Optimal Environmental Targeting in the Amazon Rainforest*

Juliano Assunção[†], Robert McMillan[‡], Joshua Murphy[§], and Eduardo Souza-Rodrigues[¶]

January 30, 2020

*We would like to thank Fernanda Brollo, Sylvain Chabe-Ferret, Francisco Costa, Clarissa Gandour, Kelsey Jack, Charles Manski, Ismael Mourifié, Marcel Oestreich, Martino Pelli, Alex Pfaff, Alberto Salvo, Paul Scott, Edson Severini, Kate Sims, and Aloysius Siow for helpful discussions, and workshop participants at Brock University, the Instituto Escolhas, the 2018 LSE-NHH Conference, the Montreal Workshop in Environmental and Resource Economics, the 2018 NBER Summer Institute, the INRA Environmental and Natural Resources Conservation Workshop in Montpellier, Toulouse School of Economics, the University of Ottawa, the University of Toronto, and the 6th World Congress of Environmental and Resource Economists for additional comments. Faisal Ibrahim provided outstanding research assistance. Financial support from the University of Toronto Mississauga is gratefully acknowledged. Assunção is gratefully thankful for financial support from the CNPq. All remaining errors are our own. The views expressed herein are those of the authors and do not necessarily reflect the views of their respective organizations.

[†]Departamento de Economia, PUC-Rio, Rua Marques de Sao Vicente, 225/F210, 22453-900 - Rio de Janeiro/RJ, Brazil, email: juliano@econ.puc-rio.br

[‡]Department of Economics, University of Toronto, Max Gluskin House, 150 St. George Street, Toronto, ON, Canada, and NBER, email: mcmillan@chass.utoronto.ca, tel. 416 978 4190.

[§]Natural Resources Canada, 580 Booth Street, Ottawa, ON K1A 0E4, Canada, email: joshua.murphy4@canada.ca

[¶]Department of Economics, University of Toronto, Max Gluskin House, 150 St. George Street, Toronto, ON, Canada, email: edusouzarod@gmail.com

Abstract

This paper sets out an empirically-driven approach for targeting environmental policies optimally in order to combat deforestation. We focus on the Amazon, the world's most extensive rainforest, where Brazil's federal government issued a 'Priority List' of municipalities in 2008, to be targeted with more intense environmental monitoring and enforcement. In this setting, we first estimate the causal impact of the Priority List on deforestation using 'changes-in-changes' (Athey and Imbens, 2006), a flexible treatment effects estimation method, finding that it reduced deforestation by 40 percent and cut emissions by 39.5 million tons of carbon. Second, we develop a novel framework for computing targeted ex-post optimal blacklists. This involves a procedure for assigning municipalities to a counterfactual list that minimizes total deforestation subject to realistic resource constraints, drawing on the ex-post treatment effect estimates from the first part of the analysis. Accounting for spillovers, we show that the ex-post optimal list resulted in carbon emissions over 7.4 percent lower than the actual list, amounting to savings of more than $900 million, and emissions over 25 percent lower (on average) than a randomly selected list. The approach we propose is relevant for assessing both targeted counterfactual policies to reduce deforestation and quantifying the impacts of policy targeting more generally.

Keywords: Policy Targeting, Optimal Regulation, Monitoring, Deforestation, Amazon, Carbon Emissions, Changes-in-Changes, Difference-in-Differences, Spillovers, Resource Constraints, Partial Identification, Minimax Ambiguity

1 Introduction

In many developing countries, weak institutions undercut the effective implementation of environmental policies, as recent research has documented clearly.[1] The unregulated, often illegal activities that prevail can cause severe duress to fragile ecosystems, producing outcomes that are both damaging and inefficient. As a prominent instance of this phenomenon, several studies (notably by Burgess et al., 2012) have highlighted the role of illegal logging and land clearing as a driving force behind tropical deforestation, widely understood to be a critical contributor to global carbon emissions (see IPCC, 2013). In settings such as these where existing institutions are over-stretched, targeted monitoring and enforcement policies may be advantageous, helping to focus limited resources where they can have higher-than-average impacts.

This paper measures the causal effects of blacklist-type government regulations – a widespread form of targeting – and then explores how such targeted regulations can be optimized. It does so in the context of deforestation, focusing on the Amazon, the world's most extensive rainforest and a vitally important ecosystem, whose fundamental roles in storing carbon, conserving biodiversity, maintaining water quality and even modulating the Earth's climate are well established (Foley et al., 2005; Stern, 2007; Bonan, 2008; Davidson et al., 2012). Deforestation in the Amazon has been a source of international concern for at least the past 30 years, spurring increased regulatory activity, especially on the part of Brazil's federal government. The regulations introduced in Brazil coincided with a marked slowdown in deforestation, the annual deforested area falling by 75 percent between 2004 and 2017. As other factors may be responsible for this decline (changing commodity prices among them), policy makers in Brazil and elsewhere are keenly interested in knowing how effective actual regulations have been in reducing deforestation, and how such regulations might be further refined. Yet the literature has not supplied a means to assess, in a systematic quantitative way, which policy configurations would be likely to have most impact in limiting future deforestation given relevant constraints: filling that gap is the central task of this paper.

Our analysis is built around an important regulatory change that occurred in 2008, when Brazil's federal government issued a blacklist of 36 municipalities (out of a total of 526) with especially high levels of deforestation – the so-called 'Priority List.' The listed municipalities were to be subject to more rigorous monitoring and stricter penalties, with the list being renewed every year subsequently.

The paper's first goal is to estimate the causal treatment effect of the Priority List on deforestation levels in the Brazilian Amazon. Given the official criteria did not specify exactly how the list was chosen, we start by investigating the effective selection rule that assigned municipalities to the Priority List. The patterns we find in the data indicate that the federal government adhered

[1] See Greenstone and Jack (2015) for a thorough review of the issues involved.

closely to a threshold rule, essentially separating municipalities based on their deforestation levels but not on their trends.[2] Indeed, we cannot reject the common trends assumption, comparing municipalities on the list (versus not) leading up to its introduction in 2008.

In considering the short-run impact of the reform over the period 2006–2010, one could estimate a standard difference-in-differences (DID) model based on that evidence. Yet in the current context, heterogeneous treatment effects are likely to be present, with the Priority List being implemented on the group with potentially higher average benefits when compared to the control group.[3] Given such heterogeneous effects, a DID strategy can only identify treatment effects on the treated (ATT), which is insufficient when trying to shed light on optimal targeting – the second goal of this paper. For the targeting exercise, we need to estimate the policy impacts on the untreated.[4]

To that end, we adopt the changes-in-changes (CIC) model proposed by Athey and Imbens (2006) (henceforth 'A&I'), which provides a nonlinear generalization of the DID model to the entire distribution of potential outcomes. In a policy evaluation context with pre- and post-policy periods, A&I show how the difference in the distribution functions of the untreated group before and after treatment can be combined with the distribution function of the treated group before treatment to predict the hypothetical distribution of the treated group in the post-treatment period, absent treatment. (In standard DID, the adjustments are to the average, not to the entire distribution function, and are implemented linearly.) Similarly, the counterfactual distribution function of the effects of treatment on the untreated can be recovered. As the two counterfactual distributions can be arbitrarily different, treatment effects are allowed to be heterogeneous across units (municipalities in our application) *and* across treatment and control groups.

In terms of the main treatment effect results, we find that the Priority List caused substantial reductions in the deforestation rate, cutting it by 40 percent in the short term (the period 2009–2010) relative to the case in which no program was enacted. This reduction led to avoided emissions of 30 million tons of carbon, with a social benefit of around \$2.2 billion, assuming a social cost of carbon of \$20/tCO_2 (Greenstone et al., 2013; Nordhaus, 2014).[5] Further, there is evidence of heterogeneous treatment effects, with the average effect on the untreated (ATU) being between 10

[2] Using only the threshold rule, we are able to replicate the actual 2008 assignments with 97 percent accuracy. The use of this threshold rule also suggests that econometric strategies commonly employed in the program evaluation literature – propensity score, matching, regression discontinuity, instrumental variables – to estimate treatment effects may not be applicable in our setting. See Appendix C.

[3] The official criteria to enter the Priority List reflect the assumption that deforestation is a persistent process: highly deforested locations in the past are expected to be more likely to be deforested in the future, so concentrating regulatory effort in highly deforested areas may result in more substantial reductions in total deforestation.

[4] We note that extrapolating results from the treated group to the untreated under the assumption of homogeneous effects would bias the estimated effects on the untreated and make any ex-post policy calculations unreliable.

[5] This is a conservative lower bound. Using the EPA's recommended current social cost of carbon estimate would double the estimated social benefit.

and 14 percent of the estimated effect on the treated.[6]

We also investigate the possibility that the Priority List generated *spillovers*. Farmers in untreated municipalities geographically close to a Priority municipality and which experienced substantial deforestation in the past might think that monitoring could also increase there.Accordingly, we split the untreated group in two (denoted the 'spillover' and 'control' groups), depending on whether untreated municipalities were more or less likely to react to the policy intervention. Estimates of the CIC model provide evidence of spillover effects, with the spillover group reducing deforestation in response to the intervention: the treatment effect for this group is smaller than the effect on the treated, but greater than the effect on the control group. Once we account for spillovers, 2,705 km^2 of deforestation were avoided in Priority municipalities directly, while the indirect impact discouraged the clearing of 618 km^2 of forested area in the spillover group, totalling 3,323 km^2 of forested area preserved in 2009–2010 as a result of the program. The total avoided emissions amounted to 39.5 million tons of carbon, with a social benefit of approximately $2.9 billion.

The paper's second goal (referenced above) is to look beyond the actual policy and compare the Priority List with an ex-post optimal blacklist. To this end, we develop a framework for exploring the assignment of municipalities to an optimized counterfactual list based on information about ex-post treatment effects drawn from the first part of the analysis. The framework allows us to investigate in a systematic way how knowledge of treatment effects – perhaps only partial in nature – can lead to better-targeted conservation policies.

We suppose the federal policy maker assigns municipalities to a counterfactual list with the objective of minimizing either total deforestation or total carbon emissions – a variety of other social objectives can be accommodated by the approach. The policy maker's decision is analyzed as a treatment choice problem under ambiguity – appropriate given that some treatment effects are not point-identified – and we use the minimax criterion, assuming the policy maker chooses the ex-post list in order to achieve the best of the worst outcomes (see the seminal contributions of Manski, 2000, 2004).[7] Further, to incorporate limited monitoring resources into the minimization problem, we consider two alternative constraints, one restricting the total area that can be monitored, and the other, the total number of municipalities on the list.[8]

Accounting for spillover effects, we show that the Priority List resulted in carbon emissions

[6] Although data limitations prevent us from point-identifying the treatment on the untreated, the estimated effect on the untreated is partially identified with informatively narrow identified sets.

[7] Manski (2005a) provides an accessible overview.

[8] We set the constraints at the same values as those corresponding to the Priority List; we also investigate the effects of relaxing these constraints. Information about the resources that were effectively allocated to monitoring is difficult, if not impossible, to obtain. Nevertheless, it is reasonable to presume that the larger the area or the number of municipalities monitored, the higher the monitoring costs.

that were *at least* 8 percent higher than the ex-post optimal lists (under either constraint), while randomly selected lists of municipalities would result in emissions that were over 34 percent higher on average. The avoided emissions translate into a lower bound for the social value of the optimal list of approximately $900 million over the period 2009–2010. As these counterfactual gains derive from the treatment effect estimates, they imply high social returns to investments in conservation policy research.

The geographic distributions of the ex-post optimal lists reveal several interesting patterns that were not imposed during the course of the estimation. First, the overlap between protected areas and the area-constrained counterfactual list is much lower than the overlap between protected areas and the original Priority List. This suggests these two policies can be made to work together in ways that could be further leveraged by the Brazilian government.[9] Second, ignoring spillover effects, the area-constrained counterfactual list is contiguous and forms a protective shield close to the deforestation frontier, which (together with protected areas) may help impede the deforestation process from continuing into more pristine areas, with benefits in the longer term. Third, when accounting for spillovers, the area-constrained optimal list becomes more geographically dispersed and less contiguous; intuitively, placing all targeted municipalities together does not exploit the potential reduction in deforestation in adjacent locations due to spillovers.

Beyond the current application, the approach we develop is relevant for assessing counterfactual targeted policies to reduce deforestation in other contexts, based around actual policy interventions. Those interventions can be used to recover heterogeneous policy impacts, our approach then allowing researchers to trace out the quantitative implications for forest cover and carbon emissions when policy makers face realistic resource constraints and only partially identified estimates. It also provides a coherent framework for assessing the quantitative impacts of policy targeting more generally, as we discuss below, using credible estimates based on a flexible treatment effects estimation approach.

The rest of the paper is organized as follows: The next section places our analysis in the context of the literature. Section 3 sets out relevant institutional background; Section 4 describes the data, along with descriptive evidence motivating the empirical model, introduced in Section 5; Section 6 presents the empirical results, including the average treatment effects; Section 7 develops our counterfactual framework and shows results from the counterfactual targeting exercises, and Section 8 concludes.[10]

[9] A similar implication can be drawn from evidence that compares protected area policies and payments for ecological services from Mexico (see Alix-Garcia et al., 2015).

[10] The Appendix supplements the main text with information about the data sources and the construction of key variables, several robustness excercises, and a detailed explanation of how the counterfactual optimal lists are calculated in practice.

2 Relation to the Literature

Our paper contributes first to a growing body of work examining environmental policy implementation in developing countries (see Greenstone and Jack (2015) for a recent survey) – a complement to the vast literature studying environmental policies in a developed country context.[11] Greenstone and Hanna (2014) argue that weak institutional arrangements in developing countries pose obstacles to effective law enforcement, showing that policies targeting improvements in air and water quality in India had varying degrees of success. In the case of climate issues, linked to the deforestation process analyzed in this paper, the available evidence is limited (Burke et al., 2016). Our analysis examines a widespread form of targeting and connects the causal impacts of targeting policies to the release of carbon into the atmosphere, in these ways contributing to both lines of research.[12]

Second, several papers examine the impact of monitoring and the role of institutions in the Amazon itself, notably Hargrave and Kis-Katos (2013), Assunção et al. (2017), and Burgess et al. (2017); payments for ecological services programs have also been studied as alternatives to command-and-control policies by Pattanyak et al. (2010), Alix-Garcia et al. (2012, 2015), Jayachandran et al. (2017), Jack and Jayachandran (2018), and Simonet et al. (2019). Compared with these papers, our analysis examines the effectiveness of an optimized counterfactual policy-targeting strategy as a way of overcoming institutional and political obstacles. This type of targeted strategy can be applied in other contexts that encompass a substantial portion of global rainforest cover – in other parts of Amazonia, the Congo and Southeast Asia.[13]

Within the Amazon context, recent papers have examined the effects of the Priority List, including Arima et al. (2014), Cisneros et al. (2015), Andrade and Chagas (2016), Harding et al. (2018), Koch et al. (2018), and Assunção and Rocha (2019). Those studies use difference-in-differences and matching methods to obtain average treatment effects similar in magnitude to the corresponding estimates in our study. Our estimation approach also allows us to recover the effects of treatment on the untreated, which we use in computing optimally targeted blacklists.

A third main strand of literature investigates the underlying causes of land use change, including tropical deforestation, due to changes in population, infrastructure, agricultural prices, political

[11] See Gray and Shimshack (2011) for a survey.

[12] Our approach also complements an earlier theoretical literature in environmental economics studying targeted regulatory strategies – see Harrington (1988) and Friesen (2003). We show how a regulator can target resources in an optimal way subject to realistic constraints using credible treatment effect estimates. Our approach allows the associated benefits to be quantified directly on the basis of econometric evidence.

[13] Jack and Jayachandran (2018) consider how targeting could be incorporated into the design of payments for ecological services (through manipulation of enrollment costs) to improve the cost-effectiveness of such programs. Targeting social programs in contexts other than environmental conservation is of broad interest; see, e.g., Hanna and Olken (2018) for a comprehensive discussion in the context of cash transfer programs to reduce poverty.

economy factors and climate-related phenomena.[14] Our results indicate that monitoring policies are important drivers of land use change and deforestation, affecting not only the municipalities that are targeted directly but also generating spillovers for neighboring areas.

Fourth, from an estimation standpoint, the flexible CIC model has not been used widely to date. It is implemented by Havnes and Mogstad (2015) in their study of child care in Norway when carrying out robustness checks, by Kottelenberg and Lehrer (2017) to assess targeted versus universal childcare, and in other supplementary analyses – see Athey and Imbens (2017). Our analysis based on ex-post treatment effects is (to the best of our knowledge) the first time the approach has been used in the environmental or regulation literatures. Beyond these existing studies, the CIC method is applicable in a variety of important settings, especially when providing policy-relevant estimates that can be used counterfactually, as we show.

