup>1\*</sup>  
SUBHRENDU K. PATTANAYAK<sup>2</sup>  
PAUL J. FERRARO<sup>3</sup>  
KEITH ALGER<sup>4</sup>

<sup>1</sup> Departamento de Ciências Florestais e Recursos Ambientais, Universidade Estadual da Carolina do Norte, EUA.

<sup>2</sup> Programa de Economia do Meio Ambiente, Saúde e Desenvolvimento, Research Triangle Institute, Carolina do Norte, EUA.

<sup>3</sup> Departamento de Economia, Escola Andrew Young de Política Pública, da Universidade Estadual da Geórgia, EUA.

<sup>4</sup> Programa Dimensões Humanas da Conservação de Biodiversidade, Centro para Pesquisa Aplicada à Biodiversidade, Conservation International, Washington, D.C.

\* e-mail: erin\_sills@ncsu.edu

## RESUMO

A importância de monitoramento e avaliação é cada vez mais reconhecida na comunidade conservacionista. Neste artigo, nós argumentamos que os métodos de 'avaliação de programas' desenvolvidos em outras disciplinas deveriam ser utilizados para testar causalidade em projetos de conservação. Esses métodos incluem técnicas estatísticas e quase-experimentais para construir os casos não observados de não-tratamento pela intervenção: o que teria acontecido sem a intervenção e o que aconteceria caso a intervenção fosse aplicada em algum outro lugar. Embora reconheçamos a natureza multidimensional de monitoramento e avaliação, enfatizamos a necessidade crescente de análises quantitativas da eficácia de intervenções, controlando fatores de confusão, com o objetivo de dar aos tomadores de decisão maior segurança de que recursos adicionais podem realmente aumentar a escala dos resultados da conservação. Qualquer projeto conservacionista que assimile os métodos e medições abordadas nesse artigo farão uma vital contribuição em direção ao preenchimento da grande lacuna no nosso conhecimento sobre os mais eficazes investimentos em conservação.

## ABSTRACT

*The importance of monitoring and evaluation is increasingly recognized in the conservation community. In this article, we argue that 'program evaluation' methods developed in other fields should be used to test for causality in conservation programs. These methods include quasi-experimental and statistical techniques to construct the unobserved counter-factuals: what would happen without the intervention, and what would happen if the intervention were applied elsewhere. While recognizing the multi-dimensional nature of monitoring and evaluation, we emphasize the increasing need for quantitative analysis of the effectiveness of interventions that controls for confounding factors. This will give policy makers greater confidence that additional resources can*

*actually scale-up conservation results. Each conservation project that incorporates the methods and measurements reviewed in the paper will make a vital contribution towards filling the large gap in our knowledge about the most effective conservation investments.*

## INTRODUÇÃO

A importância de monitoramento e avaliação é cada vez mais reconhecida na comunidade conservacionista (Christensen, 2003; Kleiman *et al.*, 2000; Sanderson 2002). Embora organizações conservacionistas costumem empregar vários métodos distintos de avaliação, alguns princípios comuns são reconhecidos e são, inclusive, postos em prática com frequência. Um desses princípios é a importância de identificar e medir os objetivos de um projeto ou programa – ou seja, não somente os insumos ou atividades (p.ex., número de domicílios ou hectares envolvidos), mas também os resultados ou metas (p.ex., aumento da riqueza familiar e biodiversidade preservada). Outro princípio comum é a necessidade de desenvolver um modelo conceitual das relações causais entre insumos ou atividades e os resultados almejados. O planejamento fundamentado em relações causais explícitas e o monitoramento baseado em indicadores de resultado são elementos chave de sistemas de monitoramento e avaliação em desenvolvimento (p.ex., a abordagem de marco lógico (*log frame*) utilizada pela USAID e descrita em Gasper 2000) e conservação (p.ex., os padrões abertos propostos pela *Conservation Measures Partnership*, CMP 2004). No entanto, esses sistemas não demonstram como testar as relações causais postuladas, utilizando os indicadores de resultados (Ferraro & Pattanayak, 2006).

Neste artigo, nós argumentamos que os métodos de ‘avaliação de programas’ desenvolvidos em outras disciplinas deveriam ser utilizados para testar causalidade em projetos de conservação. Esses métodos permitem aos pesquisadores isolar mudanças atribuíveis ao esforço conservacionista de um *background* de mudança econômica e social simultânea. Nós o ilustramos com exemplos do estado do Acre, cujo governo tem experimentado várias políticas inovadoras de desenvolvimento sustentável de base florestal (Kainer *et al.*, 2003). Isso inclui recomendações de políticas públicas populares, como o pagamento por serviços ambientais (Fearnside, 2005) e pacotes combinando construção de estradas e melhoramentos em serviços sociais e em fiscalização ambiental (Reid & Souza, 2005). Dados os limitados recursos disponíveis à conservação, é crítico que se tome lições de tais experimentos políticos e que

se identifique quais intervenções são as mais eficazes, rejeitando potenciais efeitos de confusão (CMP, 2004). Métodos de avaliação de programas alcançam isso aplicando o mesmo rigor científico e o estado da arte em metodologia tipicamente utilizados para testar hipóteses ecológicas (Ferraro & Pattanayak, 2006). Esses métodos incluem técnicas estatísticas e quase-experimentais para construir os casos não observados de não-tratamento pela intervenção: o que teria acontecido sem a intervenção e o que aconteceria caso a intervenção fosse aplicada em algum outro lugar.

Para ilustrar tanto a importância como o desafio da avaliação de programas, considere um programa como o Pró-Ambiente, que paga aos proprietários de florestas por serviços ambientais. Os proprietários são mais propensos a serem voluntários em tal programa se eles forem mais inclinados a conservar a floresta, talvez porque dão mais valor à sua floresta, ou porque lucrariam menos com usos alternativos da terra. Isso sugere que o impacto dos pagamentos seria superestimado pela comparação de florestas conservadas entre participantes e não-participantes. Por outro lado, se os gestores de programas direcionam pagamentos a áreas que são ameaçadas pelo desmatamento, a comparação de florestas conservadas em áreas cobertas pelo programa com áreas não cobertas subestimaria o impacto dos pagamentos. Tanto um caso como o outro são exemplos de seleção endógena, o que provavelmente caracteriza muitas, se não a maioria das intervenções conservacionistas.