Fifth, our counterfactual analysis draws on a burgeoning literature studying statistical treatment rules in econometrics, including Manski (2004, 2005a), Stoye (2009), Hirano and Porter (2009), Bhattacharya and Dupas (2012), Kasy (2016), and Kitagawa and Tetenov (2018), among others. (We provide a more detailed discussion of this literature in Appendix A.) Few empirical applications have appeared thus far, aside from empirical illustrations presented in some of the existing methodological papers.[15] In an applied econometric context, our analysis is novel (as explained in the appendix) in that there is no study in which all the following hold simultaneously: (a) unconfoundedness assumptions fail so that the treatment effects and the welfare objective function are partially identified; (b) the estimation of treatment effects accounts for violations of the 'Stable Unit Treatment Value Assumption' (or SUTVA); (c) the treatment choice is made under ambiguity (and also allows for spillover effects, again in violation of SUTVA); and (d) the set of admissible policies must satisfy binding capacity constraints.

3 Institutional Background and Regulations

In this section, we describe relevant background, especially relating to the institutional context – the legal environment, the introduction of satellite monitoring in 2004, and our main focus: the Priority List, introduced in 2008.

Our setting is the Brazilian Amazon, which accounts for two-thirds of the Amazon Rainforest and is itself a vast area, almost ten times the size of California. Prior to the 1960s, the forest was barely occupied; access was open, and local economic activities were based largely on subsistence

[14] See papers by Stavins (1999), Pfaff (1999), Andersen et al. (2002), Lubowski et al. (2006), Brady and Irwin (2011), Cisneros et al. (2013), Mason and Platinga (2013), Pailler (2018), and Souza-Rodrigues (2019).

[15] One important empirical study we are aware of is the analysis by Dehejia (2005), who examines the Greater Avenues for Independence (GAIN) program that began in California in 1986.

and extraction activities, primarily involving rubber and Brazil nuts.[16] During the 1960s and 1970s, the occupation of the Amazon was promoted by the military dictatorship with the explicit goals of securing national borders and developing the region,[17] although government investment was then cut in the 1980s due to economic recession and hyperinflation. In the late-1980s, ecological concerns started to shape policies in the Amazon. Notably, IBAMA (the Brazilian Environmental Protection Agency) was created in 1989, given power to execute environmental policies, and serving as the national police authority concerned with the investigation and sanctioning of environmental infractions.

The Legal Environment. Approximately half of the Amazon was under legal protection by 2010 – either indigenous lands or conservation units such as national parks, extractive reserves, and areas of ecological interest. Deforestation in those areas is subject to strict requirements. The rest of the Amazon comprises undesignated public land where no deforestation is allowed, or private land – approximately 20 percent of the total area (according to the Agricultural Census of 2006) – where deforestation has to follow the rules of the Forest Code. This code states that, among other requirements, farms in the Amazon must preserve 80 percent of their area in the form of native vegetation. While deforestation on private land can be legal if it is both authorized and accords with the Forest Code, empirical evidence suggests that compliance with the Forest Code is limited (Michalski et al., 2010; Borner et al., 2014; Godar et al., 2014); and although some deforested areas captured in our data may have been cleared legally, most deforestation in the Amazon is illegal.

In terms of environmental monitoring, IBAMA's operations in the Amazon up to the mid-2000s were based largely on information collected and processed by IBAMA's headquarters and regional offices. Although land and air patrols were used in the 1990s and early-2000s, they were limited in their effectiveness given the sheer extent of the area covered and risks posed to law enforcers.

Satellite-Based Monitoring. The adoption of satellite-based monitoring from the mid–2000s improved patrolling capabilities significantly. The first stage began in 2004, with the launch of the Action Plan for the Prevention and Control of Deforestation in the Legal Amazon (PPCDAm), setting out new procedures for monitoring and environmental control.[18]

Central to PPCDAm law enforcement has been the use of high-frequency remote sensing technology in the form of a satellite-based system, DETER, developed by the Brazilian Institute for Space Research (INPE). This increased the capacity to monitor forest-clearing activities in the

[16] See Souza-Rodrigues (2019) for more detail.

[17] Hydroelectric facilities, mines, ports, and around 60,000 km of roads were constructed during this period.

[18] The PPCDAm also led to the expansion of protected areas, mostly during its first phase (spanning 2004–2007), before the first municipalities were assigned to the Priority List in 2008.

Amazon in a significant way, processing land use images on a frequent basis, detecting areas experiencing a loss of forest cover, and in turn triggering DETER deforestation alerts for the immediate attention of law enforcers. Since being introduced in the mid-2000s, DETER has served as the primary tool for IBAMA's monitoring efforts in the Amazon. As Assunção et al. (2017) show, this satellite-based system alone has had an important impact: estimated deforestation in the absence of the system would have been more than 3.6 times greater.

The Priority List. In 2008, the government launched the second phase of the PPCDAm, the main component of which involved the creation of a blacklist to better target regulatory effort in order to combat illegal deforestation. Any Amazon municipality could be added to what became known as the 'Municípios Prioritários' List (for convenience, the 'Priority List'). Municipality-level selection criteria for this list were based on (a) total deforested area, (b) total deforested area over the past three years, and (c) the increase in the deforestation rate in at least three of the past five years, although the exact rules followed are not in the public domain and so have to be inferred.[19]

Municipalities on the Priority List were subjected to more intense environmental monitoring and law enforcement, with IBAMA devoting a greater share of its resources to them (MMA, 2009). Fines were also increased in Priority municipalities, which became subject to a series of further administrative measures that imposed additional costs to being blacklisted.[20]

The Ministry of the Environment's Ordinance 28, issued in January 2008, listed 36 municipalities making up the initial Priority List – 7 percent of the total number of municipalities in the Brazilian Amazon. The original list was expanded to include an additional seven municipalities in 2009. A further six were placed on the list in 2011, followed by two more in 2012. By then, just six municipalities had been removed from the list (one in 2010, another in 2011, and four in 2012). There were no changes 2013 and 2017 (when eight new municipalities entered the list). In total, 59 municipalities were eventually placed on the Priority List between 2008 and 2017, while 467 municipalities did not enter the list during the same period.

[19]The legal basis for targeting certain municipalities was set out in the Presidential Decree 6,321 in December 2007. Exiting the Priority List depended on reducing deforestation in a significant way and having at least 80 percent of the municipal private area registered in the Rural Environmental Registry system.

[20]These included more stringent conditions applying to the approval of subsidized credit contracts, and the requirement to develop local plans for sustainable production (see Brito et al., 2010; Maia et al., 2011; Arima et al., 2014). Private land titles were also revised in a bid to identify fraudulent documentation and illegal occupancy, and licensing requirements were made stricter for rural establishments.

4 Data and Descriptive Evidence

We have assembled a municipality-year panel data set that combines information about Priority status, land use (including the location of public protected areas), and other possible determinants of deforestation.[21] Our analysis focuses on the time period 2006–2010, with the pre-treatment period covering 2006 and 2007, and the post-treatment period, 2009–2010.[22] The official list of Priority municipalities comes from the Ministry of the Environment. The treatment group comprises municipalities that entered and remained on the list from 2008 to 2010 inclusive (with the exception of one municipality that exited in 2010). The control group consists of the set of municipalities that did not enter the list before 2010.[23]

The main variables of interest are listed in Table 1 – sample statistics in the table (discussed below) are provided for 2007, the last year prior to the policy's introduction. Our municipality-year panel includes annual measures of the forested area remaining, cumulative deforestation, and incremental deforestation in each municipality, drawn from the Brazilian government's satellite-based forest monitoring program, PRODES. Factors affecting deforestation other than the Priority List include rainfall, temperature, protected areas, prices of beef and crops, and local gross domestic product. We have also assembled data on crop area, the number of cattle, deforestation alerts and fines issued, and measures of above-ground carbon stock. We end up with a balanced panel of 490 (out of a possible 526) municipalities within the Amazon Biome, as there are few instances of missing data.

Aggregate Trends and Geography. Around 20 percent of the Brazilian Amazon has been deforested to date – an area totalling over 700,000 square kilometres, which is larger than Texas. Cleared areas are used mainly for agriculture: approximately two thirds of the deforested area comprises pasture, and around 8 percent is used for crops – see Almeida et al. (2016).[24]

[21] Further information about data sources and the construction of key variables is provided in Appendix B.

[22] Because of a structural break in 2004–2005 associated with the first phase of PPCDAm, comparing deforestation before the first phase of PPCDAm and after the implementation of the Priority List in the pre-treatment period would capture the combined effect of both regulatory changes.

[23] Because there are few municipalities entering and exiting the Priority List from 2009 on, there is not much that can be said with any accuracy about the impact of the policy in these cases. Given more data, it would be possible to estimate causal effects that varied depending on the length of exposure to the program and whether the policy affected municipalities at different times. In the current case, standard DID exploiting variation across groups of units receiving treatment at different times would estimate a weighted sum of different average treatment effects (Goodman-Bacon, 2018; de Chaisemartin and D'Haultfoeuille, 2018), which would bias our analysis of the counterfactual optimal list. See Abbring and Heckman (2007) for a thorough discussion of dynamic treatment effects, Han (2019) for recent identification results, and Callaway and Sant'Anna (2018) for implications relating to treatment effect estimation in the context of DID models.

[24] Almeida et al. (2016) also show that 20 percent of the cleared area currently takes the form of secondary vegetation. The remaining areas correspond to mining, urban areas, 'other,' and 'unobserved' (i.e., areas whose land usage cannot be interpreted due to cloud cover or smoke from recent forest burning).

Figure 1 presents aggregate deforestation trends. This reveals two pronounced downward steps coinciding the main phases of the PPCDAm, in 2004 and 2008. It is clear that deforestation fell considerably in 2004 and the years immediately following, and again after 2008, with the rate stabilizing subsequently. In total, annual deforestation declined by approximately 75 percent over the period 2004–2017.[25]

Figure 2 presents initial evidence relating to contributing factors, showing the evolution of deforestation levels together with the international prices of soybeans and beef. The figure suggests a positive correlation between deforestation and prices prior to 2008, consistent with the fact that most of the deforested area in the Brazilian Amazon is used for pasture (grazing cattle being reared mainly for beef) and crops (mostly soybeans and corn). After 2008, the correlation appears to be much weaker, suggesting that the Priority List may have helped preserve the rainforest even when international prices were rising.

The location of the municipalities on the Priority List within the Amazon is shown in Figure 3, with Priority municipalities being found mostly in the Amazon's southern and eastern regions – an area known as the "Arc of Deforestation." Figure 4 shows where the incremental deforestation occurred each year between 2006 and 2010, and also presents the cumulated deforestation by 2010 (together with Priority municipalities overlaid). These figures make clear that new deforestation is a persistent process. In such circumstances, a targeted policy may be effective, concentrating monitoring and enforcement in locations where deforestation is more likely to occur.

Selection onto the Priority List. The Priority status of a municipality depends on the three official selection criteria noted above: the total amount of forested land cleared in municipality m from its inception up to and including year $t-1$ (labelled Z^1_{mt-1}); the amount of forested land cleared in municipality m in the three-year period ending in year $t-1$ (Z^2_{mt-1}); and an indicator for whether municipality m experienced year-on-year growth in new deforestation at least three times in the five-year period ending with year $t-1$ (Z^3_{mt-1}). Of these, the first two selection criteria relate to long-run and more recent deforestation, while the third relates to whether deforestation accelerated in recent years.

Under the assumption that these variables fully determine Priority status, the selection equation can be written:

$$G_{mt} = g\left(Z^1_{mt-1}, Z^2_{mt-1}, Z^3_{mt-1}\right), \tag{1}$$

[25] Total incremental deforestation by year is shown in Table 13 in Appendix G, together with the number of fines issued, the expansion of protected areas, and the number of municipalities added to the Priority List. Incremental deforestation does not incorporate potential forest clearing in unobserved/clouded areas, while the official aggregate deforestation rates include estimates of deforestation in unobserved areas, based on local extrapolations; see Appendix B. This distinction reconciles the profiles in Figures 1, 7 and 8 below exactly.

where $G_{mt} \in \{0, 1\}$ indicates whether municipality m is on the Priority List in year t. Given that the precise rules determining selection are not stated publicly, we seek to infer them by exploring whether the vector summarizing the three criteria, $Z_{mt-1} \equiv \left(Z^1_{mt-1}, Z^2_{mt-1}, Z^3_{mt-1}\right)$, determines Priority status fully. To that end, Figure 5 plots all combinations of Z^1_{mt-1} and Z^2_{mt-1} for a given value Z^3_{mt-1}: the scatterplot in panel (a) holds $Z^3_{mt-1} = 0$, and the scatterplot in panel (b) holds $Z^3_{mt-1} = 1$. In both panels, municipality-year observations with $G_{mt} = 0$ (not on the list) and $G_{mt} = 1$ (on the list) are marked with crosses and dots, respectively.

From the two panels, it is clear that regulators adhered closely to a threshold rule involving the first and second criteria: both Z^1_{m2007} and Z^2_{m2007} had to cross pre-determined thresholds in order for municipality m to qualify for the Priority List, while the third criterion (whether deforestation accelerated in recent years) is not important. It is possible to define threshold values for the first and second selection criteria (indicated by vertical and horizontal lines in each panel) that almost completely separate Priority municipalities from non-Priority municipalities. Specifically, the thresholds drawn in both panels of Figure 5 are 2,700 km^2 for Z^1_{mt-1} and 220 km^2 for Z^2_{mt-1}. Using only these inferred thresholds, we are able to replicate the actual 2008 assignments with 97 percent accuracy. This is highly suggestive that a strict threshold selection rule is followed in practice.

One important consequence is that factors such as local political influence are unlikely to lead to manipulation close to the relevant thresholds determining the Priority List's initial composition. This is perhaps surprising, given evidence that corruption is an important and widespread problem in Brazil, with documented consequences for deforestation – see, for example, Cisneros et al. (2013) who find that mayors caught engaging in corrupt behavior allow more deforestation. A careful study by Pailler (2018) finds evidence that deforestation rates increase 8–10 percent on average in election years when an incumbent mayor runs for re-election (noting that, in our setting, 2008 was a mayoral election year). Yet she does not find significant effects in the years leading up to or following the election year, suggesting that re-election incentives have not affected deforestation differentially in Priority and non-Priority municipalities. We find that the fraction of municipalities in which the mayor is affiliated with the political coalition of the Brazilian president is the same among Priority and non-Priority municipalities (approximately 40 percent in each group), suggesting the policy was not used as punishment against political enemies at the local level.[26]

The empirical form taken by the selection function $g(\cdot)$ has important implications for the viability of several widely used identification strategies. Because there is very little overlap in the data among Priority and non-Priority groups given Z_{mt-1}, selection-on-observables techniques

[26] We are grateful to Fernanda Brollo for generously providing the data on political coalitions. This finding contrasts with evidence on the use of federal transfers documented by Brollo and Nannicini (2012).

(matching or propensity scores) are problematic in this context. Further, the use of a regression discontinuity (RD) design is limited by the fact that there are few observations close to the threshold frontier (in addition to which an RD does not identify the policy treatment effect of interest in this paper); and while the criteria variables in Z_{mt-1} might seem to be natural instruments for Priority status, they are invalid when the unobservables affecting deforestation decisions are serially correlated. (We discuss these points in more detail in Appendix C.) These issues motivate the use of difference-in-differences and changes-in-changes approaches in order to estimate causal impacts of the policy on deforestation.

Fines and Penalties. The purpose of the Priority List was to focus monitoring and enforcement efforts on municipalities with high levels of deforestation, and presumably where further deforestation was most likely. To get a sense of whether the environmental police were more active in Priority municipalities, we compare the extent of deforested areas and the number of fines issued by IBAMA before and after 2008. Figure 6 plots fines as a function of contemporaneous deforestation, separately for municipalities in treated and untreated groups.[27] The evidence suggests that the Priority List led to more intense enforcement in Priority municipalities relative to non-Priority municipalities.[28] This is consistent with the results of Assunção and Rocha (2019), who find evidence suggesting that law enforcement is the main channel through which the policy affected deforestation. We investigate these channels further in Section 6.

Comparing Treated versus Untreated Municipalities. Next, we compare Priority and non-Priority municipalities descriptively. First, we consider summary statistics for a 'baseline' cross section from 2007, right before the policy's introduction and considered separately by Priority status, presented in Table 1. As expected given the graphical evidence relating to the selection criteria, the two groups differ in important ways. In Priority municipalities, incremental deforestation and total historical deforestation – the first selection criterion considered by the Ministry of the Environment when assigning Priority status – are higher. Priority municipalities are also larger and have higher local agricultural GDP, higher carbon stocks per hectare, and are subject to more stringent policy measures.

[27] The points are non-parametric predictions from local linear regressions that use a rectangular kernel and a bandwidth of 25 square kilometres.

[28] Prior to 2008, the number of fines issued for a given level of deforestation did not differ by Priority status (panels (a)–(c)). As soon as the Priority List was introduced, a clear upward shift occurs in the number of fines issued for a given amount of deforestation in Priority municipalities, while no shift is apparent in non-Priority municipalities (panel (d)). (We note that between 2006 and 2007, the number of fines increased in both treatment and control groups and deforestation fell, relative to previous years, likely reflecting the first phase of the government's plan to control deforestation in the Amazon. There is no discernible targeting of municipalities with historically high rates of deforestation, however.)

Beyond baseline differences, aggregate trends comparing the two groups are informative. Panel (a) of Figure 7 compares the evolution of deforestation among treated and untreated municipalities in levels. Differences in deforestation levels are apparent, but for both groups, new deforestation fell after 2005, and increased slightly in 2006–2008. There are no signs of any anticipation effects.

We also compare the evolution of the log odds ratios among the two groups in panel (b) of Figure 7, given that the outcome variable in our empirical framework is the log odds ratio of deforestation shares (which we estimate from a logit model, as explained below in Section 5). The same pattern emerges, with differences in levels but similar movements before 2008. This aggregate evidence suggests that the selection rule effectively separated municipalities based on their deforestation levels, not on their trends (consistent with the evidence that the third selection criterion, Z^3_{mt-1}, capturing acceleration in deforestation, does not help predict Priority status). As further corroboration, we cannot reject the common trends assumption based on the log odds ratio before treatment (shown in Section 6). Further, although deforestation slowed down in both treatment and control groups after 2008, the aggregate slowdown among Priority municipalities was more marked, providing initial evidence that deforestation may have responded to the blacklist policy.

Spillover Effects. Next we consider the possibility that the Priority List generated spillover effects, working in two distinct ways. First, by concentrating monitoring in areas where a disproportionate amount of deforestation occurred (so-called 'hot spots'), the intervention might simply shift, rather than reduce, total deforestation.[29] The extent to which deforestation could relocate geographically (a problem known as 'leakage') depends on how costly it is to move and then deforest in other areas. Such costs make it unlikely that such leakage would be important in the short run (although it may be important in the longer term). Indeed, supporting this view, Figure 4 shows no clear evidence that new deforestation was accumulating after 2008 in non-Priority municipalities that were close by the municipalities placed on the list.[30]

A second potential spillover effect can work in the opposite direction: farmers in untreated municipalities may deforest less if they expect the intervention to increase monitoring in non-targeted locations.[31] Indeed, Figures 4 and 7 suggest that deforestation declined in both treated and untreated municipalities following the treatment.

[29] This relates to the literature on criminal deterrence, and more specifically to the impact of 'hot-spots' policing – see Chalfin and McCrary (2017) for an excellent review of that literature.

[30] This strongly suggests that leakage is not a first-order issue for the time period covered in the data. The empirical results presented in Section 6 are also consistent with this view.

[31] In the 'hot spots' policing literature, the majority of studies find no evidence of the displacement of crime to adjacent neighborhoods, and a substantial number of the studies have found instead a tendency for crimes to fall in non-treated adjacent locations (Chalfin and McCrary, 2017).

To investigate whether such deterring spillovers may be present, we split the untreated group in two, depending on whether untreated municipalities are more or less likely to react to the policy intervention. Specifically, we consider two plausible conditions for designating 'spillover' municipalities: (i) whether a municipality shares a border with a treated municipality (i.e., adjacent locations), and (ii) whether a municipality has high levels of deforestation historically. Formally, we define our second condition for splitting the untreated group (those with 'high levels of historical deforestation') on the basis of the threshold criteria that were (implicitly) adopted by the Brazilian government, shown in Figure 5.[32] We call the group of untreated municipalities satisfying both conditions – being a neighbor of a Priority municipality and having high levels of past deforestation – the 'spillover' group. The summary statistics in Table 2 confirm that the spillover group falls between the treated and control groups in virtually all instances. In turn, Figure 8 compares the evolution of deforestation among the three groups. Panel (a) presents deforestation in levels, while panel (b) shows the log odds ratio of deforestation shares. The spillover group features deforestation levels between the other two groups, while its log odds ratios are slightly above those of the treated group. Again, the evolution profiles are similar, especially after 2005. Of note, while deforestation slowed down in the three groups after 2008, the slowdown among the spillover municipalities is not as pronounced as that observed among Priority units, but it is more prominent than among municipalities in the control group, which suggests the presence of spillover effects in the current context.[33]

5 Empirical Framework

In this section, we set out a framework that underlies our approach to studying targeted environmental regulations in the Amazon. We start, as a benchmark, with standard difference-in-differences (DID), then describe the more general changes-in-changes (CIC) model proposed by A&I. In the process, we lay out our empirical strategy and the parameters of interest – the average treatment effects.

Our empirical approach is shaped by particular data constraints. Given that we do not observe the land use decisions of individual farmers, but rather have land use panel data at the munici-

[32] Given the threshold criteria from the figure, we split the untreated group depending on whether Z^1_{mt-1} and Z^2_{mt-1} exceed 70 percent of the thresholds – that is, whether $Z^1_{mt-1} \geq 0.7 \times 2,700$ km^2 and $Z^2_{mt-1} \geq 0.7 \times 220$ km^2. The empirical results presented in Section 6 are robust to diffferent definitions of how close past deforestation is to the threshold criteria. (See Appendix F.)

[33] It is worth mentioning that municipalities with deforestation levels near the selection threshold criteria and that do not have a neighbor treated may also react to the Priority List, in anticipation of possibly stricter monitoring in future. As there are only 13 municipalities satisfying this condition, we cannot split the untreated group further to investigate this case with any degree of accuracy.

pal level, we focus on municipal-level deforestation and treat this as a function of the regulatory environment, among other factors (commodity prices, local climatic conditions etc.). On the policing side, we have only limited information about the intensity of monitoring, so we use a binary measure of treatment – assignment to the Priority List – and follow a treatment effects approach, given that modeling the decisions of individual farmers and regulators directly at the micro-level is not feasible. These data constraints notwithstanding, our empirical approach allows us to obtain causal treatment effects based on aggregate data and credible policy variation, described next.

5.1 Difference-in-Differences

We make use of the standard potential outcomes notation in describing the empirical approach, with capital letters denoting random variables, and lower case letters denoting corresponding realized values. Each municipality m belongs to a group $G_m \in \{0, 1\}$, where group 0 is the control group and group 1 is the treatment group – extensions to more than two groups are straightforward. Let A_{mt} denote the total forested area in municipality m at the beginning of year t, and let D_{mt} be the amount of deforestation that occurred in m during the same year. The share of newly deforested area Y_{mt} is the ratio of D_{mt} to A_{mt}. We use superscript $j \in \{0, 1\}$ to indicate the potential outcome that arises under the policy regime j. The observed share of deforestation for municipality m at time t can then be written:

$$Y_{mt} = (1 - G_m) \times Y^0_{mt} + G_m \times Y^1_{mt}.$$

We adopt a logistic regression framework, as is common in the empirical land use literature (Stavins, 1999; Pfaff, 1999; Souza-Rodrigues, 2019). In the standard DID model, the regression formulation is given by:

$$\log\left(\frac{Y_{mt}}{1 - Y_{mt}}\right) = X'_{mt}\beta + \delta_t + \tau_1 (G_m \times \delta_{2009}) + \tau_2 (G_m \times \delta_{2010}) + \alpha_m + \eta_{mt}, \tag{2}$$

where X_{mt} is a municipality-level vector of observed factors, including prices and agro-climatic conditions (see Section 4); δ_t are time dummies; α_m is a municipality-level fixed effect; η_{mt} is a time-varying unobservable factor; and (β, τ_1, τ_2) are the parameters to be estimated. The parameters τ_1 and τ_2 equal the average treatment effect (in terms of the log odds ratio of deforestation shares) among the treated municipalities during the first and the second year of the program, respectively, thus allowing for time-varying treatment effects. Given that the Priority List should reduce deforestation, one would expect $\tau_1 \leq 0$ and $\tau_2 \leq 0$. The parameters can be estimated consistently based on (2) provided that the common trends assumption holds – we provide formal

econometric evidence below.

The logistic model is appealing, both conceptually and given its measurement properties, in this context. It can be motivated based on a continuum of farmers who make binary choices (to deforest or not), aggregated up to the municipality level; as such, one can trace the share of deforestation back to underlying individual decisions, which is helpful in interpreting the empirical results.[34] From a measurement perspective, in contrast to a standard linear model, it does not predict negative deforestation. This is particularly important in our setup because the estimated ex-post optimal list depends crucially on having reasonable predictions for counterfactual deforestation, yet there are many municipalities with low levels of deforestation in the data (as expected, given that deforestation is a costly process), and the linear model predicts negative deforestation for over 14 percent of these observations – a non-negligible portion. This may lead to biased ATT estimates and in turn produce misleading results when constructing the counterfactual optimal list, concerns that motivate our use of the CIC model that follows.

5.2 Changes-in-Changes

The CIC model developed by A&I is a nonlinear generalization of the DID model to the entire distribution of potential outcomes. Formally, potential share of deforestation Y_{mt}^{j} – whether in the presence or absence of the policy intervention – is given by the nonparametric specification:

$$Y_{mt}^{j} = h^{j}\left(X_{mt}, U_{mt}, t\right),$$

for $j = 0, 1$, where U_{mt} is a municipality-level unobservable term that can incorporate municipality fixed effects (reflecting permanent differences across m in terms of, say, unmeasured soil quality, climatic conditions, topography, etc.) in addition to time-varying unobservables; for instance, we allow for (but are not restricted to) a decomposition of the form $U_{mt} = \alpha_m + \eta_{mt}$. The function h^j allows for very flexible time trends. In terms of the impact of the policy, one might expect $h^1(x, u, t) \leq h^0(x, u, t)$ for any (x, u, t), given that the Priority List increases monitoring and enforcement intensity.

[34] Specifically, the relevant individual farmer's land use choice model takes the following form: Consider a parcel of forested land i located in municipality m at time period t. Let Y_{imt} equal one if the plot is cleared and zero otherwise. The farmer deforests the plot when $Y_{imt} = 1\{X_{mt}'\beta + v_{mt} > \varepsilon_{imt}\}$, where v_{mt} incorporates all variables on the right hand side of (2) except the control vector X, and ε_{imt} reflects unobserved heterogeneity within the municipality capturing the farmer's idiosyncratic abilities, effort and other influences on farmers' decisions to deforest. When ε_{imt} follows a logistic distribution, the probability that the plot of land i in municipality m at time t is deforested conditional on X_{mt} and v_{mt} is given by the logit formula, $e^{X_{mt}'\beta+v_{mt}}/(1+e^{X_{mt}'\beta+v_{mt}})$, which in turn implies equation (2). Note that this assumes that the distribution of ε_{imt} is not affected by the treatment, which is reasonable given that selection into treatment does not occur at the level of the farmer and is not part of the farmer's choice set, absent moving.

We impose four assumptions on the model. Following A&I, we first make

Assumption 1 Strict Monotonicity: *The functions $h^j(x, u, t)$ – for $j = 0, 1$ – are strictly increasing in u.*

This assumption is satisfied by the DID model, which assumes u enters the function h^0 additively (having constructed the log odds ratio using the share of deforestation). While imposing strict monotonicity of h^j on the unobservables u involves a loss of generality, it allows for more flexible functional forms than a purely additive function – interactions between the time trend and the municipality-level unobservables, for instance. Allowing for such interactions is important because conversion costs may increase and/or land quality may decrease as deforestation in a municipality progresses – if, for example, farmers opt to deforest first in locations with lower conversion costs or higher land quality.

We do not restrict the way in which the functions h^j are affected by treatment status j. Municipalities at different stages of the deforestation process may respond differently to the policy intervention, resulting in heterogeneous treatment effects. Further, because h^0 and h^1 can both change flexibly over time, the intervention may have dynamic impacts. For example, farmers' decisions to deforest might differ in municipalities that have been on the Priority List longer: monitoring could change based on the length of time on the list,[35] or it may take some time for potential deforesters to update their beliefs about the probability of being caught and fined.

Assumption 2 Time Invariance Within Groups: *Conditional on each group G, (i) the unobservable U is independent of X, and (ii) U has an identical distribution over time.*

Assumption 2(i) is the typical extension of the zero correlation assumption from linear to nonlinear models. Note that because of the conditioning on groups, the assumption allows the distribution of X_{mt} to vary by group and with time. Put differently, we do not need the groups to be balanced (nor to reweight and balance them) in terms of their observable characteristics in order to estimate treatment effects.

Assumption 2(ii) requires any unobservable differences between Priority and non-Priority municipalities to be stable over time. That is, the distribution of U_{mt} among the Priority municipalities must be the same in different time periods, and the same for non-Priority municipalities. This is a key condition for the CIC model, playing a role similar to the common trends assumption in the standard DID model: in order to construct counterfactual predictions based on the observable distributions, some form of stability over time is necessary.

[35] Differences in the intensity of regulatory effort across municipalities and over time may also result in heterogeneous treatment effects.

We note that this assumption is less demanding than it might appear. First, the realizations of U_{mt} may vary over time, and can be serially correlated (for instance, due to the presence of fixed effects), although they must come from the same distribution.[36] Second, the distribution of unobservables does not have to be the same across treatment and control groups; treatment effects can be heterogeneous across municipalities *and* across groups G. Recall that the selection rule discussed in Section 4 is based on the assumption that deforestation is a persistent process: highly deforested locations in the past are expected to be more likely to deforest more in the future. This suggests that systematically higher unobservables lead to both higher levels of new deforestation as well as to a higher probability of being placed on the Priority List (through past deforestation), consistent with there being systematic unobservable differences across groups. Assumption 2 thus allows for policy interventions targeted at a group with potentially higher average benefits.

Identification. In discussing identification, we adapt key results in A&I to our context. Denote by $F_{Y^j_{gt}}$ the conditional distribution function of potential share of deforestation Y^j_{mt} given $G = g$ and $X = x$ (to simplify notation, we will omit the conditioning variables X). Let the inverse distribution be given by $F^{-1}_{Y^j_{gt}}(q)$ for any quantile $q \in [0, 1]$. (When it is sufficiently clear from the context, we also use the short-cut notation Y^j_{gmt} to denote the potential outcome variable for a municipality in group g.)

To simplify exposition, take two consecutive periods t and $t + 1$, before and after treatment. Athey and Imbens (2006, Theorem 3.1 and Corollary 3.1) show that under Assumptions 1 and 2, the counterfactual distribution of Y^0_{1mt+1} (i.e., the distribution for the treated group $g = 1$ in the absence of the policy intervention, $j = 0$, at $t + 1$) is identified on the support of Y_{0mt+1} (i.e., the support of the control group at $t + 1$) and is given by

$$F_{Y^0_{1t+1}}(y) = F_{Y_{1t}}\left(F^{-1}_{Y_{0t}}\left(F_{Y_{0t+1}}(y)\right)\right), \tag{3}$$

where $y \in Supp\,(Y_{0mt+1})$. In words, the counterfactual distribution $F_{Y^0_{1t+1}}$ can be calculated based on the distribution of three *observable* variables: the distribution of deforestation shares for the same group but prior to the treatment ($F_{Y_{1t}}$), and the distributions of the share of deforestation for the control group both before and after the treatment ($F_{Y_{0t}}$ and $F_{Y_{0t+1}}$). Note that the distribution of Y for the treated group under the treatment at $t + 1$ (after treatment) is trivially identified: $F_{Y^1_{1t+1}} = F_{Y_{1t+1}}$. By comparing the observed $F_{Y_{1t+1}}$ with the counterfactual $F_{Y^0_{1t+1}}$, we can obtain

[36] This is less restrictive than the parallel trend assumption underlying the DID estimator: while the CIC model allows group and time effects to differ across individuals with different (observed and unobserved) characteristics, the DID model implicitly imposes constant group and time effects.

various treatment effects on the treated (average effects, quantile effects, etc.).

Equation (3) is the nonparametric nonlinear analog of the counterfactual expected deforestation from the DID model. Intuitively, it uses 'double-matching' (as A&I explain clearly – see their Figure 1, page 442) to construct the counterfactual distribution: a treated municipality that deforested a fraction y of its forested area during period t is first matched to an untreated municipality that deforested the same fraction during the same time period. Then the untreated municipality is matched to its rank counterpart (i.e., in the same quantile) among untreated units in period $t+1$. Let y' denote the fraction deforested by this last unit during $t+1$, and define $\Delta \equiv y' - y$. The difference between the shares of deforestation of the treated unit during t and during $t+1$ *in the absence of treatment* is then given by the difference between the deforestation shares of the untreated units *with the same rank* before and after treatment. That is, the counterfactual share of deforestation of the treated unit in the absence of treatment is given by $y + \Delta$.[37] This is similar to the adjustment in the standard DID model, although in the DID case, the adjustment is linear and to the mean, given by:

$$E(Y^0_{1mt+1}) = E\left(Y_{1mt}\right) + \left[E\left(Y_{0mt+1}\right) - E\left(Y_{0mt}\right)\right].$$

When the data set covers one time period before treatment, the model is just-identified. With more than one pre-treatment time period, there is more than one way to identify $F_{Y^0_{1t+1}}$: the model becomes overidentified and the equality in (3) is testable. (Note that $F_{Y^0_{1t+1}}$ is identified only on the support of Y_{0mt+1} for the control group at $t+1$: outside this support, $F_{Y^0_{1t+1}}$ is not identified.)