Descrevemos aqui monitoramento e avaliação convencionais, ilustrando as suas limitações no contexto de dois exemplos do Acre. Inspirando-nos em Ferraro & Pattanayak (2006), revisamos os métodos quase-experimentais de avaliação de programas, incluindo desenho amostral, coleta de dados e análise dos impactos de intervenções não-aleatórias, como programas pagando por serviços ambientais. Nós traçamos exemplos de avaliações de políticas sociais no Brasil (p.ex., fornecimento de água e saneamento, redução de pobreza, nutrição) e sugerimos como os métodos poderiam ser aplicados no Acre. Essencialmente, todos esses métodos são elaborados para compensar a falta de controles experimentais, portanto concluímos com uma discussão sobre o potencial de realização de experimentos reais de políticas públicas.

## MONITORAMENTO E AVALIAÇÃO CONVENCIONAL DOS IMPACTOS PROGRAMÁTICOS

Sistemas de monitoramento e avaliação convencionais servem a propósitos múltiplos, desde à responsabilização dos executores, à garantias de prestação de contas, ao aprendizado e à orientação de políticas públicas. (Hailey & Sorgenfrei, 2003). Em conservação, monitoramento e avaliação são freqüentemente conceitualizados como parte do Manejo Adaptativo (CMP, 2004; Margolis & Salafsky, 1998). A *Conservation Measurement Partnership* recomenda um sistema que possa “isolar e avaliar os efeitos de sua intervenção”, considerando “o processo e impactos assim como potenciais fatores de confusão” (CMP, 2004). No campo mais amplo do setor não-governamental e de desenvolvimento, Hailey & Sorgenfrei (2003) também apontam uma crescente ênfase na avaliação de impactos e eficácia. Eles percebem que há algum “consenso sobre as características chave de sistemas de avaliação de impacto eficazes. Estas incluem participação de *stakeholders*<sup>1</sup>, auto-avaliação, triangulação e uso de uma mescla de métodos e ferramentas, assim como sólida análise e comunicação através da disseminação das conclusões”. Essa mescla freqüentemente inclui “medidas não-financeiras, qualitativas e baseadas em processos”, assim como as tradicionais medidas quantitativas de performance. Embora reconheçamos a natureza multidimensional de monitoramento e avaliação, enfatizamos a necessidade crescente de análises quantitativas da eficácia de intervenções, controlando fatores de confusão, com o objetivo de dar aos *policy makers*<sup>2</sup> maior segurança de que recursos adicionais podem realmente aumentar a escala dos resultados da conservação.

Avaliação de impactos requer a identificação do objetivo ou resultado almejado e de indicadores mensuráveis, e o emprego de uma técnica apropriada para coleta de dados. Por exemplo, se o resultado almejado é a manutenção dos serviços ambientais fornecidos por florestas, o indicador pode ser área florestal e a técnica para coleta de dados, o sensoriamento remoto. Por sua vez, se o objetivo é a manutenção da biodiversidade, um dos indicadores poderia ser relacionado à atividade de caça, sendo os dados coletados através de um

levantamento domiciliar. Freqüentemente, esses dados são coletados somente de participantes (ou equivalentemente, áreas focais) em um programa. Isso pode fornecer a base para um valioso estudo de caso, mas gera pouca informação quantitativa sobre impactos. Avaliações mais rigorosas, portanto, comparam o indicador tanto ‘antes como depois’ ou ‘com e sem’ programa. O primeiro caso requer planejamento adiantado para que os participantes sejam entrevistados antes de uma intervenção, estabelecendo uma *baseline*<sup>3</sup>. O segundo requer a coleta de informações sobre um grupo mais abrangente tanto de participantes como de não-participantes do programa.

O problema com comparações simples baseadas nesses desenhos amostrais comuns é que eles não capturam efeitos de confusão – efeitos que são contemporâneos à intervenção e que podem, plausivelmente, afetar o resultado e assim mascarar o seu efeito. Como descrito por Ferraro & Pattanayak (2006), eles incluem tendências históricas, programas ou políticas não relacionados, características ambientais e sociais não observadas e seleção endógena. Eles mencionam que “as análises correntes tipicamente não consideram as implicações relacionadas ao que leva uma área a ser selecionada para uma intervenção e outra rejeitada, ou a de que alguns indivíduos sejam ‘voluntários’ e outros não. Em qualquer programa com aplicação não-aleatória, as características que influenciam a variável resultado também influenciam com freqüência a probabilidade de serem selecionadas no programa. A inobservância da questão da seleção endógena pode conduzir a estimativas enviesadas da eficácia de um programa”.

Por exemplo, um estudo na África constatou que “o sucesso de áreas protegidas não estava diretamente correlacionado com benefícios relativos à geração de empregos na comunidade vizinha, à educação conservacionista, às associações conservacionistas, ou à presença e alcance de Programas Integrados de Conservação e Desenvolvimento (PICDs)” (Struhsaker *et al.*, 2005). Em resposta, Ferraro & Pattanayak (2006) argumentam que “intervenções como os PICDs e programas de educação conservacionista não são alocados aleatoriamente através da paisagem. Intervenções com base comunitária são mais propensas de serem adotadas

<sup>1</sup> Não há tradução literal ao português. Stakeholders são o conjunto de pessoas e entidades interessadas – direta e indiretamente – num determinado projeto, ou seja, aqueles que estão envolvidos, são afetados ou atingidos pelo resultado final de um projeto.

<sup>2</sup> Não há tradução literal. Policy makers são aqueles responsáveis por planejar políticas públicas.

<sup>3</sup> Sem tradução literal, significa uma linha base, de referência.

em áreas que estão enfrentando fortes pressões humanas. Portanto, comparando-se resultados médios da conservação em áreas onde as intervenções beneficiam comunidades locais (sob forte pressão) com resultados médios em áreas onde há poucas intervenções de tal tipo (sob pouca pressão), gera-se uma estimativa enviesada (para baixo) da influência da conservação em tentativas de beneficiar moradores vizinhos a áreas protegidas”.