A similar expression to (3) holds for the control group under the same assumptions (Athey and Imbens, 2006, Theorem 3.2):

$$F_{Y^1_{0t+1}}(y) = F_{Y_{0t}}\left(F^{-1}_{Y_{1t}}\left(F_{Y_{1t+1}}(y)\right)\right), \tag{4}$$

where $y \in Supp\left(Y_{1mt+1}\right)$. Thus equation (4) provides information about treatment effects on the untreated. (As before, the counterfactual distribution for the untreated $F_{Y^1_{0t+1}}$ is not identified outside the support of the treated group Y_{1mt+1}.)

Support Conditions and Partial Identification. When the support conditions are not satisfied, we cannot identify the counterfactual distributions at the lower and upper tails. However,

[37] Note that the 'double-matching' here is based on the outcome variable, while selection-on-observables methods perform matching based on covariates (or on propensity scores).

we can obtain worst-case bounds in a spirit similar to Manski (2003).[38] To do so, we need prior information relating to the counterfactual support for Y^1_{0mt+1}. Assumption 3 provides such prior information, and has been implemented previously in the empirical literature (see, e.g., Ginther, 2000; Lee, 2009).

Assumption 3 Support: *Assume* $Supp(Y^j_{gmt}) = Supp\,(Y_{gmt})$ *for* $j, g = 0, 1$, and for any t.

Assumption 3 implies that while the policy intervention may affect the distribution of deforestation shares, it does not affect the support of the distribution. By putting all mass outside $Supp\,(Y_{1mt+1})$ at the left and right end points of $Supp\,(Y_{0mt+1})$, we obtain the lower and upper bounds for $F_{Y^1_{0t+1}}$, denoted by $F^L_{Y^1_{0t+1}}$ and $F^U_{Y^1_{0t+1}}$ respectively (the same reasoning applies to $F_{Y^0_{1t+1}}$).[39]

Note that under Assumption 3, we cannot point identify the counterfactual distributions of both treated and untreated groups simultaneously when $Supp(Y_{1mt+1}) \neq Supp(Y_{0mt+1})$. Further, if $Supp(Y_{1mt+1}) \subset Supp(Y_{0mt+1})$, we can point identify the counterfactual distribution for the treated group $F_{Y^0_{1t+1}}$, but not the control group, $F_{Y^1_{0t+1}}$. In this case, we identify the average treatment on the treated, but we can only partially identify the average treatment on the untreated.

Semiparametric Specification. Although the CIC model can be estimated completely nonparametrically (Athey and Imbens, 2006; Melly and Santangelo, 2015), we adopt a semiparametric specification because of data limitations. The simplest and most parsimonious procedure is to partial-out the covariates X_{mt} and apply the CIC model to the residuals, as A&I suggest.

For comparability with DID, we adopt the logit model. In addition to the reasons given in Subsection 5.1, it is worth adding that the logit specification is useful when estimating the CIC model because it allows for heterogeneous effects of X_{mt} on deforestation, helpful when selecting the ex-post list; if heterogeneous effects were restricted to depend only on unobservables, the ex-post list would only select all municipalities in the group with the higher average impact of treatment. Further, the logit model has a convenient functional form that makes it easy to partial-out the covariates in order to estimate the CIC model. A fully nonparametric model would require estimating all conditional distribution functions given X in equations (3) and (4) nonparametrically,

[38] For instance, if $Supp\,(Y_{1mt+1}) \subset Supp(Y^1_{0mt+1})$, then $F_{Y^1_{0t+1}}$ is identified on the subset $Supp\,(Y_{1mt+1})$, and we place the remaining probability mass outside $Supp\,(Y_{1mt+1})$ at the end points of $Supp(Y^1_{0mt+1})$. We assume the supports of the observed variables are connected, so that *only* at the tails is there no information about $F_{Y^1_{0t+1}}$.

[39] These are worst-case bounds because they do not incorporate possible additional restrictions such as continuity or smoothness on counterfactual distributions. In order to minimize the impact of outliers, we follow the literature and trim observations below the 3rd and above the 97th percentiles (Ginther, 2000; Lee, 2009). The empirical results are robust to the trimming, for example dropping observations below and above the percentiles $[2.5, 97.5]$ and $[3.5, 96.5]$. See Appendix F.

which is not practical in our setting.[40]

We assume the following

Assumption 4 Semiparametric Model: *The potential share of newly deforested area, Y_{mt}^j, for $j = 0, 1$, in a municipality m at t is given by*

$$Y_{mt}^j = \frac{\exp\left[X_{mt}'\beta + V_{mt}^j\right]}{1 + \exp\left[X_{mt}'\beta + V_{mt}^j\right]}, \tag{5}$$

where V_{mt}^j are unobservable variables such that (a) $V_{mt}^j = v^j(U_{mt}, t)$, where the functions $v^j(u, t)$ satisfy Assumption 1 (i.e., strict monotonicity on u), and (b) V_{mt}^j satisfy the support condition in Assumption 3.

By regressing the log odds ratio of the share of deforestation on covariates, we can identify and estimate the coefficients β (by Assumption 2). We can therefore back out the residuals V_{mt}, and apply the CIC model to them.[41]

5.3 Average Treatment Effects

We now discuss how we calculate the average treatment effects. Start with the logistic function $\varphi(x, v) = \exp(x'\beta + v) / (1 + \exp(x'\beta + v))$. From (5), the potential share of new deforestation is given by $Y_{mt}^j = \varphi(X_{mt}, V_{mt}^j)$. The expected deforestation under intervention j, D_{mt}^j, conditional on observables (X_{mt} and A_{mt}) and on the group $G = g$, is given by

$$E\left[D_{mt}^j | X_{mt}, A_{mt}, G_m = g\right] = \int \left[\varphi(X_{mt}, v) \times A_{mt}\right] dF_{V_{gt}^j}(v), \tag{6}$$

where the distribution $F_{V_{gt}^j}$ is either observed (from the residuals of the log odds ratio regression) or is obtained from the CIC model (i.e., from either (3) or (4) applied to the residuals V_{mt}). Given (6), average treatment effects are defined in the standard way. When the support conditions are violated, the counterfactual distributions are not identified, in which case we bound the conditional

[40] An alternative solution, proposed by Kottelenberg and Lehrer (2017), is to reweight the observations based on the propensity scores. Although appealing, this solution is of limited use in the current context because of the lack of common support on propensity scores induced by the selection rule (see Section 4).

[41] More specifically, as A&I note, let I_{mt} be a vector of dummy variables indicating group status (control versus treatment) interacted with time dummies. In the first stage, we estimate the regression $\log\left(\frac{Y_{mt}}{1-Y_{mt}}\right) = X_{mt}'\beta + I_{mt}'\gamma + \nu_{mt}$, then construct the residuals with the group-time effects left in: $\log\left(\frac{Y_{mt}}{1-Y_{mt}}\right) - X_{mt}'\widehat{\beta} = I_{mt}'\widehat{\gamma} + \widehat{\nu}_{mt}$.

expectations as

$$\begin{aligned} & \int \left[\varphi\left(X_{mt}, v\right) \times A_{mt}\right] dF^{L}_{V^{j}_{gt}}\left(v\right) \\ \leq \quad & E\left[D^{j}_{mt} | X_{mt}, A_{mt}, G_{m} = g\right] \\ \leq \quad & \int \left[\varphi\left(X_{mt}, v\right) \times A_{mt}\right] dF^{U}_{V^{j}_{gt}}\left(v\right). \end{aligned} \tag{7}$$

Bounds on average treatment effects follow naturally from (7). Given that the evolution of the remaining forested area depends on deforestation in previous periods, we take dynamics into account when calculating counterfactual deforestation (see Appendix D).

In turn, to measure the carbon emissions that result from the deforestation process, we consider the equality $E^{j}_{mt} = D^{j}_{mt} \times CS_{mt}$, where E^{j}_{mt} are the potential carbon emissions under policy j, and CS_{mt} is the average difference in carbon stock comparing forested and deforested areas within municipality m.[42]

6 Empirical Results

In this section, we present estimates of the Priority List's effects. We first show difference-in-differences results, followed by results based on the changes-in-changes model. We then provide evidence relating to possible mechanisms that might link the Priority List and deforestation.

Difference-in-Differences. Table 3 presents the coefficients from estimating the DID regression model specified in equation (2). The first column does not include covariates, while the second column does. In the third and fourth columns, we incorporate potential spillover effects in the estimation strategy, splitting the untreated municipalities into 'control' and 'spillover' groups; column (3) does not include covariates while column (4) does.[43]

In all specifications in the table, the Priority List appears to have reduced deforestation substantially. First when ignoring spillovers, the coefficients on Priority status after treatment are statistically significant, and show an average reduction in the odds ratio of the deforestation share of approximately 45% in 2009 and 90% in 2010. The impacts are robust to the inclusion (or exclusion) of the covariates in the specification, with the greater impact in 2010 possibly due to farmers

[42] For simplicity, we ignore carbon decay and assume all carbon stock is immediately released into the atmosphere once a plot of land is deforested.

[43] Recall that we split the untreated group according to whether a municipality is likely to be affected by the intervention or not. We consider two criteria: (a) if a municipality shares a border with a treated municipality, and (b) if a municipality has high levels of past deforestation (determined by how close Z^{1}_{mt-1} and Z^{2}_{mt-1} are to the threshold values that the Brazilian government (implicitly) adopted in the selection rule). The municipalities satisfying these two criteria are referred to as the 'spillover' group.

updating their beliefs about the new policy regime. Next, when potential spillover effects are taken into account, the coefficients on Priority status presented in columns (3) and (4) are slightly greater than the corresponding estimates ignoring potential spillovers. The average impact on the odds ratio is around 47% in 2009 and 92% in 2010. This greater impact is attributable to the fact that lower average reductions in deforestation after treatment now arise in the control group, given that it does not include those municipalities more likely to respond to the policy intervention. Overall, the estimated impacts of the Priority List on deforestation are economically and statistically significant, and robust across all specifications.

Columns (3) and (4) of the table also shed light on the estimated *spillover* effects of the Priority List following treatment. We find these effects to be negative – small in magnitude and not statistically significant in 2009, but becoming larger (and significant) in 2010, suggesting that spillovers may have taken time to emerge. The evidence indicates that untreated municipalities with a treated neighbor and high levels of past deforestation reduce their deforestation rates in response to the establishment of the Priority List (perhaps following belief-updating by farmers). The odds ratio of the share of deforestation declines, on average, by approximately 24% in 2009 and 62% in 2010. Again, the impacts are robust to whether or not the covariates are included in the specification.[44]

Testing. As is standard, we test whether the trends in the outcome variables are parallel in the pre-treatment period. Table 4 presents the results, the first column ignoring potential spillover effects, the second incorporating them. The table provides no evidence that the common trends assumption is violated before treatment: the coefficients on the time dummies interacted with Priority and spillover status are not statistically significant before 2008. This evidence accords with the discussion in Section 4, where we noted that the government's criteria for entering the Priority List indicated a rule that selected municipalities based principally on the level (rather than the trend) of past deforestation.[45]

Changes-in-Changes. We now turn to the estimates of the CIC model, first ignoring then accounting for spillovers. The estimates ignoring spillovers are reported in Table 5. In the top panel, we present the estimated effects on deforestation, as explained in Section 5. The columns in the panel give the estimated ATT, ATU, and ATE, respectively, which are provided (in the rows) separately for 2009 and 2010, and also the total cumulative deforestation in those two years

[44] Table 15 in Appendix G presents the coefficients on all regressors, including the covariates.

[45] The pre-treatment parallel trend assumption is robust to including more years (2003–2007) in the panel data regressions, notably the time period covering the PPCDAm structural break in 2004–2005. (See Table 16 in Appendix G.)

(summed over all municipalities). Results are provided using 2006 and 2007 as alternative baseline years, noting that the CIC model is over-identified when more than one time period before the treatment is available in the data.

The bottom panel of the table reports the estimated total cumulative treatment effects in terms of carbon emissions. We label these CTT, CTU, and CTE – the cumulative analogs of ATT, ATU, and ATE. That panel also reports the value of the total emissions avoided, assuming a social cost of carbon of \$20/tCO_2 (Greenstone et al., 2013; Nordhaus, 2014). The numbers in square brackets are lower and upper bound estimates for the partially identified sets, and the numbers in parentheses are 95 percent confidence intervals.[46]

We start with the treated group. All results are statistically significant and are robust to the choice of the baseline year. The estimated ATT for 2009 is between -22 km^2 and -25 km^2 (depending on the baseline), and between -51 km^2 and -54 km^2 for 2010. The pattern of increasing effects over time is similar to that in the DID regression model.[47] According to the CIC estimates using 2006 as the baseline year (which turns out to provide more conservative estimates), the treated group would have deforested a total of 6,570 km^2 in the period 2009–2010 in the absence of treatment, which is 63 percent higher than the amount of deforestation observed in the data. Thus, the estimates indicate that the Priority List led 2,540 km^2 of forested area to be preserved, and emissions of 30 million tons of carbon to be avoided in the same period. The estimated social benefit of the program in terms of avoided emissions is approximately \$2.21 billion. Compared to the combined budget allocated to IBAMA and INPE – about \$600 million – the results suggest the program was highly beneficial, and that further investments in monitoring and enforcement would be worthwhile.

Treatment effects on the untreated are not point-identified, but the identified sets are highly informative, the effects being statistically significant and robust to the choice of the baseline year.[48]

[46] For ATT, the 95 percent confidence intervals are computed based on the standard i.i.d. nonparametric bootstrap, where the i.i.d. resampling occurs in the cross-sectional dimension. The ATU and ATE are both partially identified (as we discuss below), the confidence intervals being based on the procedure in Imbens and Manski (2004) for the parameter of interest (not for the identified set), where the bootstrap replications are used to compute standard errors for the lower and upper bound estimators. (We implemented 500 bootstrap replications.)

[47] The DID estimator is not scale-invariant, so there is no *a priori* reason to expect the DID and CIC estimators will generate similar point estimates. We find estimated effects from using CIC that are somewhat smaller when computed on a comparable basis. (For instance, applying the CIC estimator to the log odds ratio of deforestation shares not conditioning on covariates, which is directly comparable to the coefficients on Priority status presented in column (1) of Table 3, implies a reduction in the odds ratio of 42% in 2009 and 67% in 2010, on average – smaller than the corresponding DID estimates.)

[48] As discussed in Section 5, the counterfactual distribution of the control group, $F_{V^1_{0t}}$, is identified on the support of the treated group. Because the treated group has a substantially smaller number of observations than the control group in the data, the estimated support of the treated group is strictly contained in the support of the control group (see Table 14 in Appendix G). This implies that the counterfactual distribution $F_{V^1_{0t}}$ is identified only on a subset of its support, and not at the tails. This means the ATU can only be partially identified. (See Figures 13 and 14 in Appendix G for the estimated factual and counterfactual distribution functions of the residuals V^J_{gt}, for both treated

The estimated effects range from -3.3 km^2 to -4.7 km^2 for 2009, and increase to between -5 km^2 and -7.4 km^2 for 2010.

The difference between the estimated ATT and ATU provides evidence of heterogeneous treatment effects, suggesting that the government did indeed select municipalities with potentially higher average impacts. Such results could not be obtained using a DID strategy given that it only identifies effects on the treated. Further, extrapolating results from the treated group to the untreated under the assumption of homogeneous effects would bias up the effects on the untreated.

Table 6 presents the estimated treatment effects based on the CIC model, now incorporating potential spillovers. Similar to the DID results, we find that the ATTs are now slightly greater than the ones estimated ignoring potential spillover effects. Using 2006 as the baseline year, the estimates indicate that the Priority List avoided the clearing of 2,705 km^2 of forested area and emissions of 32 million tons of carbon during 2009–2010, which implies a social benefit of the program on the order of approximately \$2.4 billion.

The estimated ATUs are also in line with the estimates obtained when we assumed away any spillovers effects. The average treatment effects on the spillover group, denoted ATS, are statistically significant and robust to the choice of baseline year. While the ATSs are partially identified, the estimated sets are very informative: the effects range from -11 km^2 to -16 km^2 for 2009, and increase to between -15 km^2 and -25 km^2 for 2010. Similar to the other groups, impacts are greater during the second year of the program. The magnitudes of the ATS fall between the estimated ATT and ATU, constituting further evidence of heterogeneous effects.[49]

In sum, according to the CIC estimates that account for spillovers, the direct impact of the program avoided 2,705 km^2 of deforestation in Priority municipalities in 2009–2010, while the indirect impact in the same period discouraged the clearing of 618 km^2 of forested area in spillover municipalities. This amounts to a total of 3,323 km^2 of forested area preserved (treating 2006 as the baseline year). In turn, the program avoided 39.5 million tons of carbon emissions, with a social benefit of approximately \$2.9 billion: these are our preferred summary estimates of the policy impacts.

Testing. We now discuss the results of three tests applied to the CIC model. First, we assess whether the actual distribution of deforestation shares equals the counterfactual distribution when imposing the policy intervention (falsely) in 2007, one year early; this serves as a placebo test – the

and control groups in 2009 and in 2010.)

[49] Note that the ATS presented in Table 6 differs conceptually from the coefficients on spillover status presented in Table 3. The ATS measures the average effect (on deforestation) of including a spillover municipality on the Priority List, while the coefficients on spillover status presented in Table 3 estimate how farmers in spillover municipalities reacted to the existence of the policy intervention itself.

CIC analog to the DID pre-treatment common trend test. Second, we test whether the Priority List affects the entire distribution of outcomes – similar to the placebo test but using the correct timing of the intervention. Third, we test whether the counterfactual distribution is everywhere below the actual distribution, as would be the case if the absence of treatment resulted in more deforestation everywhere – a stochastic dominance test.[50]

Table 7 presents the results. We apply each test to both the log odds ratio of deforestation shares not conditioning on covariates and to the residuals (V_{mt}) after partialling out the covariates, as explained in Section 5. In all cases, the p-values correspond to both the Kolmogorov-Smirnov and the Cramer-von Mises statistics. For the two outcomes and associated test statistics, we fail to reject the null of 'no impact' when the policy intervention is wrongly imposed in 2007 – i.e., the placebo test passes. In contrast, we reject the null of no impact when the policy intervention is set correctly in 2008, and we find strong evidence in favor of stochastic dominance, as one might expect given the estimated treatment effects discussed above.

Possible Channels. We now seek to shed light on the channels through which the Priority List may have affected deforestation. The list consisted of a bundle of provisions, as documented above, with farmers in Priority municipalities becoming subject to more rigorous monitoring and law enforcement; they also faced more stringent conditions when seeking to obtain subsidized credit contracts, along with stricter licensing and geo-referencing requirements. At the same time, the government might have expanded protected areas strategically, taking into account the location of Priority municipalities.

Given this bundle of provisions, we are interested in exploring (subject to data limitations) whether particular elements of the bundle appear to have been especially important in cutting deforestation. We focus on observables, investigating how the number of alerts given out by INPE (a proxy for monitoring), the evolution of the number of fines issued by IBAMA (a proxy for enforcement), the total volume of rural credit concessions, and the share of protected areas all vary by treatment status – treated, spillover, or control.[51]

Figure 9 presents the evolution of these variables over time by treatment status. Taking our

[50] All three tests were proposed and developed by Melly and Santangelo (2015). We apply them only to the treated group because the estimated distributions for the spillover and control groups are not point-identified. (To the best of our knowledge, the corresponding formal testing that would cover partially identified cases has not yet been developed.) We implemented 500 bootstrap replications and, given that the choice of the baseline years for comparisons do not affect results substantially, we combine them to increase power.

[51] While unlikely to be a first-order concern, licensing requirements are harder to pin down because of data limitations; for instance, we cannot establish whether stricter licensing requirements were implemented differentially across treated and untreated municipalities. Similarly, we do not have information concerning the Soy Moratorium – an agreement operating since 2006 among global soy traders not purchase soy grown on farmland that does not accord with the Forest Code – specifically, whether it affected Priority and non-Priority municipalities differently.

enforcement proxy first (see panel (a)), the number of fines increased substantially among treated municipalities in 2008 and then fell subsequently within this group, likely in response to the lower deforestation rates observed there following the policy intervention. In contrast, the number remained reasonably stable in spillover and control groups during 2006–2010. In combination, the enforcement evidence points to an overall increase in enforcement intensity, concentrated on Priority municipalities, rather than a substitution occurring away from untreated municipalities. In terms of our monitoring proxy (panel (b)), the number of alerts is higher for treated municipalities, declining markedly after 2008, again likely reflecting the deforestation slowdown among Priority municipalities, while the average number of alerts issued in untreated municipalities is stable before and after the establishment of the Priority List. (Unlike fines, there is no clear upward shift in the number of alerts by 2008 – instead, it oscillates in the pre-treatment period.) Again, the graphical evidence does not suggest any major substitution of monitoring effort away from untreated municipalities. Beyond monitoring and enforcement, there are no clear distinguishing patterns by treatment group status in terms of total rural credit (other than differences in levels), and shares of protected areas by treatment status show almost no variation over time.

Table 8 presents corresponding regression evidence. Specifically, we regress the four observable measures – fines, alerts, credit, and protected area share – on Priority and spillover status indicators interacted with time dummies and on covariates, taking 2006 as the baseline year. The regression estimates paint a similar picture, suggesting an increase in enforcement (fines) among treated municipalities, and more so than an increase in monitoring effort (alerts), while no clear decrease in enforcement or monitoring is apparent among the untreated. We find no clear evidence that more stringent conditions were applied to limit the concession of subsidized rural credit in Priority municipalities, nor is there any evidence of the strategic placement of new protected areas.[52]

Together, this suggestive evidence helps clarify whether state capacity increased or instead whether there was simply a reallocation of fixed resources following the Priority List's introduction. The graphical and regression evidence is consistent with more focused targeting being associated with an increase in state capacity to implement environmental regulations, in turn altering municipal-level behavior (as reflected in aggregate deforestation).

7 Optimal Policy Targeting

In this section, we develop a counterfactual framework for targeting regulations optimally based on the estimated treatment effects just reported, then present the empirical results from various

[52]This complements (and is consistent with) the analysis presented in Assunção and Rocha (2019).

counterfactual targeting exercises.

7.1 Policy Targeting Framework

Suppose a policy maker wishes to assign municipalities to the Priority List in order to minimize total deforestation (or total emissions), and that she has information about the conditional average treatment effects estimated above, along with the covariates. Denote the counterfactual assignment rule in time period t by $\phi_t = (\phi_{1t}, ...\phi_{Mt})$, which maps the treatment to municipalities $m = 1, ..., M$ and which can be either deterministic $\phi_{mt} \in \{0, 1\}$ or probabilistic $\phi_{mt} \in [0, 1]$. For a given time period, the policy maker solves the problem

$$\min_{\phi_t \in [0,1]^M} \sum_{m=1}^{M} \left[\phi_{mt} E\left[D_{mt}^1 | X_{mt}, A_{mt}, G_m \right] + (1 - \phi_{mt}) E\left[D_{mt}^0 | X_{mt}, A_{mt}, G_m \right] \right]. \tag{8}$$

The minimum deforestation is achieved (trivially) by a singleton rule that allocates m to the treatment when $E\left[D_{mt}^1 | X_{mt}, A_{mt}, G_m\right] \leq E\left[D_{mt}^0 | X_{mt}, A_{mt}, G_m\right]$; when the equality holds, any random allocation is optimal.

The minimization problem in (8) abstracts from two important considerations. The first involves constraints. The original Priority List had the intention of directing limited resources where they were expected to have the greatest impact. Given that data on the resources effectively allocated in practice to monitoring are difficult (if not impossible) to obtain, we incorporate limited monitoring resources into the policy maker's minimization problem by means of two alternative constraints. One constraint limits the total area $\overline{S}$ that can be monitored under the Priority List (given the plausible notion that the costs of monitoring and punishing illegal deforestation increase with the *total area* covered by the policy). We write this as:

$$\sum_{m=1}^{M} s_m \times \phi_{mt} \leq \overline{S}, \tag{9}$$

where s_m is the area of municipality m. The alternative constraint applies to the *total number* of municipalities $\overline{M}$ that can be placed on the list:

$$\sum_{m=1}^{M} \phi_{mt} \leq \overline{M}. \tag{10}$$

This constraint is reasonable when monitoring costs are primarily a function of the *number* of districts that inspectors must visit.[53]

[53] Ideally, we would have precise information on the expected monitoring costs for each municipality m in each

The second aspect concerns partial identification: when the support conditions are violated, we can only partially identify counterfactual expected deforestation. This means that an ex-post policy evaluation must be analyzed as a treatment choice problem under ambiguity (Manski, 2005a). We consider the minimax criterion, assuming the policy maker chooses the ex-post list in order to minimize total deforestation in the worst-case scenario.

Formally, let all the feasible values that $E[D^j_{mt}|X_{mt}, A_{mt}, G_m]$ can take be indexed by $\gamma \in \Gamma$ (given by the inequality (7)). The policy maker's problem under the minimax criterion is

$$\min_{\phi_t \in [0,1]^M} \sup_{\gamma \in \Gamma} \sum_{m=1}^{M} \left[\phi_{mt} E_\gamma \left[D^1_{mt}|X_{mt}, A_{mt}, G_m\right] + (1-\phi_{mt}) E_\gamma \left[D^0_{mt}|X_{mt}, A_{mt}, G_m\right]\right], \tag{11}$$

subject either to the 'total area' constraint (9), or to the 'total number of municipalities' constraint (10). The minimization problem (11) subject to either constraint is a linear programming problem that is straightforward to solve numerically. In the empirical exercise, when using constraint (9), we set $\overline{S}$ equal to the total area occupied by the municipalities that were effectively put in the list in 2008 (i.e., the treated group). Similarly, when using constraint (10), we set $\overline{M} = 35$, which is the number of municipalities in the treated group. We do so because we can then assess how close the observed Priority List was to the ex-post optimal assignment.[54]

To provide intuition for the assignment of municipalities to the optimal list, we start from the observation that in the absence of any constraint, the minimum deforestation is achieved by following a simple rule: a municipality m that was originally in the control group ($G_m = 0$) is assigned to the optimal list when the expected deforestation in the absence of treatment, $E\left[D^0_{mt}|X_{mt}, A_{mt}, G_m = 0\right]$, is greater than the maximum possible amount of expected deforestation under the treatment. In turn, a municipality m that was originally in the treatment group ($G_m = 1$) should not be on the list when the expected deforestation under treatment, $E\left[D^1_{mt}|X_{mt}, A_{mt}, G_m = 1\right]$, is greater than the maximum possible amount of deforestation in the absence of treatment. Note that the assignment rule differs depending on observed priority status because the objects that are partially

time period t, in both the absence and presence of the treatment. Then we could replace the constraints (9) and (10) with the restriction

$$\sum_{m=1}^{M} \left[\phi_{mt} E\left[MC^1_{mt}|X_{mt}, G_m\right] + (1-\phi_{mt}) E\left[MC^0_{mt}|X_{mt}, G_m\right]\right] \leq K_t,$$

where MC^j_{mt} are the monitoring and enforcement costs, and K_t is the government's budget constraint. (We note that our framework can accommodate other objective functions – e.g., using $R^j = P \times D^j + MC^j$ in the social cost function (8), where P is the social cost of deforesting one parcel of land.) Such an approach is not feasible, however. Although we do know IBAMA's and INPE's total budgets, in practice we do not have information about the true budget constraint K_t. Further, we do not know how much of the total is allocated to monitoring, nor do we have data indicating how monitoring costs are distributed across municipalities.

[54] The fact that the constrained minimization problem can be specified as a linear programming problem is convenient: in the data, the number of possible lists under the constraint $\overline{M} = 35$ is $\binom{490}{35} \approx 4 \times 10^{53}$.

identified differ. When the constraints are taken into account, the estimated magnitudes of the treatment effects for *all* municipalities matter in the minimization problem.[55]

Given that the optimal list is based on ex-post knowledge of the treatment effects, the differences in the amount of deforestation and carbon emissions under the minimax optimal assignment rule and the observed assignment rule provide lower bounds for the social value of the ex-post information about the treatment effects. Put another way, the difference measures the minimum amount that the policy maker (or the society) would be willing to pay to obtain the ex-post information.

In the presence of spillover effects, the objective function is non-linear and non-differentiable in ϕ, so that we cannot solve the minimax problem using standard methods. Instead, to find the global minimum, we use a stochastic search algorithm – more precisely, a genetic algorithm that allows for integer optimization in high-dimensional constrained minimization problems. (See Appendix E for detailed description of the way the ex-post optimal list is calculated in this case.)

7.2 Policy Targeting Results

We now present results absent potential spillovers, then show how targeted policies are affected once spillover effects are taken into account.

'No Spillovers' Case. Table 9 compares the original Priority List with the ex-post optimal list obtained by solving the relevant constrained minimizations: the left panel considers the total area $\overline{S}$ that can be monitored as the constraint (see equation (9)), while the right panel fixes the number of municipalities $\overline{M}$ (equation (10)).[56]

Overall, the proportion of municipalities that appear on both lists is high: 83.7 percent when the constraint involves the total area, and 93.5 percent when the constraint is a maximum number of municipalities. According to this latter metric, the Priority List is already close to the corresponding ex-post optimal list. When the policy maker is constrained to 'police' a pre-specified overall area, she can reduce deforestation in the worst-case scenario by replacing seven large municipalities on the Priority List with 73 municipalities that are smaller in size but that would help reduce total deforestation (based on the ex post treatment effect estimates). In contrast, when the restriction applies to the number of municipalities, the policy maker would do better by replacing small municipalities (comprising almost half of the Priority List) by municipalities that are larger in size. Indeed, the total area covered by this list is 41 percent larger than the original list.

[55] See Appendix E for details. We do not select a list that changes over time as this complicates the problem substantially, given the combinatorics involved.

[56] We present results using the baseline year 2006. Results treating 2007 as the baseline are similar.

Figure 10 presents the geographic distribution of municipalities on the various lists. For reference, the top left panel presents the actual Priority List and the top right shows the Priority List together with protected areas (composed of conservation units and indigenous reserves). The bottom left panel then shows the optimal list when the constraint is the total area covered, and in the bottom right panel, the counterfactual list when the constraint is the number of municipalities. (The bottom panels also depict protected areas.)

Two interesting patterns emerge – features that were not imposed in the course of the estimation strategy. First, the overlap between the protected areas and the area-constrained counterfactual list is much smaller than the overlap between the protected areas and the original Priority List; indeed, the former comprises an area approximately half the latter. This suggests that, together, these two policies could be further leveraged by the Brazilian government. Second, more specifically, the geographic distribution of the area-constrained counterfactual list traces out a protective shield close to the deforestation frontier; that frontier, the 'Arc of Deforestation,' is located along the southeastern edge of the Amazon Biome. In the current context, the Priority List may therefore serve to work alongside the protected areas in impeding the deforestation process from continuing into more pristine regions.

To shed some light on which observable factors might be more important in determining whether a municipality is placed on the optimal list or not, we estimate simple reduced-form regressions, regressing the optimal list on the covariates X, the 'criteria' variables Z, and the Priority status indicator G. Based on that exercise, the two most important factors predicting the optimal list are the share of protected areas (consistent with our analysis of the geographic distribution of the optimal list – see Figure 10), and Priority status itself (not surprisingly, given the overlap between the two lists presented in Table 9).[57] As an aside, we note that these suggestive reduced-form regressions ignore the fact that the assignment of a municipality onto the optimal list depends on the characteristics of *all* municipalities in the presence of capacity constraints (and spillovers, discussed below).

Next, we seek to quantify the *consequences* of optimally targeted policies. We do so by comparing both the maximum possible deforestation and the carbon emissions achieved under the optimal list with the corresponding outcomes under the Priority List, along with another benchmark: a list composed of municipalities that are selected randomly.[58] Table 10 presents the results. Compared to the area-constrained optimal list, the Priority List results in around 6 percent more deforestation

[57] The criteria variables Z are relevant only when we exclude G from the regressions, which is as one would expect given the threshold selection rule. (In particular, criterion Z^2, capturing more recent deforestation, is the only 'criteria' variable that becomes statistically significant in the absence of G.) (Results are available upon request.)

[58] We simulated 1000 random lists with $\overline{M} = 35$ and computed the average resulting counterfactual deforestation and emissions.

and 5 percent higher carbon emissions in 2009–2010. The estimated avoided emissions translate into a social value of at least $562 million for that two-year span alone. Again, we estimate high social returns to investments that generate information concerning the effects of conservation policies.

We find that the ex-post optimal list fixing the number of municipalities performs slightly better than the area-constrained optimal list. But since it covers a much larger area, monitoring costs are likely to be significantly higher in the former case. In comparison, randomly selecting 35 municipalities onto the list would result, on average, in 23–25 percent more deforestation and 26–29 percent higher emissions than the number-constrained optimal list.

Overall, although the Priority List results in higher deforestation and emissions compared to the two alternative ex-post optimal lists, the magnitudes are not substantially greater. While the ex-post optimal lists were designed to minimize the worst-case scenario, and so should be expected to result in less deforestation and emissions than those presented in Table 10, the estimated performance of the Priority List is (perhaps surprisingly) fairly close to the minimax ex-post optimal lists, especially given that the government made decisions without knowing the potential treatment effects of this policy. The Priority List also compares favourably to a completely random rule. Still, our results indicate that there is clear room for improvement.

Next, we are interested to see how much the minimax solution for carbon emissions is affected by relaxing the constraints we have been imposing. Figure 11 shows the results. The top panel presents the level of emissions at the optimum for the total area constraint, while the bottom panel shows the results when we change the number of municipalities that can be included on the optimal list. In each panel, the vertical lines show the maximum $\overline{S}$ and $\overline{M}$ that correspond to the area covered by, and the number of municipalities on the Priority List respectively. The horizontal lines correspond to the amount of carbon emissions estimated directly from the data for 2009–2010.