### **Florestania no Acre**

Para demonstrar a presença muito comum de efeitos de confusão, consideramos dois programas do estado do Acre. Escolhemos o Acre não porque ele é de qualquer forma melhor ou pior em programas de monitoramento e avaliação, mas porque a gestão do Governador Jorge Viana decidiu fazer do estado um exemplo de desenvolvimento com base florestal, com estratégias que incluem certificação florestal, áreas protegidas de uso múltiplo e um subsídio estadual para a produção de borracha, que é tratado como um pagamento por serviços ambientais. Essas iniciativas atraíram interesse, abrangendo desde estados vizinhos considerando programas similares, até a cobertura em noticiários internacionais (p.ex., Rohter, 2002). Portanto, é evidente a importância de levar em consideração a forma como se poderia avaliar a eficácia desses programas, incluindo a sua contribuição para o objetivo duplo de conservação florestal e bem-estar rural no estado.

### **Programa de Desenvolvimento Sustentável do Acre (PDSA)**

A gestão Viana adotou Governo da Floresta como seu slogan, e florestania – serviços públicos e infra-estrutura melhorados para os cidadãos da floresta – como uma política chave. Essa infra-estrutura inclui estradas, com a gestão afirmando que Viana “já entrou para a história do Acre como o governador que mais asfaltou e que mais se preocupou com estradas no Estado” (SEFE, 2004). Embora estradas sejam geralmente um anátema para a comunidade conservacionista devido a sua óbvia correlação com desmatamento, ainda há um debate aberto sobre a direção de causalidade – a questão das “estradas endógenas”, os efeitos do melhoramento de estradas em fronteiras antigas e o potencial de combinar e concatenar adequadamente estradas com outros serviços públicos e fiscalização ambiental (Weinhold & Reis, 2003). Esse potencial é a base para o PDSA, financiado pelo Banco Interamericano de Desenvolvimento (BID).

O PDSA busca melhorar a qualidade de vida no estado e preservar sua riqueza natural a longo prazo. Para alcançar esse objetivo duplo, o governo estadual propôs pavimentar a BR-364 e, simultaneamente, estabelecer um ambiente político e institucional que encorajasse o manejo florestal sustentável. A afirmação é de que com políticas apropriadas estabelecidas, a estrada facilitará o desenvolvimento com base florestal através do melhoramento do acesso a mercados para produtos florestais e a serviços públicos aos moradores rurais tradicionais. As políticas adequadas incluem serviços públicos (educação, saúde, informação sobre mercados, extensão e treinamento), crédito acessível, aplicação da legislação ambiental e suporte a mecanismos que compensem os proprietários das florestas por serviços ecológicos (p.ex., a certificação de produtos florestais). Representantes do governo do Acre argumentam que a afirmação é particularmente provável de ser favorável aos moradores tradicionais das florestas, cujas preferências e capital (humano, social, físico e natural) os predispõe ao desenvolvimento de base florestal. Esses grupos têm uma longa história de organização social e ativismo no estado.

Com os melhoramentos da estrada já em vias de execução, o governo estabeleceu vários mecanismos para o monitoramento de impactos ao longo da estrada. De acordo com o documento de empréstimo, os indicadores a serem monitorados incluem (a) contagem de domicílios pobres “beneficiados”, (b) área de “pastos recuperados”, (c) aumento de renda nos domicílios beneficiados e (d) aumento na produção agroflorestal. Embora esse monitoramento ‘antes-depois’ seja crítico para prestação de contas, ele tem várias limitações em relação ao aprendizado oriundo do projeto. Primeiramente, na área do projeto todos são afetados tanto pelo melhoramento da estrada como pelo resto do “pacote de políticas públicas” e, portanto, não será possível separar tais influências. Em segundo lugar, como mencionado acima, o Acre tem uma história única de organização social, que certamente foi um fator importante na obtenção do empréstimo do BID, e provavelmente afetará o resultado do PDSA: um problema clássico de seleção endógena. Em terceiro lugar, vários indicadores listados no documento de empréstimo medem insumos em vez de resultados. Como insumos são geralmente mais fáceis de serem medidos, haverá uma tentação em focar-se neles em detrimento dos indicadores de resultado de conservação e desenvolvimento, que são mais difíceis porém mais necessários de serem medidos.

### Lei Chico Mendes

Durante a primeira gestão de Viana, a agência estadual Secretaria Executiva de Floresta e Extrativismo (SEFE) foi responsável por desenvolver alternativas econômicas compatíveis com a conservação para melhorar o bem-estar de populações rurais tradicionais. O ponto central dos esforços da SEFE foi o Programa da Borracha, ou Programa de Desenvolvimento da Cadeia Produtiva da Borracha, e no cerne desse programa estava um subsídio governamental à borracha, também conhecido por Lei Chico Mendes. O objetivo imediato do subsídio é melhorar os padrões de vida dos seringueiros o que, por sua vez, espera-se que reduza o desmatamento por encorajá-los tanto a permanecerem na floresta, prevenindo a invasão de fazendeiros e criadores de gado migrantes, como a se concentrarem na borracha, em vez de expandirem suas atividades agrícolas.

O estado argumenta que o subsídio da borracha é justificado como um pagamento por serviços ambientais. A maior parte da produção de borracha no Acre ocorre em florestas nativas, e as poucas plantações de seringueira no estado também fornecem serviços ambientais, como proteção de bacias hidrográficas e seqüestro de carbono, embora com biodiversidade significativamente menor. A idéia de que esses serviços sejam produzidos conjuntamente a produtos florestais não-madeireiros não é exclusiva ao Acre. Por exemplo, a Coordenadoria de Agroextrativismo do Ministério do Meio Ambiente afirma que um de seus objetivos é o crescente reconhecimento dos serviços ambientais fornecidos por comunidades tradicionais. No caso da borracha, pode haver inclusive benefícios à conservação *in situ* de diversidade genética, que também não é refletida no preço de mercado (Davis, 1997).

A maior crítica à Lei Chico Mendes é que ela subsidia uma atividade que não é mais viável economicamente. A maioria dos economistas argumentaria que é mais eficiente pagar diretamente pelos serviços ambientais desejados (p.ex., Ferraro, 2001). Da perspectiva do governo, a dificuldade encontra-se na elaboração de um mecanismo para selecionar, monitorar e pagar moradores rurais, através de todo o estado, pela conservação de florestas. Um desafio é assegurar que os pagamentos alcancem os seringueiros, já que um dos objetivos-chave é aumentar o seu padrão de vida. Um segundo desafio é assegurar que os recebedores dos pagamentos estejam de fato conservando a floresta. O subsídio da borracha pode ser interpretado como um mecanismo de incentivo prático que se dirige a ambos desafios. A não ser que o subsídio seja substancialmente aumentado, é improvável a atração de novas pessoas

para a atividade seringueira, ou o encorajamento de novas plantações de borracha. Portanto, o subsídio é pelo menos, em certa medida, direcionado a si próprio: somente seringueiros (os beneficiários almejados) extraem a borracha, e autofiscalizável: a produção de borracha requer uma floresta em pé.

Baseando-nos em registros da SEFE, sabemos que entre 1999 e 2002 o governo estadual subsidiou mais de 6,5 milhões de quilos de borracha em nome de 6.750 seringueiros em 37 diferentes associações, a um valor de R\$0,40 por quilo (posteriormente aumentado para R\$0,60 e R\$0,70). Baseando-nos em estimativas estaduais da população de seringueiros (ZEE, 2000), o subsídio alcançou mais que um terço das famílias de seringueiros do Acre. Dependendo dos pressupostos e períodos de tempo, o subsídio estadual gerou ao redor de R\$12,00 a R\$20,00 por família por mês em média, ou 16 – 25% do salário mínimo oficial. Todos esses são indicadores de insumos do programa. As questões mais fundamentais de avaliação são: quão mais altas são as rendas sob o subsídio da borracha e qual é o impacto que isso possui sobre a floresta.

Para avaliar esses impactos, dados sobre os recebedores do subsídio da borracha podem ser combinados com pesquisas junto a moradores rurais do Acre na execução de uma análise 'com-sem' subsídio. Por exemplo, dados de uma pesquisa de 2002 demonstram que a quantidade possuída de gado e a venda de produtos pecuários (carne e leite) estão altamente correlacionados à participação no programa de subsídio da borracha (Sills & Saha, 2004). Mas será que isso significa que os seringueiros que haviam começado a investir em gado são direcionados de volta à extração de borracha pelo programa (um resultado ambiental positivo), ou que a renda incrementada pela borracha está sendo investida em gado (um resultado ambiental negativo)? Novamente, esse é um problema de seleção endógena, com o resultado de interesse (gado) possivelmente afetando a participação no programa. Na verdade, nós sabemos que os seringueiros são mais propensos a participar do programa se eles tiverem uma família grande e se estiverem explorando um número grande de seringueiras, mas não reivindicam uma área grande de terra (Sills & Saha, 2004). Isso parece sensato: aqueles com maiores probabilidades de beneficiar-se do programa são mais propensos a participar.

Para ilustrar os potenciais vieses resultantes da auto-seleção, assumamos que os seringueiros que têm mais a ganhar do programa também têm inicialmente menores rendas e mais florestas. E assumamos que renda e cobertura florestal são indicadores dos resultados al-

mejados pelo programa. Nesse caso, numa avaliação baseada na comparação desses indicadores ‘antes-depois’ da participação do programa, o impacto na renda seria provavelmente superestimado e o impacto na cobertura florestal subestimado. Por outro lado, numa comparação ‘com e sem’ programa (uma comparação de participantes e não-participantes), o efeito sobre a renda seria provavelmente subestimado e o efeito sobre a cobertura florestal superestimado.

## MÉTODOS DE AVALIAÇÃO DE PROGRAMAS

O campo da avaliação de programas fornece ferramentas que isolam os efeitos causais de projetos, programas e políticas, controlando efeitos de confusão, como a seleção endógena. Embora a avaliação de programas tenha sido aplicada a uma ampla variedade de intervenções relacionadas a políticas públicas (p. ex., redução de pobreza, capacitação para o trabalho, reabilitação de criminosos, saúde pública), ela raramente tem sido utilizada na avaliação de políticas ambientais (Benneer & Coglianese, 2005). Ferraro & Pattanayak (2006) sugerem razões pelas quais esses métodos não têm sido aplicados no campo da conservação e argumentam que isso representa uma oportunidade perdida de aprender da experiência e de direcionar, de melhor forma, os limitados recursos da conservação.

### Experimentos naturais

Uma estratégia de avaliação de programas é aproveitar os conhecidos ‘experimentos naturais’ – situações em que natureza ou acaso criaram grupos ‘tratamento’ e ‘controle’, p.ex., participantes e não-participantes no programa sendo avaliado (Rosenzweig & Wolpin, 2000). Ferraro & Pattanayak (2006) sugerem que desastres naturais podem “criar barreiras naturais que protegem ou expõem florestas a pressões sociais de forma diferente. Ambos lados das barreiras naturais fornecem locais comparáveis para avaliações. A natureza, em vez de pessoas, seleciona as unidades com base no acaso, e assim elimina o viés na seleção”. Mais frequentemente, algum elemento aleatório de política governamental cria um experimento. Por exemplo, Wood & Walker (2004) afirmam que políticas emergenciais de regularização fundiária na região do Araguaia-Tocantins resultaram em uma alocação aleatória de títulos de terra para alguns fazendeiros e não a outros, permitindo-lhes inferir sobre o impacto do *status* de posse sobre o uso da terra. Olinto *et al.* (2003) e Morris *et al.* (2004) estudaram o impacto da Bolsa Alimentação na nutrição,

através da comparação entre participantes do programa e pessoas elegíveis a participar, mas que foram acidentalmente excluídas das listas devido a erros no banco de dados dos computadores.

Fronteiras políticas e administrativas são outro tipo de experimento natural. O empréstimo do BID para o PDSA fornece ao Acre meios de melhorar o acesso a estradas e, simultaneamente, implementar políticas que tragam benefícios, regulamentação, crédito e outros mercados à região. Embora estados vizinhos possam seguir políticas similares no longo prazo, o empréstimo criará um experimento natural no curto prazo: tanto o Acre como o estado vizinho do Amazonas serão impactados pela estrada, mas somente o Acre beneficiar-se-á de outras políticas elaboradas para promover o manejo florestal sustentável. Além disso, no Acre as atividades do programa serão direcionadas a áreas específicas, permitindo comparações entre municípios dentro do estado. Com um desenho amostral cuidadoso, uma pesquisa poderia incluir domicílios que apresentam diferenças no acesso a mercados e em outras políticas, nas escalas temporal e espacial, criando uma oportunidade de identificar os efeitos do ‘tratamento’ somente com o melhoramento do acesso à estrada, e com o ‘pacote modernizador’ mais completo proposto pelo PDSA.