The minimax carbon emissions decrease rapidly when a small area is covered by the optimal list and level off for large $\overline{S}$ eventually, indicating that the benefits of including additional municipalities on the list decrease with $\overline{S}$. Because monitoring costs should increase with $\overline{S}$, concentrating efforts on a strategically selected subregion of the Amazon rainforest emerges as a suitable policy. Furthermore, the minimum area needed for the optimal list to generate the same amount of emissions as the original Priority List is 535 km^2, which is approximately 70 percent of the area covered by the original list (763 km^2); this corresponds to the point in the top panel figure at which the minimax carbon emissions curve crosses the horizontal line. This finding is important, drawing attention to substantial monitoring cost savings that are available, holding the level of observed emissions fixed.[59]

[59]The same type of reasoning applies when we change the number of municipalities, $\overline{M}$, allowed on the optimal list, with the minimum number of municipalities generating the same amount of observed emissions as the Priority

'Spillovers' Case. We now discuss the ex-post optimal lists when spillovers are incorporated into both the estimation procedure and the minimization problem.

First, Table 11 compares the Priority List with the ex-post optimal lists based on the two different constraints discussed above, although now allowing for spillovers. Under the total area constraint, the optimal list replaces a greater number of large municipalities by small municipalities when compared to the optimal list with no spillovers. This leads to a smaller proportion of cases that appear in both the optimal list and original Priority List – now 79 percent. Intuitively, such an assignment takes advantage of the fact that a larger number of small municipalities treated can have wider impacts because of spillover effects.

Figure 12 presents the geographic distribution of the resulting optimal lists. When compared to the no-spillover case in Figure 10, the area-constrained optimal list is now more geographically dispersed, with fewer municipalities being contiguous. Such a pattern is also attributable to the operation of spillover effects: placing all targeted municipalities together does not exploit the potential reduction in deforestation in adjacent locations that arises when spillovers are operating. (A similar pattern is observed for the number-constrained optimal list.)

Table 12 compares the levels of deforestation and emissions associated with the alternative lists. Because the optimal lists now take advantage of potential spillover effects, they can achieve lower levels of forest loss in the worst-case scenario. The Priority List now results in around 9–14 percent more deforestation and 8–14 percent higher carbon emissions in 2009–2010 than the area-constrained optimal list (depending on the baseline year). This places a lower bound on the value of the optimal list of approximately $900 million. Results are similar when we consider the optimal list constrained by the number of municipalities that can be included. Randomly selected municipalities now result in 34–44 percent more emissions than the number-of-municipalities constrained optimal list.

More generally, by relaxing the constraints of the minimization problem, we find that the minimum area covered that is needed for the optimal list to result in the same amount of carbon emissions as the Priority List is 326 km^2, which is approximately 43 percent of the area covered by the original list. This suggests even greater monitoring cost savings once spillovers are taken into account. (Similarly, the minimum number of municipalities generating the same amount of observed emissions as the Priority List is 17 – less than half of the original blacklist.)

List being 18, which is half of the original list.

8 Conclusion

In this paper, we have developed a new approach for assessing the efficacy of targeted, blacklist-type policies for slowing deforestation – itself a primary contributor to global carbon emissions and a source of considerable concern worldwide. Focusing on the Priority List introduced by the federal government in the Brazilian Amazon in 2008, we first showed that the policy had a substantial causal impact, using the flexible changes-in-changes approach of Athey and Imbens (2006): deforestation was cut by 40 percent in municipalities placed on the list (relative to the case in which no policy was introduced), and also generated non-trivial spillover effects in the form of lower deforestation elsewhere.

We then used the treatment effect estimates in a counterfactual policy framework that allowed us to compute an ex-post optimal list, reflecting realistic resource constraints faced by the regulatory agency in its ability to monitor behavior and enforce environmental protections. Comparing ex-post optimal lists (on two alternative bases) with the actual Priority List, we showed that optimal targeting can generate significant additional gains. Carbon emissions would be at least 8 percent higher under the Priority List than under the optimal list, even though emissions under the actual list are still significantly lower than they would be under a randomly chosen set of municipalities.

From a regulation perspective, our approach provides a means to quantify, based on credible econometric estimates, the gains to the environment from optimally targeted policies aimed at countering tropical deforestation. More generally, our counterfactual approach using ex-post treatment effects is applicable in a variety of other settings where targeted regulations have been introduced, and can help policy makers to assess which policy configurations (accounting for realistic resource constraints) are likely to have most environmental impact.

References

Abbring, J. H. and J. J. Heckman (2007). Econometric evaluation of social programs, part iii: Distributional treatment effects, dynamic treatment effects, dynamic discrete choice, and general equilibrium policy evaluation. *Handbook of Econometrics 6*, 5145–5303.

Alix-Garcia, J. M., E. N. Shapiro, and K. Sims (2012). Forest conservation and slippage: evidence from Mexico's national payments for ecosystem services program. *Land Econ. 88*, 613–38.

Alix-Garcia, J. M., K. R. E. Sims, and P. Yaez-Pagans (2015). Only one tree from each seed? Environmental effectiveness and poverty alleviation in Mexico's payments for ecosystem services program. *American Economic Journal: Economic Policy 7*(4), 1–40.

Almeida, C. A., A. C. Coutinho, J. C. D. M. Esquerdo, M. Adami, A. Venturieri, C. G. Diniz, N. Dessay, L. Durieux, and A. R. Gomes (2016). High spatial resolution land use and land cover mapping of the Brazilian Legal Amazon in 2008 using Landsat-5/tm and MODIS data. *Acta Amazonica 46(3)*, 291–302.

Andersen, L. E., C. W. Granger, E. Reis, D. Weinhold, and S. Wunder (2002). *The Dynamics of Deforestation and Economic Growth in the Brazilian Amazon.* Cambridge University Press, U.K.

Andrade, L. and A. L. S. Chagas (2016). Spillover effects of blacklisting policy in the Brazilian Amazon. Working Paper 2016-32, Department of Economics, University of São Paulo.

Arima, E. Y., P. Barreto, E. Araújo, and B. Soares-Filho (2014). Public policies can reduce tropical deforestation: Lessons and challenges from Brazil. *Land Use Policy 41*, 465–473.

Armstrong, T. B. and S. Shen (2015). Inference on optimal treatment assignments. Working Paper, Yale University.

Assunção, J., C. Gandour, and R. Rocha. (2017). Detering deforestation in the Amazon: Environmental monitoring and law enforcement. CPI Working Paper, Climate Policy Initiative.

Assunção, J. and R. Rocha (2019). Getting greener by going black: the effect of blacklisting municipalities on Amazon deforestation. *Environment and Development Economics 24*(2), 115–137.

Athey, S. and G. W. Imbens (2006). Identification and inference in nonlinear difference-in-differences models. *Econometrica 74(2)*, 431–97.

Athey, S. and G. W. Imbens (2017). The state of applied econometrics: Causality and policy evaluation. *Journal of Economic Perspectives 31(2)*, 3–32.

Athey, S. and S. Wager (2018). Efficient policy learning. Working Paper, Cornell University.

Baccini, A., S. J. Goetz, W. S. Walker, N. T. Laporte, M. Sun, D. Sulla-Menashe, J. Hackler, P. S. A. Beck, R. Dubayah, M. A. Friedl, S. Samanta, and R. A. Houghton (2012). Estimated carbon dioxide emissions from tropical deforestation improved by carbon-density maps. *Nature Climate Change 2*, 182–185.

Bhattacharya, D. and P. Dupas (2012). Inferring welfare maximizing treatment assignment under budget constraints. *Journal of Econometrics 167*(1), 168–196.

Bonan, G. B. (2008). Forests and climate change: Forcings, feedbacks, and the climate benefits of forests. *Science 320*(5882), 1444–1449.

Borner, J., S. Wunder, S. Wertz-Kanounnikoff, G. Hyman, and N. Nascimento (2014). Forest law enforcement in the Brazilian Amazon: Costs and income effects. *Global Environmental Change 29*, 294–305.

Brady, M. and E. Irwin (2011). Environmental effects of land use change revealed through spatially explicit empirical models: Recent developments and challenges ahead. *Environmental and Resource Economics 48(3)*, 487–509.

Brito, B., C. Souza Jr, and P. Amaral (2010). *Reducing emissions from deforestation at municipal level: a case study of Paragominas*, Chapter 8. Brasil: Defra, British Embassy Brasília.

Brollo, F. and T. Nannicini (2012). Tying your enemy's hands in close races: The politics of federal transfers in Brazil. *American Political Science Review 106*(4), 742–761.

Burgess, R., F. J. M. Costa, and B. A. Olken (2017). The power of the state: National borders and the deforestation of the Amazon. Working Paper, London School of Economics.

Burgess, R., M. Hansen, B. A. Olken, P. Potapov, and S. Sieber (2012). The political economy of deforestation in the tropics. *The Quarterly Journal of Economics 127(4)*, 1707–1754.

Burke, M., M. Craxton, C. D. Kolstad, C. Onda, H. Allcott, E. Baker, L. Barrage, R. Carson, K. Gillingham, J. Graff-Zivin, M. Greenstone, G. Heal, S. Hsiang, B. Jones, D. L. Kelly, K. Kopp, M. Kotchen, R. Mendelsohn, K. Meng, G. Metcalf, J. Moreno-Cruz, R. Pindyck, S. Rose, I. Rudik, J. Stock, and R. S. J. Tol (2016). Opportunities for advances in climate change economics. *Science 352*, 292–293.

Callaway, B. and P. H. C. Sant'Anna (2018). Difference-in-differences with multiple time periods and an application on the minimum wage and employment. Working Paper, Vanderbilt University.

Camara, G., D. M. Valeriano, and J. V. Soares (2006). Metodologia para o cálculo da taxa anual de desmatamento na Amazônia Legal. *INPE, São José dos Campo*.

Chalfin, A. and J. McCrary (2017). Criminal deterrence: A review of the literature. *Journal of Economic Literature 55(1)*, 5–48.

Chamberlain, G. (2011). *Bayesian Aspects of Treatment Choice*, Chapter 1, pp. 11–39. New York: Oxford University Press.

Cisneros, E., J. Hargrave, and K. Kis-Katos (2013). Unintended consequences of anti-corruption strategies: Public fiscal audits and deforestation in the Brazilian Amazon. Working Paper, BIOECON Conference.

Cisneros, E., S. L. Zhou, and J. Börner (2015). Naming and shaming for conservation: Evidence from the Brazilian Amazon. *PloS One 10*(9), e0136402.

Davidson, E. A., A. C. de Arajo, P. Artaxo, J. K. Balch, I. F. Brown, M. M. C. Bustamante, M. T. Coe, R. S. DeFries, M. Keller, M. Longo, J. W. Munger, W. Schroeder, B. S. Soares-Filho, C. M. Souza, and S. C. Wofsy (2012, January). The Amazon Basin in transition. *Nature 481*, 321–328.

de Chaisemartin, C. and X. D'Haultfoeuille (2018). Two-way fixed effects estimators with heterogeneous treatment effects. Working Paper, UC Santa Barbara.

Deep, K., K. P. Singh, M. Kansal, and C. Mohan (2009). A real coded genetic algorithm for solving integer and mixed integer optimization problems. *Applied Mathematics and Computation 212(2)*, 505–518.

Dehejia, R. H. (2005). Program evaluation as a decision problem. *Journal of Econometrics 125*(1), 141–173.

Foley, J. A., R. DeFries, G. P. Asner, C. Barford, G. Bonan, S. R. Carpenter, F. S. Chapin, M. T. Coe, G. C. Daily, H. K. Gibbs, J. H. Helkowski, T. Holloway, E. A. Howard, C. J. Kucharik, C. Monfreda, J. A. Patz, I. C. Prentice, N. Ramankutty, and P. K. Snyder (2005). Global consequences of land use. *Science 309*(5734), 570–574.

Friesen, L. (2003). Targeting enforcement to improve compliance with environmental regulations. *Journal of Environmental Economics and Management 46*(1), 72 – 85.

Gandour, C. C. (2018). *Forest Wars: A Trilogy on Combating Deforestation in the Brazilian Amazon*. Ph. D. thesis.

Ginther, D. K. (2000). Alternative estimates of the effect of schooling on earnings. *The Review of Economics and Statistics 82(1)*, 103–116.

Godar, J., T. A. Gardner, E. J. Tizado, and P. Pacheco (2014). Actor-specific contributions to the deforestation slowdown in the Brazilian Amazon. *Proceedings of the National Academy of Sciences 111(43)*, 15591–15596.

Goodman-Bacon, A. (2018). Difference-in-differences with variation in treatment timing. NBER Working Paper 25018, National Bureau of Economic Research.

Gray, W. B. and J. P. Shimshack (2011). The effectiveness of environmental monitoring and enforcement: A review of the empirical evidence. *Review of Environmental Economics and Policy 5*, 3–24.

Greenstone, M. and R. Hanna (2014). Environmental regulations, air and water pollution, and infant mortality in India. *American Economic Review 104(10)*, 3038–3072.

Greenstone, M. and B. K. Jack (2015). Envirodevonomics: A research agenda for an emerging field. *Journal of Economic Literature 53*(1), 5–42.

Greenstone, M., E. Kopits, and A. Wolverton (2013). Developing a social cost of carbon for US regulatory analysis: A methodology and interpretation. *Review of Environmental Economics and Policy 7(1)*, 23–46.

Han, S. (2019). Identification in nonparametric models for dynamic treatment effects. Working Paper, University of Texas at Austin.

Hanna, R. and B. A. Olken (2018). Universal basic incomes versus targeted transfers: Anti-poverty programs in developing countries. *Journal of Economic Perspectives 32*(4), 201–26.

Harding, T., J. Herzberg, and K. Kuralbayeva (2018). Environmental regulation and commodity prices: Evidence from deforestation in Brazil. Working Paper, London School of Economics.

Hargrave, J. and K. Kis-Katos (2013). Economic causes of deforestation in the Brazilian Amazon: A panel data analysis for the 2000s. *Environmental & Resource Economics 54(4)*, 471–494.

Harrington, W. (1988). Enforcement leverage when penalties are restricted. *Journal of Public Economics 37*(1), 29 – 53.

Havnes, T. and M. Mogstad (2015). Is universal child care leveling the playing field? *Journal of Public Economics 127C*, 100–114.

Heckman, J. J. and E. Vytlacil (2005). Structural equations, treatment effects, and econometric policy evaluation. *Econometrica 73*(3), 669–738.

Hirano, K. and J. R. Porter (2009). Asymptotics for statistical treatment rules. *Econometrica 77*(5), 1683–1701.

Imbens, G. W. and C. F. Manski (2004). Confidence intervals for partially identified parameters. *Econometrica 72*, 1845–1857.

INPE (2017). Monitoramento da floresta amazônica brasileira por satélite – projeto prodes. *Available at http://www.obt.inpe.br/prodes/*.

IPCC (2013). *Climate Change 2013: The Physical Science Basis.* Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Jack, B. K. and S. Jayachandran (2018). Self-selection into payments for ecosystem services programs. *Proceedings of the National Academy of Sciences*.

Jayachandran, S., J. de Laat, E. F. Lambin, C. Y. Stanton, R. Audy, and N. E. Thomas (2017). Cash for carbon: A randomized trial of payments for ecosystem services to reduce deforestation. *Science 357(6348)*, 267–273.

Kasy, M. (2016). Partial identification, distributional preferences, and the welfare ranking of policies. *Review of Economics and Statistics 98*(1), 111–131.

Kitagawa, T. and A. Tetenov (2018). Who should be treated? Empirical welfare maximization methods for treatment choice. *Econometrica 86*(2), 591–616.

Koch, N., E. K. H. J. Ermgassen, J. Wehkamp, F. Oliveira, and G. Schwerhoff (2018). Agricultural productivity and forest conservation: Evidence from the Brazilian Amazon. Working Paper, University of Cambridge.

Kottelenberg, M. J. and S. F. Lehrer (2017). Targeted or universal coverage? Assessing heterogeneity in the effects of universal child care. *Journal of Labor Economics 35*(3), 609–653.

Lee, D. S. (2009). Training, wages, and sample selection: Estimating sharp bounds on treatment effects. *Review of Economic Studies 76(3)*, 1071–1102.

Lubowski, R. N., A. J. Plantinga, and R. N. Stavins (2006). Land-use change and carbon sinks: Econometric estimation of the carbon sequestration supply function. *Journal of Environmental Economics and Management 51*(2), 135–152.

Maia, H., J. Hargrave, J. J. Gómez, and M. Röper (2011). Avaliação do plano de ação para prevenção e controle do desmatamento na Amazonia legal: PPCDAm 2007–2010.

Manski, C. (2000). Identification problems and decisions under ambiguity: Empirical analysis of treatment response and normative analysis of treatment choice. *Journal of Econometrics 95*(2), 415–442.

Manski, C. (2003). *Partial Identification of Probability Distributions.* New York: Springer-Verlag.

Manski, C. (2004). Statistical treatment rules for heterogeneous populations. *Econometrica 72(4)*, 1221–1246.