A desvantagem de experimentos naturais é que eles raramente envolvem uma aleatorização perfeita. Por exemplo, os mesmos fatores que auxiliaram o Acre a receber o empréstimo do BID – histórico de organização social das populações dependentes da floresta, apoio popular a políticas de conservação de florestas – também poderiam influenciar os resultados. Em termos técnicos, domicílios são heterogêneos tanto em características observáveis (p. ex., capital natural) como em não-observáveis (p. ex., preferências por tipo de trabalho). Essa heterogeneidade pode estar relacionada aos limites municipais e estaduais, sendo o ‘tratamento’ neste caso não distribuído aleatoriamente através das famílias (ou comunidades) e, portanto, sujeito ao viés de seleção. Controlar tais diferenças requer outros passos: medir as condições de *baseline* (p. ex., desmatamento e bem-estar), examinar covariáveis relevantes nos domicílios (outros fatores que afetem o desmatamento e o bem-estar, e fatores que influenciem o alcance da política) e aplicar métodos estatísticos para construir grupos controle.

### Baselines, covariáveis, e grupos controle

*Baselines*, mensuração de covariáveis e grupos controle são os três métodos interdependentes básicos para lidar com efeitos de confusão em qualquer pesquisa cien-

tífica (Ferraro & Pattanayak, 2006). *Baselines* medem condições e comportamentos anteriores à intervenção, e assim controlam condições iniciais que poderiam afetar as medidas de eficácia do programa. Covariáveis são fatores observáveis que também influenciam a medida de resultado; esses fatores podem ser socioeconômicos, biofísicos, econômicos ou institucionais. Grupos controle são indivíduos, comunidades ou áreas não submetidos à intervenção, mas que em outros aspectos são similares (em geral). Esses métodos são aplicáveis mesmo que um experimento natural não esteja disponível. Para organizar as idéias, discutimos a Lei Chico Mendes.

Como descrito anteriormente, ao combinar registros administrativos do subsídio da borracha no Acre com dados de levantamentos domiciliares, podemos fazer comparações entre domicílios beneficiados ou não pelo subsídio. Como a participação do programa é voluntária, necessitamos informações adicionais para construir um grupo controle que seja válido. Tais informações poderiam ser uma *baseline*, p. ex., informações sobre o indicador de resultado relativo à posse de gado antes do subsídio da borracha. No entanto, ainda não seríamos capazes de identificar se os domicílios participantes foram motivados por (ou focados devido a) fatores não observados específicos. Portanto, poderíamos utilizar informações sobre as características dos domicílios que são determinantes exógenas da participação do programa de subsídio e de resultado. Alguns estudos tentam aproximar-se desse desenho amostral ao perguntar questões retrospectivas (p. ex., Walker *et al.*, 2000; McCracken *et al.*, 2002), porém a confiabilidade de dados recordados não é bem fundamentada (Rindfuss *et al.*, 2003).

Outra estratégia de obtenção de *baselines* e covariáveis é definir a unidade de análise – como um município ou cidade – e usar dados secundários de levantamentos, como a Pesquisa Nacional por Amostragem de Domicílios (PNAD). Obviamente, tal estratégia requer que o resultado e as covariáveis de interesse sejam registrados no levantamento. Por exemplo, Pianto & Soares (2004) avaliam o impacto do Programa de Erradicação do Trabalho Infantil (PETI) utilizando dados da PNAD e aproveitando-se do fato de que o PETI foi executado em fases, em municípios distintos em anos diferentes. Clark *et al.* (2004), da mesma forma, utilizam dados secundários para avaliar o impacto da privatização da água e de sistemas de saneamento no Paraná. Em ambos casos, os autores têm acesso aos resultados médios municipais (p. ex., número de crianças trabalhando em atividades degradantes ou perigosas, número de domicílios com água encanada) antes e depois da interven-

ção política. Eles utilizam informações sobre os níveis *baseline* dos resultados e covariáveis relevantes para identificar grupos controle e construir o controle experimental *ex post facto*, empregando os métodos estatísticos que nós descrevemos adiante.

### Análise estatística

As técnicas estatísticas – ou econométricas – utilizadas em avaliação de programas permitem que relações causais hipotéticas entre as atividades do programa e os resultados almejados sejam testadas. Essas técnicas levam em conta tanto o cenário observado como cenários hipotéticos, ou seja, tanto o efeito do programa de intervenção sobre os participantes como o que ocorreria aos mesmos participantes em situações hipotéticas ‘sem intervenção’. Nós focamos em métodos de “emparelhamento” (*matching*), assim chamados por combinar (ou, emparelhar) áreas (ou domicílios) de um programa com áreas (ou domicílios) “muito similares” que não participam do programa. Esses pares não-participantes servem como um controle experimental *ex post facto*. Antes de nos voltarmos ao emparelhamento, revisamos primeiramente um método econométrico mais convencional: Variáveis Instrumentais (VI). Baker (2000) fornece uma visão geral e não técnica de ambos métodos no contexto de avaliação de programas.

O método da Variável Instrumental depende de fatores exógenos que influenciam a intervenção, mas não o resultado da conservação. É similar ao ‘experimento natural’ pois o ‘instrumento’ (como a ‘natureza’) quebra a relação circular entre a intervenção e o resultado. Por exemplo, no exemplo discutido em Experimentos Naturais, os limites estaduais ou municipais são instrumentos eficazes que são correlacionados à intervenção, mas não necessariamente ao resultado. Uma das mais bem conhecidas aplicações da abordagem VI a recursos naturais é um estudo sobre como o manejo florestal comunitário afeta a coleta de lenha no Nepal. Edmonds (2002) utiliza a presença de programas de extensão e postos de fiscalização florestal para explicar a formação de grupos de manejo locais. Ele demonstra que essas variáveis não estão relacionadas à coleta de lenha e, portanto, são instrumentos válidos. Ele então testa se esses grupos de manejo influenciam a coleta de lenha. A abordagem ‘VI’ e sua semelhante ‘função controle’ são freqüentemente utilizadas em programas sociais, por exemplo, em estimativas de regressos ao sistema educacional (Card, 2001; Heckman & Navarro-Lozano, 2004). No entanto, como Ferraro & Pattanayak (2006) advertem, “em geral, boas variáveis instrumentais são difíceis de serem encontradas. O uso

da abordagem VI tipicamente requer uma mescla de intuição teórica clara, dados secundários de boa qualidade e uma sólida compreensão das condições de campo”.

O objetivo dos métodos de emparelhamento (*matching*) é identificar um grupo de comparação que seja “muito similar” ao grupo submetido à intervenção (Heckman *et al.*, 1997). Na verdade, idealmente haveria somente uma diferença chave: o grupo de comparação não teria participado do programa. Considere o caso do subsídio da borracha no Acre. O método de emparelhamento pressupõe que a participação do programa pode ser explicada segundo as características observáveis dos participantes. Conseqüentemente, o analista pode combinar cada domicílio que recebe o subsídio com o domicílio equivalente, ou seja, um domicílio com as mesmas características, mas que não tenha recebido o subsídio. Os domicílios em cada par combinado têm idênticas probabilidades de participar do programa da borracha. O impacto do subsídio da borracha nos participantes – o efeito causal do programa da borracha – pode ser calculado como a diferença média dos resultados entre os domicílios subsidiados e os não subsidiados, em cada par combinado.

Um dos melhores e mais utilizados métodos de emparelhamento é o de emparelhamento por nota de propensão (*propensity score matching*). Ele aborda um problema prático do emparelhamento: quando o número de características utilizadas na combinação é aumentado para melhorá-la, torna-se impossível encontrar um par exato para cada participante. Rosenbaum & Rubin (1983) demonstraram que, em vez de parear participantes e não-participantes em todas as características observadas, poder-se-ia parear baseando-se na probabilidade de participação. A probabilidade é a nota de propensão (*propensity score*), tipicamente estimada de um modelo estatístico de participação em função de fatores ecológicos, socioeconômicos, institucionais e geográficos (Ferraro & Pattanayak, 2006). Uma característica importante da abordagem de emparelhamento é que ela é não-paramétrica e assim evita os pressupostos sobre distribuição e as restrições de modelos de regressão que estimam resultados como função de características observadas.

Métodos de emparelhamento têm sido amplamente utilizados em países em desenvolvimento na identificação e estimativa de impactos causais de políticas sociais. Estudos publicados investigam o impacto de

programas escolares na educação infantil (Behrman *et al.*, 2005), transferências de dinheiro sobre o crescimento de crianças (Morris *et al.*, 2004), fornecimento de água encanada sobre diarreia infantil (Jalan & Ravallion, 2003a), fundos de investimento sociais sobre infraestrutura pública e saúde (Rawlings *et al.*, 2004) e de um programa antipobreza sobre a renda (Jalan & Ravallion, 2003b). Por exemplo, Morris *et al.* (2004) e Olinto *et al.* (2003) empregaram métodos de emparelhamento para demonstrar que, nas primeiras fases, o programa Bolsa Alimentação resultou em um consumo de alimentos maior e mais diverso, mas em menor ganho de peso entre crianças. Eles sugerem que isso possa ser devido ao fato de que o acesso a programas governamentais anteriores (ou seja, o Incentivo para o Combate de Carências Nutricionais) havia sido condicionado a crianças com peso abaixo do normal. Pianto & Soares (2004) também utilizam métodos de emparelhamento para avaliar o impacto do PETI. Eles descobriram que o programa realmente diminuiu o trabalho infantil. Renda e matrícula escolar também aumentaram em municípios onde o PETI havia sido implementado, mas quando esses municípios foram comparados com municípios similares mas não beneficiados pelo PETI, não havia diferença. Em ambos casos, métodos de emparelhamento desvendaram os efeitos de importantes programas sociais, identificando relações causais de resultados de interesse.

No campo dos recursos naturais, estamos somente conscientes de alguns poucos trabalhos que demonstram o uso do método de emparelhamento por nota de propensão para avaliar os impactos de distúrbios florestais sobre serviços florestais (Pattanayak, 2004), da queima preventiva sobre incêndios descontrolados (Butry *et al.*, 2005), da gestão descentralizada sobre a cobertura florestal na Índia (Somanathan *et al.*, 2005) e do *Endangered Species Act*<sup>4</sup> sobre a recuperação de espécies (Ferraro *et al.*, 2005). Em um estudo em andamento, estamos examinando proprietários de terras com floresta na Costa Rica utilizando vários métodos de amostragem para encontrar os melhores pares potenciais entre participantes e não-participantes do programa de pagamentos por serviços ambientais, e estamos estimando o impacto dos pagamentos sobre a cobertura florestal (Sills *et al.*, 2005). Edmonds (2002) e Bui Dung (2005) apresentam outras formas de emparelhamento

<sup>4</sup> Lei americana de proteção às espécies ameaçadas.

para avaliar o efeito de grupos de usuários da floresta sobre a coleta de lenha no Nepal, e o de pagamentos em dinheiro sobre a adoção de práticas de exploração florestal sustentável no Vietnã.

## EXPERIMENTOS DE POLÍTICAS PÚBLICAS

Todas as sofisticadas ferramentas da econometria descritas anteriormente essencialmente compensam a falta de controles experimentais, o que sugere que uma abordagem mais simples seja a implementação de experimentos de campo nos quais intervenções sejam designadas aleatoriamente através de indivíduos, comunidades ou regiões (Greenberg *et al.*, 2003). Se feito de maneira correta, o desenho experimental assegura que potenciais fatores que possam gerar confusão sejam equilibrados através de unidades controle e unidades sob intervenção, sendo que assim quaisquer diferenças nos resultados entre as duas possam ser atribuídas à intervenção. No campo do desenvolvimento econômico, tem havido experimentos que testam os efeitos do micro-crédito sobre o bem-estar familiar e da desparasitação infantil sobre a performance escolar. Profissionais da conservação poderiam implementar experimentos aleatorizados para examinar a eficácia de intervenções, como pagamentos por serviços ambientais ou educação conservacionista.