Manski, C. (2005a). *Social Choice with Partial Knowledge of Treatment Response.* Princeton University Press.

Manski, C. F. (2005b). Optimal search profiling with linear deterrence. *American Economic Review 95*(2), 122–126.

Manski, C. F. (2006). Search profiling with partial knowledge of deterrence. *The Economic Journal 116*(515), F385–F401.

Mason, C. F. and A. J. Platinga (2013). The additionality problem with offsets: Optimal contracts for carbon sequestration in forests. *Journal of Environmental Economics and Management 66*, 1–14.

Matsuura, K. and C. Willmott (2012). Terrestrial precipitation: 1900–2010 gridded monthly time series (1900 - 2010). *University of Delaware. http://climate.geog.udel.edu/climate/*.

Melly, B. and G. Santangelo (2015). The changes-in-changes model with covariates. Working Paper, University of Bern.

Michalski, F., D. Norris, and C. A. Peres (2010). No return from biodiversity loss. *Science 329 (5997)*, 1282.

MMA, M. d. M. A. (2009). Plano de ação para prevenção e controle do desmatamento na Amazonia Legal (PPCDAm). 2a fase (2009-2011), rumo ao desmatamento ilegal zero.

Nordhaus, W. (2014). Estimates of the social cost of carbon: Concepts and results from the DICE–2013R model and alternative approaches. *Journal of the Association of Environmental and Resource Economists 1(1/2)*, 273–312.

Pailler, S. (2018). Re-election incentives and deforestation cycles in the Brazilian Amazon. *Journal of Environmental Economics and Management 88*, 345–365.

Pattanyak, S., S. Wunder, and P. Ferraro (2010). Show me the money: Do payments supply environmental services in developing countries? *Review of Environmental Economics and Policy 4*, 254–274.

Pfaff, A. (1999). What drives deforestation in the Brazilian Amazon? Evidence from satellite and socioeconomic data. *Journal of Environmental Economics and Management 37*(1), 26–43.

Simonet, G., J. Subervie, D. Ezzine-de Blas, M. Cromberg, and A. E. Duchelle (2019). Effectiveness of a REDD+ project in reducing deforestation in the Brazilian Amazon. *American Journal of Agricultural Economics 101*(101), 211–229.

Souza-Rodrigues, E. (2019). Deforestation in the Amazon: A unified framework for estimation and policy analysis. *Review of Economic Studies, forthcoming*.

Stavins, R. N. (1999). The costs of carbon sequestration: A revealed-preference approach. *American Economic Review 89(4)*, 994–1009.

Stern, N. (2007). The economics of climate change: The Stern review. *New York: Cambridge University Press*.

Stoye, J. (2009). Minimax regret treatment choice with finite samples. *Journal of Econometrics 151*(1), 70–81.

Table 1: Summary Statistics for 2007 Cross-Section by Priority Status

	Total Sample ($N = 490$)		Treated Group ($N = 35$)		Untreated Group ($N = 455$)	
	Mean	SD	Mean	SD	Mean	SD
Land Use (km^2)						
Deforested Area	21	60	148	169	12	20
Cumulative Deforested Area	1,270	1,436	4,413	2,437	1,028	978
Forested Area	6,499	15,507	15,990	27,549	5,769	13,953
Municipal Area	8,726	16,716	21,815	29,409	7,719	14,899
Deforested Share (%)	1.21	1.91	1.72	1.41	1.17	1.94
Policy Measures						
Number of Alerts	57	213	509	616	22	60
Fines Issued	9	19	40	44	7	14
Share of Protected Area (%)	28	33	22	22	28	34
Agriculture and Ranching						
GDP (million Reais)	179	1,013	180	399	179	1,046
Agricultural GDP (million Reais)	19	24	39	22	18	24
Cattle (thousands)	105	148	363	290	85	108
Crop Area	109	455	289	563	95	443
Total Rural Credit (million Reais)	7	13	19	16	6	12
Other Variables						
Rainfall (mm)	2,206	613	1,948	195	2,227	633
Temperature ($^\circ C$)	26	1	26	1	26	1
Carbon Stock in Forested Areas (tC/ha)	189	68	212	35	187	69
Carbon Stock in Deforested Areas (tC/ha)	117	48	101	23	118	49

Notes: This table reports municipality-level means and standard deviations (SD) for the variables used in the empirical analysis, for the year 2007. A unit of observation is a municipality in the Brazilian Amazon. Land use data are drawn from satellite images (areas are measured in square kilometres). Deforested Area measures incremental deforestation during the year; Cumulative Deforested Area adds past deforestation up to and including 2007; Forested Area measures the total area covered by forests at the beginning of the year; Deforested Share divides incremental deforestation in 2007 by the forested area at the beginning of the year. Share of Protected Area is the proportion of the municipal area that is under legal protection (either indigenous lands or conservation units). GDP consists of the municipalities' total GDPs. Agricultural GDP includes crop and livestock production. All monetary amounts are expressed in December 2011 Reais. Annual rainfall is measured in millimetres (mm), while annual temperature is measured in degrees Celsius ($^\circ C$). Carbon stocks are measured in tons of carbon per hectare (tC/ha).

Table 2: Summary Statistics for 2007 Cross-Section by Group

	Treated Group ($N = 35$)		Spillover Group ($N = 24$)		Control Group ($N = 431$)	
	Mean	SD	Mean	SD	Mean	SD
Land Use (km^2)						
Deforested Area	148	169	49	36	9	17
Cumulative Deforested Area	4,413	2,437	2,832	941	927	878
Forested Area	15,990	27,549	6,043	11,378	5,754	14,093
Total Area	21,815	29,409	10,408	12,499	7,569	15,020
Deforested Share (%)	1.72	1.41	2.03	2.14	1.12	1.92
Policy Measures						
Number of Alerts	509	616	141	144	15	44
Fines Issued	40	44	23	20	6	13
Share of Protected Area (%)	22	22	16	22	29	34
Agriculture and Ranching						
GDP (million Reais)	180	399	210	380	177	1,071
Agricultural GDP (million Reais)	39	22	41	46	16	21
Cattle (thousands)	363	290	248	126	76	99
Crop Area	289	563	250	396	87	444
Total Rural Credit (million Reais)	19	16	16	15	6	12
Other Variables						
Rainfall (mm)	1,948	195	2,038	293	2,237	645
Temperature ($^\circ C$)	26	1	26	2	26	1
Carbon Stock in Forested Areas (tC/ha)	212	35	203	45	187	70
Carbon Stock in Deforested Areas (tC/ha)	101	23	97	24	119	50

Notes: This table breaks the Untreated Group from Table 1 into Spillover and Control Groups. The Spillover Group consists of the untreated municipalities that (a) share a border with a treated municipality, and (b) have high levels of past deforestation – specifically, determined by whether the 'selection criteria' variables Z^1_{mt-1} and Z^2_{mt-1} exceed 70 percent of the thresholds values that the Brazilian government (implicitly) adopted in the selection rule; that is, whether $Z^1_{mt-1} \geq 0.7 \times 2,700$ km^2 and $Z^2_{mt-1} \geq 0.7 \times 220$ km^2. The format of the table is otherwise identical to Table 1.

Table 3: Difference-in-Differences Results

	(1) Log odds	(2) Log odds	(3) Log odds	(4) Log odds
Treated Group x Year=2009	-0.456**	-0.457**	-0.471**	-0.471**
	(0.150)	(0.147)	(0.152)	(0.150)
Treated Group x Year=2010	-0.928***	-0.887***	-0.964***	-0.922***
	(0.160)	(0.156)	(0.162)	(0.159)
Spillover Group x Year=2009			-0.283	-0.243
			(0.149)	(0.143)
Spillover Group x Year=2010			-0.682***	-0.624***
			(0.188)	(0.186)
Year=2006	-0.370***	-0.569***	-0.370***	-0.575***
	(0.0661)	(0.0994)	(0.0661)	(0.0997)
Year=2007	-0.467***	-0.564***	-0.467***	-0.565***
	(0.0659)	(0.0729)	(0.0660)	(0.0729)
Year=2009	-0.948***	-0.920***	-0.933***	-0.909***
	(0.0769)	(0.0763)	(0.0801)	(0.0794)
Year=2010	-0.672***	-0.730***	-0.636***	-0.698***
	(0.0806)	(0.0830)	(0.0839)	(0.0865)
Covariates	NO	YES	NO	YES
R^2	0.098	0.113	0.101	0.115
Observations	2450	2450	2450	2450

Notes: An observation is a municipality in the Brazilian Amazon. The dependent variable is the log odds ratio of deforestation shares. All regressions include municipality fixed effects. Robust standard errors in parentheses are clustered at the municipality level.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table 4: Pre-Treatment 'Common Trends' Test, 2006–2007

	(1) Log odds	(2) Log odds
Treated Group x Year=2007	-0.0890	-0.0911
	(0.140)	(0.143)
Spillover Group x Year=2007		-0.0300
		(0.137)
Year=2007	0.305*	0.308*
	(0.128)	(0.130)
Lagged Rainfall	0.122	0.122
	(0.0708)	(0.0709)
Lagged Rainfall Squared	-0.00399**	-0.00400**
	(0.00135)	(0.00136)
Lagged Temperature	0.458	0.460
	(0.418)	(0.418)
Share of Protected Areas	0.811	0.807
	(0.608)	(0.612)
Price of Beef Lagged	-0.00466	-0.00469
	(0.0321)	(0.0321)
Price of Crops Lagged	-3.303*	-3.303*
	(1.651)	(1.651)
Lagged GDP	-0.106	-0.106
	(0.375)	(0.376)
R^2	0.050	0.050
Observations	980	980

Notes: An observation is a municipality in the Brazilian Amazon. The dependent variable is the log odds ratio of deforestation shares. Rainfall is measured in millimetres (mm) and temperature is measured in degrees Celsius ($^\circ C$). The price of beef is a weighted average of international beef prices weighted by the ratio of head of cattle to municipal area. The price of crops is the price index based on a principal component analysis applied to individual weighted prices of the most predominant crops in the Brazilian Amazon (the weights are given by the share of the municipal area used to cultivate the crop). For all agricultural products, the weights are fixed in the period 2000–2001. Municipal GDP is measured in million Reais. All monetary amounts are expressed in December 2011 Reais. The coefficient on the constant term is omitted. All regressions include municipality fixed effects. Robust standard errors in parentheses are clustered at the municipality level. (The pre-treatment 'common trend' test covering the longer span (2003–2007) is presented in Appendix G.)
$^{*}\ p < 0.05, ^{**}\ p < 0.01, ^{***}\ p < 0.001.$

Table 5: Average and Cumulated Treatment Effects, CIC Model without Spillovers – Deforestation and Carbon Emissions

Average Treatment Effects: Deforestation			
	ATT	ATU	ATE
2009			
Baseline 2006	-21.61	[-4.74, -3.30]	[-5.94, -4.61]
	(-24.35, -18.88)	(-4.84, -3.21)	(-6.07, -4.50)
Baseline 2007	-24.91	[-4.70, -3.62]	[-6.15, -5.14]
	(-28.28, -21.55)	(-4.81, -3.52)	(-6.28, -5.01)
2010			
Baseline 2006	-50.94	[-7.40, -5.38]	[-10.51, -8.63]
	(-55.30, -46.58)	(-7.52, -5.28)	(-10.67, -8.49)
Baseline 2007	-53.93	[-7.42, -5.89]	[-10.74, -9.32]
	(-58.58, -49.29)	(-7.53, -5.79)	(-10.90, -9.17)
Cumulative 2009-2010			
Baseline 2006	-2539	[-5524, -3948]	[-8064, -6488]
	(-2765, -2314)	(-5618, -3873)	(-8190, -6377)
Baseline 2007	-2760	[-5514, -4327]	[-8274, -7086]
	(-3019, -2501)	(-5608, -4244)	(-8408, -6960)

Cumulated Treatment Effects, 2009-2010: Avoided Carbon Emissions			
	CTT	CTU	CTE
Emissions			
Baseline 2006	-30.17	[-58.45, -41.78]	[-88.62, -71.96]
	(-33.06, -27.29)	(-59.47, -40.98)	(-90.06, -70.70)
Baseline 2007	-32.80	[-58.34, -45.79]	[-91.14, -78.58]
	(-36.09, -29.50)	(-59.37, -44.90)	(-92.67, -77.15)
Value (U$ 20/tCO2)			
Baseline 2006	2.21	[3.06, 4.29]	[5.28, 6.50]
	(2.00, 2.42)	(3.00, 4.36)	(5.18, 6.60)
Baseline 2007	2.41	[3.36, 4.28]	[5.76, 6.68]
	(2.16, 2.65)	(3.29, 4.35)	(5.66, 6.80)

Notes: 95% confidence intervals are in parentheses. For ATT and CTT, the intervals are computed based on the standard i.i.d. nonparametric bootstrap, where the i.i.d. resampling occurs in the cross-sectional dimension. For ATU, ATE, CTU, and CTE, they are based on Imbens and Manski (2004). We implemented 500 bootstrap replications. Deforestation is measured in square kilometres. Emissions are measured in millions of tons of carbon. Values are measured in billion US$, assuming a social cost of carbon of US$ 20/tCO2. The calculation uses the fact that 1 tC = (44/12) tCO2.

Table 6: Average and Cumulated Treatment Effects, CIC Model with Spillovers – Deforestation and Carbon Emissions

Average Treatment Effects: Deforestation				
	ATT	ATU	ATS	ATE
2009				
Baseline 2006	-24.65	[-4.46, -2.80]	[-11.52, -11.51]	[-6.25, -4.78]
	(-27.59, -21.71)	(-4.56, -2.72)	(-13.74, -9.31)	(-6.38, -4.67)
Baseline 2007	-29.27	[-4.46, -3.29]	[-16.69, -16.69]	[-6.83, -5.80]
	(-33.02, -25.52)	(-4.56, -3.20)	(-20.05, -13.35)	(-6.98, -5.66)
2010				
Baseline 2006	-52.63	[-6.79, -4.97]	[-22.42, -15.28]	[-10.83, -8.88]
	(-57.16, -48.10)	(-6.90, -4.88)	(-24.97, -12.38)	(-10.99, -8.73)
Baseline 2007	-58.57	[-6.85, -5.57]	[-25.56, -18.43]	[-11.46, -9.98]
	(-63.49, -53.65)	(-6.96, -5.47)	(-28.95, -14.69)	(-11.63, -9.82)
Cumulative 2009-2010				
Baseline 2006	-2705	[-4849, -3347]	[-814, -643]	[-8368, -6695]
	(-2945, -2464)	(-4931, -3282)	(-915, -540)	(-8499, -6577)
Baseline 2007	-3074	[-4877, -3815]	[-1014, -843]	[-8965, -7732]
	(-3356, -2793)	(-4957, -3746)	(-1156, -702)	(-9107, -7596)

Cumulated Treatment Effects, 2009-2010: Avoided Carbon Emissions				
	CTT	CTU	CTS	CTE
Emissions				
Baseline 2006	-32.19	[-50.10, -34.59]	[-9.69, -7.64]	[-91.99, -74.42]
	(-35.28, -29.11)	(-50.99, -33.91)	(-11.00, -6.32)	(-93.48, -73.08)
Baseline 2007	-36.59	[-50.39, -39.43]	[-12.09, -10.04]	[-99.08, -86.06]
	(-40.21, -32.98)	(-51.26, -38.70)	(-13.92, -8.24)	(-100.73, -84.49)
Value (U$ 20/tCO2)				
Baseline 2006	2.36	[2.54, 3.67]	[0.56, 0.71]	[5.46, 6.75]
	(2.13, 2.59)	(2.49, 3.74)	(0.46, 0.81)	(5.36, 6.86)
Baseline 2007	2.68	[2.89, 3.70]	[0.74, 0.89]	[6.31, 7.27]
	(2.42, 2.95)	(2.84, 3.76)	(0.60, 1.02)	(6.20, 7.39)

Notes: 95% confidence intervals are in parentheses. For ATT and CTT, the intervals are computed based on the standard i.i.d. nonparametric bootstrap, where the i.i.d. resampling occurs in the cross-sectional dimension. For ATU, ATS, ATE, CTU, CTS, and CTE, they are based on Imbens and Manski (2004). We implemented 500 bootstrap replications. Deforestation is measured in square kilometres. Emissions are measured in millions of tons of carbon. Values are measured in billion US$, assuming assuming a social cost of carbon of US$ 20/tCO2. The calculation uses the fact that 1 tC = (44/12) tCO2.

Table 7: Tests based on Changes-in-Changes Model

	Placebo Test		'No Effect' Test		Stochastic Dominance Test	
	KS	CM	KS	CM	KS	CM
Unconditional	0.734	0.498	0.052	0.010	1.000	1.000
Residuals	0.660	0.592	0.002	0.000	1.000	1.000

Notes: The Placebo Test compares the factual and counterfactual distributions when we wrongly impose that the policy intervention was set in 2007; the null hypothesis states that the two distributions are equal to each other. The 'No Effect' Test is similar to the Placebo Test, but uses the correct timing of the intervention. The Stochastic Dominance Test assesses whether the counterfactual distribution is everywhere below the factual distribution. The test statistics are the Kolmogorov-Smirnov (KS) and the Cramer-von Mises (CM) statistics. We apply each test both on the log odds ratio of deforestation shares not conditioning on covariates (corresponding to the Unconditional row), and on the residuals, after partialling the covariates out (corresponding to the Residuals row). The cells present the p–values based on 500 bootstrap replications. The tests are proposed and developed by Melly and Santangelo (2015).