Portanto, por que não se executa somente experimentos reais de políticas públicas? Afinal de contas, a maioria dos projetos e programas exclui alguns beneficiários potenciais. Se essa exclusão fosse aleatória e os dados fossem coletados sobre os excluídos, haveria evidência direta e inequívoca dos impactos. Mas experimentos são difíceis, pois a localização de intervenções reflete demandas sociais e cálculos políticos. Alguns dos desafios de avaliações experimentais incluem (a) a questão ética e política ao negar-se a intervenção no grupo 'controle'; (b) dificuldade na definição de um grupo controle caso o experimento seja numa escala nacional ou ecoregional; (c) contaminação cruzada do grupo controle ou tratamento devido ao movimento de pessoas ou espécies; e (d) recursos insuficientes para considerar todas as permutações e combinações potenciais de políticas públicas. Portanto, os métodos quase-experimentais e econométricos de avaliação de programas podem representar a abordagem mais viável. Como argumentamos anteriormente, eles podem ser utilizados para testar e identificar rigorosamente quais intervenções são mais eficazes, especialmente com a coleta de dados *baseline* em uma amostra cuidadosamente planejada, que inclua tanto grupos tratamento como controle.

## SÍNTESE

Neste artigo, introduzimos métodos de avaliação quase-experimentais ilustrados por exemplos de políticas conservacionistas no estado do Acre. Esses métodos de avaliação de programas abrangem desenho experimental, coleta de dados e análise estatística. Não há uma solução genial que funcionará em todas situações; por exemplo, experimentos naturais ou boas variáveis instrumentais estarão disponíveis somente em alguns casos. Como descrito por Ferraro & Pattanayak (2006), profissionais da conservação poderiam fazer melhoramentos crescentes em avaliação, primeiro, simplesmente coletando dados sobre covariáveis; segundo, se os recursos permitirem, coletando dados adicionais sobre grupos controle; e terceiro, caso seja possível o planejamento antecipado, coletando dados *baseline*. Os recursos dedicados à avaliação de programas claramente irão (e devem) variar através de diferentes intervenções conservacionistas. Nossa preocupação é a de que há pouca ou nenhuma aplicação atual à conservação da biodiversidade, dessa forma frustrando a nossa capacidade de identificar, planejar e justificar intervenções eficazes. Isso é mais crítico para programas novos em suas fases piloto, como o Pró-Ambiente, ou gestões como a do Governo da Floresta no Acre, que busca estabelecer um exemplo para o resto da região. No entanto, qualquer projeto conservacionista que assimile os métodos e medições revisados nesse artigo farão uma vital contribuição em direção ao preenchimento da grande lacuna no nosso conhecimento sobre os mais eficazes investimentos em conservação.

## AGRADECIMENTOS

Gostaríamos de agradecer o apoio à pesquisa no Acre à Fundação Ford, ao IMAZON, ao governo do Acre, e à Universidade Estadual de Carolina do Norte (EUA). Shubhayu Saha contribuiu para a análise e Gardênia de Oliveira Salas e Maria Lúcia R. Santos coletaram dados sobre o Programa de Borracha. A pesquisa recebeu apoio também da Fundação Gordon e Betty Moore, através do Centro para Pesquisa Aplicada à Biodiversidade.

## REFERÊNCIAS BIBLIOGRÁFICAS

Baker J. 2000. Evaluating the impact of development projects on poverty: handbook for practitioners. World Bank. Disponível em: <http://web.worldbank.org/WBSITE/EXTERNAL/TOPICS/EXTPOVERTY/EXTISPMA/0,,contentMDK:20194198~>

- pagePK:148956~piPK:216618~theSitePK:384329,00.html (acessado em janeiro de 2006).
- Behrman, J., P. Sengupta & P. Todd. 2005. Progressing through progress: an impact assessment of a school subsidy experiment. *Economic Development and Cultural Change* 54: 237-275.
- Benbear, L. & C. Coglianese. 2005. Measuring progress: Program evaluation of environmental policies. *Environment* 47(2): 22-39.
- Bui Dung, T. 2005. Payments for environmental services in Vietnam: a trial of sustainable forest management. EEPSEA Final Report.
- Butry, D., S. Pattanayak, E. Sills & D. Mercer. 2006. Applying program evaluation methods to natural resource policy: are wildfire suppression expenditures worth it? Presented at the AERE session of the ASSA Meeting, Chicago, EUA.
- Card, D. 2001. Estimating the return to schooling: progress on some persistent econometric problems. *Econometrica* 69(5): 1127-1160.
- Christensen, J. 2003. Auditing conservation in an age of accountability. *Conservation in Practice* 4(3): 12-19.
- Clark, R., K. Katrina & S. Wallsten. 2004. Has private participation in water and sewerage improved coverage? Empirical evidence from Latin America. World Bank Policy Research Working Paper N°3445. World Bank, Development Research Group, Washington, D.C.
- CMP (Conservation Measures Partnership). 2004. Open standards for the practice of conservation. Disponível em: <http://www.conservationmeasures.org/CMP/> (acessado em janeiro de 2006).
- Davis, W. 1997. The Rubber industry's biological nightmare. *Fortune Magazine* 86.
- Edmonds, E. 2002. Government initiated community resource management and local resource extraction from Nepal's forests. *Journal of Development Economics* 68(2): 89-115.
- Fearnside, P. 2005. Desmatamento na Amazônia brasileira: história, índices e conseqüências. *Megadiversidade* 1(1).
- Ferraro, P. 2001. Global habitat protection: limitations on development interventions and a role for conservation performance payments. *Conservation Biology* 15(4): 990-1000.
- Ferraro, P., C. McIntosh & M. Ospina. 2005. The effectiveness of listing under the U.S. endangered species act: an econometric analysis using matching methods. Association of Environmental and Resource Economists Workshop "Natural Resources at Risk," Grand Teton National Park, WY. 6/05. Audubon Society, Atlanta, EUA.
- Ferraro, P. & S. Pattanayak. 2006. Money for nothing? A call for evidence-based practice in biodiversity conservation. Working Paper. RTI International, Research Triangle Park, North Carolina, EUA.
- Gasper, D. 2000. Evaluating the logical framework approach towards learning-oriented development evaluation. *Public Administration and Development* 20 (1):17-28.
- Greenberg D., D. Links & M. Mandell. 2003. Social experimentation and public policymaking. Urban Institute Press, Washington, D.C.
- Kainer, K.A., M. Schmink, A.C.P. Leite & M.J. da Silva Fadell. 2003. Experiments in forest-based development in western Amazonia. *Society and Natural Resources* 16: 869-886.
- Hailey, J., & M. Sorgenfrei. 2003. Measuring success. *Issues in performance management. ITRAC Occasional Papers Series* N° 44.
- Heckman, J., H. Ichimura & P. Todd. 1997. Matching as an econometric evaluation estimator: evidence from evaluating a job training programme. *Review of Economic Studies* 64: 605-654.
- Heckman, J. & S. Navarro-Lozano. 2004. Using matching, instrumental variables, and control functions to estimate economic choice models. *The Review of Economics and Statistics* 86(1): 30-57.
- Jalan, J. & M. Ravallion. 2003a. Does piped water reduce diarrhea for children in rural India? *Journal of Econometrics* 111(1): 153-173.
- Jalan, J. & M. Ravallion. 2003b. Estimating the benefit incidence of an anti-poverty program by propensity-score matching. *Journal of Business and Economic Statistics* 21(1): 19-30.
- Kleijn D. & W.J. Sutherland. 2003. How effective are European agri-environment schemes in conserving and promoting biodiversity? *Journal of Applied Economic Letters* 40: 947-969.
- Kleiman D.G., R.P. Reading, B.J. Miller, T.W. Clark & J.M. Scott. 2000. Improving the evaluation of conservation programs. *Conservation Biology* 14(2): 356-365.
- Margolis, R. & N. Salafsky. 1998. Measures of success-designing, managing and monitoring conservation and development projects. Island Press, Washington, D.C.
- McCracken, S., A.D. Siqueria, E.F. Moran & E.S. Brondizio. 2002. Land-use patterns on an agricultural frontier in Brazil: insights and examples from a demographic perspective. In: Wood, C. & R. Porro (eds.). *Deforestation and land use in the Amazon*, University Press of Florida, Gainesville, EUA.
- Morris, S., P. Olinto, R. Flores, E. Nilson & A. Figueiro. 2004. Conditional cash transfers are associated with a small reduction in the rate of weight gain of preschool children in northeast Brazil. *Journal of Nutrition* 134: 2336-2341.
- Olinto, P., R. Flores, S. Morris & A. Veiga. 2003. The impact of the bolsa alimentação program on food consumption. Mimeo. Disponível no banco de dados de avaliação de pobreza do Banco Mundial: <http://www1.worldbank.org/prem/poverty/ie/evaluationdb.cfm>
- Pattanayak, S.K. 2004. Forest amenities and aesthetics: an econometric evaluation using North Carolina FIA Data. Documento não publicado. RTI International, Research Triangle Park, EUA.
- Pianto, D.M. & S. Soares. 2004. Use of survey design for the evaluation of social programs: The PNAD and the program for the eradication of child labor in Brazil. Mimeo. Disponível no banco de dados de avaliação de pobreza do Banco Mundial: <http://www1.worldbank.org/prem/poverty/ie/evaluationdb.cfm>
- Rawlings, L.B., L. Sherburne-Benz & J.V. Domelen. 2004. Evaluating social funds: a cross-county analysis of community investments. The World Bank, Washington, D.C.
- Reid, J. & W. Sousa Jr. 2005. Investimentos em infra-estrutura e políticas de conservação no Brasil. *Megadiversidade* 1(1): 189-197.
- Rindfuss, R., S. Walsh, V. Mishra, J. Fox & G. Dolcemascolo. 2003. Linking household and remotely sensed data: methodological

- and practical problems. In: Fox, J., R.R. Rindfuss, S.J. Walsh, & V. Mishra (eds.). *People and the environment: approaches for linking household and community surveys to remote sensing and GIS*. Kluwer Academic Publishing Group, Dordrecht.
- Rohter, L. 2002. Discovering Amazon rain forest's Silver Lining. *The New York Times Section A*: 8.
- Rosenbaum, P.R. & D.B. Rubin. 1983. The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1): 41-55.
- Rosenzweig M.R. & K.I. Wolpin. 2000. Natural "natural experiments" in economics. *Journal of Economic Literature* 38: 827-874.
- Sanderson, S. 2002. The future of conservation. *Foreign Affairs* 81: 162-173.
- SEFE (Secretaria Executiva de Florestas e Extrativismo). 2004. Governo reabre a BR-364 pelo sexto ano consecutivo. Disponível em: [http://www.sefe.ac.gov.br/secom/noticias/maio2004/n01\\_12maio2004.htm](http://www.sefe.ac.gov.br/secom/noticias/maio2004/n01_12maio2004.htm) (acessado em janeiro de 2006).
- Sills, E. & S. Saha. 2004. Rubber subsidies for extractive reserves: sign of failure or success? presented at the Yale ISTF Conference People and Parks, New Haven, EUA.
- Sills, E., S. Pattanayak, P. Ferraro, R. Arriagada, L. Carrasco, K. Caldwell & S. Cordero. 2005. Costa Rica's payments for environmental services: using matched samples and linked surveys and remote sensing to evaluate program effectiveness. Interim Status Report.
- Somanathan, E., R. Prabhakar & B.S. Mehta. 2005. Does decentralization work? Forest conservation in the Himalayas. Discussion Paper 05-04. Indian Statistical Institute, Planning Unit, Delhi, India.
- Struhsaker, T.T., P.J. Struhsaker & K.S. Siex. 2005. Conserving Africa's rain forests: problems in protected areas and possible solutions. *Biological Conservation* 123(1): 45-54.
- Walker, R., E. Moran & L. Anselin. 2000. Deforestation and cattle ranching in the Brazilian Amazon: external capital and household processes. *World Development* 28 (4): 683-699.
- Weinhold, D. & E. Reis. 2003. Land use and transportation costs in the Brazilian Amazon. Documento de trabalho.
- Wood, C. & R. Walker. 2004. Land titles, tenure security and resource use among small farmers in the Brazilian Amazon. Documento de trabalho.
- ZEE. 2000. Zoneamento ecológico-econômico do Acre. SECTMA, Rio Branco, AC.