Table 8: Priority Status and Spillovers: Other Outcomes

	(1) Fines	(2) Alerts	(3) Total Credit	(4) PA Share
Treated Group x Year=2007	7.738	-405.3***	-2039.3	-0.00375
	(5.102)	(101.2)	(1978.3)	(0.00511)
Treated Group x Year=2008	28.05**	-95.12	4888.7	0.0121
	(8.480)	(82.75)	(2665.0)	(0.0156)
Treated Group x Year=2009	14.55	-398.0***	-4215.2	0.00893
	(7.791)	(111.3)	(2188.7)	(0.0160)
Treated Group x Year=2010	0.619	-605.1***	484.0	0.00860
	(6.216)	(101.9)	(3069.8)	(0.0160)
Spillover Group x Year=2007	-9.996	-83.02	-5113.9	-0.00866**
	(6.912)	(44.03)	(3141.2)	(0.00274)
Spillover Group x Year=2008	-1.841	-13.13	492.1	-0.00658*
	(8.940)	(45.02)	(2333.2)	(0.00314)
Spillover Group x Year=2009	-10.65	-81.95	-998.6	-0.0104**
	(10.30)	(52.45)	(3498.2)	(0.00374)
Spillover Group x Year=2010	-8.945	-110.7*	371.5	-0.0117**
	(7.057)	(50.00)	(2589.5)	(0.00385)
Year=2007	0.678	-37.13***	-361.8	0.0110**
	(1.644)	(9.282)	(581.6)	(0.00356)
Year=2008	4.783*	-5.210	823.8	0.0190***
	(2.265)	(9.780)	(745.9)	(0.00493)
Year=2009	2.654	-19.95	-383.4	0.0185***
	(2.807)	(10.65)	(690.1)	(0.00454)
Year=2010	2.073	-33.18***	1587.2*	0.0204***
	(2.213)	(9.808)	(693.5)	(0.00479)
Lagged Rainfall	-0.568	8.346**	-431.7*	0.000730
	(0.698)	(2.989)	(185.0)	(0.000791)
Lagged Rainfall Squared	0.000198	-0.106*	8.530**	-0.0000219
	(0.0106)	(0.0415)	(2.987)	(0.0000142)
Lagged Temperature	-2.370	29.42**	-274.3	-0.00663
	(1.460)	(9.251)	(369.1)	(0.00444)
Price of Beef Lagged	-0.142	-0.335	-15.04	-0.000849***
	(0.0883)	(0.395)	(33.51)	(0.000206)
Price of Crops Lagged	15.91*	120.7*	-17437.6***	-0.0233*
	(7.591)	(54.76)	(2503.8)	(0.00932)
Lagged GDP	-17.01	-49.47	-410.7	-0.00840
	(9.220)	(45.18)	(3164.6)	(0.00716)
R^2	0.073	0.297	0.094	0.052
Observations	2450	2450	2325	2450

Notes: This table shows estimates from regressions of the outcome variables listed at the top of each column on Priority Status and Spillovers indicators, interacted with time dummies, along with other observables. An observation is a municipality in the Brazilian Amazon. The dependent variables are (1) the number of fines, (2) the number of alerts, (3) the total rural credit, and (4) the share of protected areas. Rainfall is measured in millimetres (mm) and temperature is measured in degrees Celsius ($^{\circ}C$). The price of beef is a weighted average of international beef prices weighted by the ratio of head of cattle to municipal area. The price of crops is the price index based on a principal component analysis applied to individual weighted prices of the most predominant crops in the Brazilian Amazon (the weights are given by the share of the municipal area used to cultivate the crop). For all agricultural products, the weights are fixed in the period 2000–2001. Municipal GDP is measured in million Reais. All monetary amounts are expressed in December 2011 Reais. The coefficient on the constant term is omitted. All regressions include municipality fixed effects. Robust standard errors in parentheses are clustered at the municipality level.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 9: Percent Correctly Predicted, Ex-post Deforestation Optimal Rule without Spillovers

Constraint:	*Total Area*			*Number of Municipalities*		
	Optimal			*Optimal*		
	0	1		0	1	
Observed			*Percent Correct*			*Percent Correct*
0	382	73	83.9	439	16	96.4
1	7	28	80.0	16	19	54.3
		Overall	83.7		*Overall*	93.5

Notes: The baseline year is 2006.

Table 10: Comparing Ex-Post Optimal, Priority, and Randomly Selected Lists – without Spillovers

Constraint:	*Total Area*		*Number of Municipalities*			
	Observed vs Optimal		Observed vs Optimal		Random vs Optimal	
	Ratio	Value	Ratio	Value	Ratio	Value
Total Deforestation						
Baseline 2006	1.06	-	1.06	-	1.23	-
Baseline 2007	1.06	-	1.06	-	1.25	-
Total Carbon Emissions						
Baseline 2006	1.05	562	1.07	870	1.26	2,704
Baseline 2007	1.05	620	1.08	957	1.29	2,951

Notes: 'Ratio' divides total deforestation (total emissions) evaluated at the observed list by the ex-post optimal total deforestation (total emissions). 'Value' takes their difference. Values are measured in million US$, assuming a social cost of carbon of US$ 20/tCO2.

Table 11: Percent Correctly Predicted, Ex-post Deforestation Optimal Rule with Spillovers

Constraint:	*Total Area*			*Number of Municipalities*		
	Optimal			*Optimal*		
	0	1		0	1	
Observed			*Percent Correct*			*Percent Correct*
0	365	90	80.2	436	19	95.8
1	12	23	65.7	19	16	45.7
		Overall	79.2		*Overall*	92.2

Notes: The baseline year is 2006.

Table 12: Comparing Ex-Post Optimal, Priority, and Randomly Selected Lists – with Spillovers

Constraint:	*Total Area*		*Number of Municipalities*			
	Observed vs Optimal		Observed vs Optimal		Random vs Optimal	
	Ratio	Value	Ratio	Value	Ratio	Value
Total Deforestation						
Baseline 2006	1.09	-	1.07	-	1.29	-
Baseline 2007	1.14	-	1.10	-	1.39	-
Total Carbon Emissions						
Baseline 2006	1.08	897	1.08	934	1.34	2,651
Baseline 2007	1.14	1,445	1.12	1,314	1.44	3,122

Notes: 'Ratio' divides total deforestation (total emissions) evaluated at the observed list by the ex-post optimal total deforestation (total emissions). 'Value' takes their difference. Values are measured in million US$, assuming a social cost of carbon of US$ 20/tCO2.

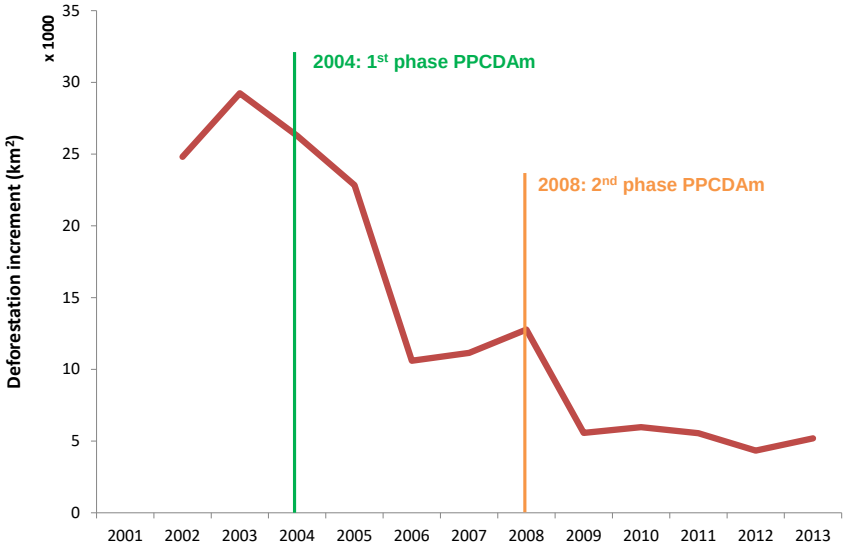


Figure 1: Incremental Deforestation and Key Policy Changes

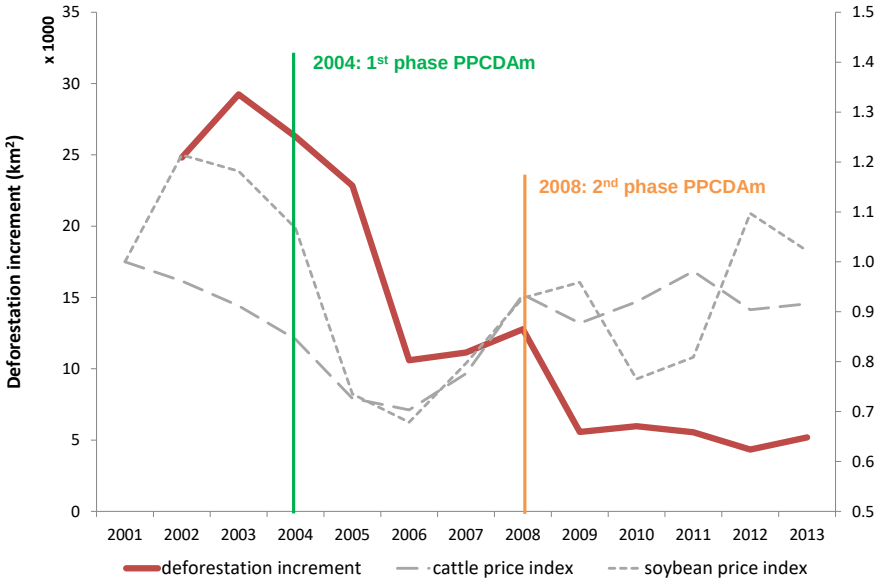


Figure 2: Incremental Deforestation and Agricultural Commodity Prices, by year

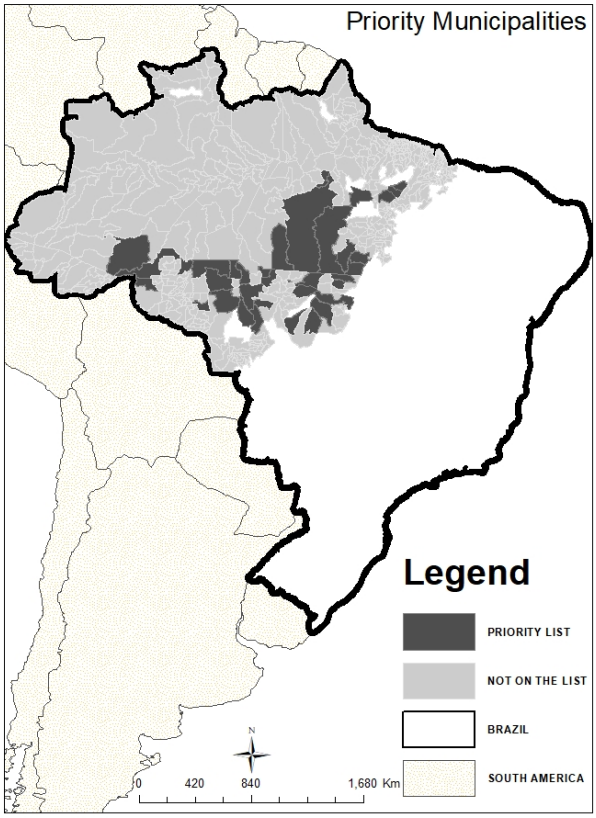
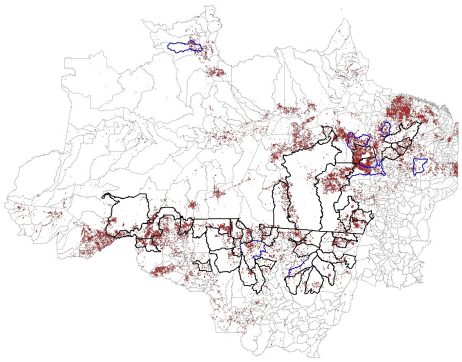
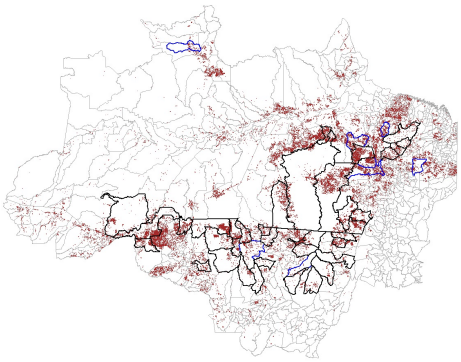


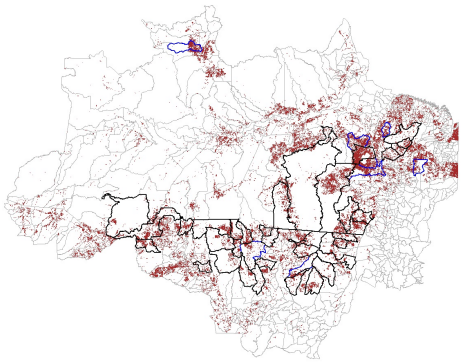
Figure 3: Map of Brazil, Amazonia, and the Location of the Priority List



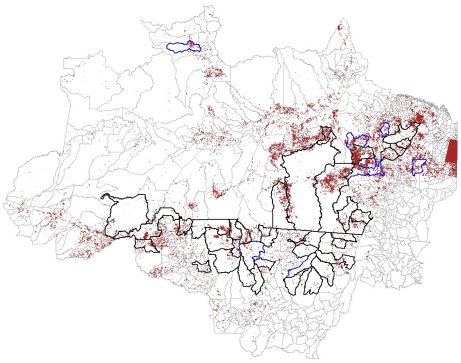
(a) Deforestation, 2006



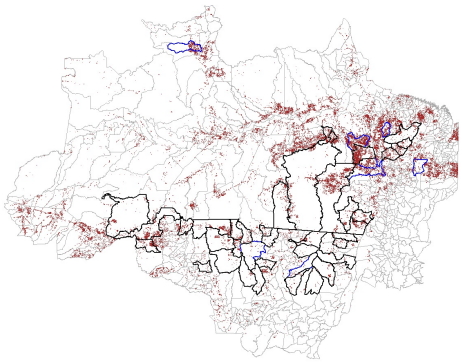
(b) Deforestation, 2007



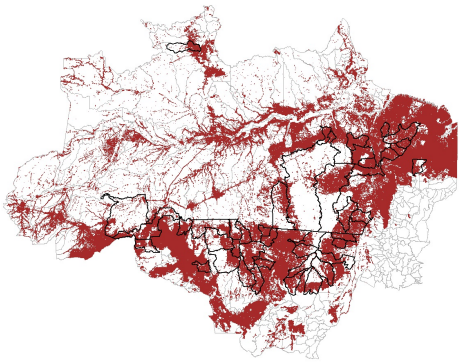
(c) Deforestation, 2008



(d) Deforestation, 2009

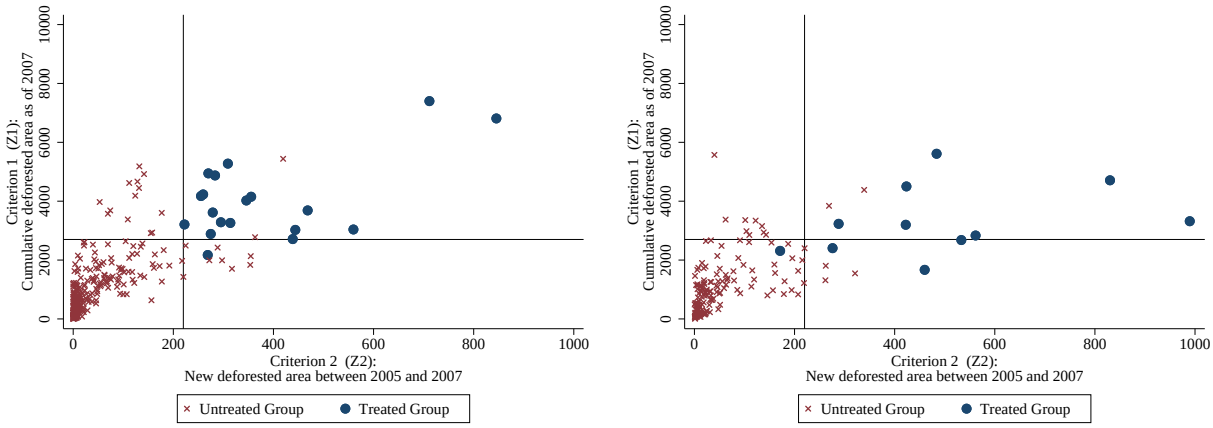


(e) Deforestation, 2010



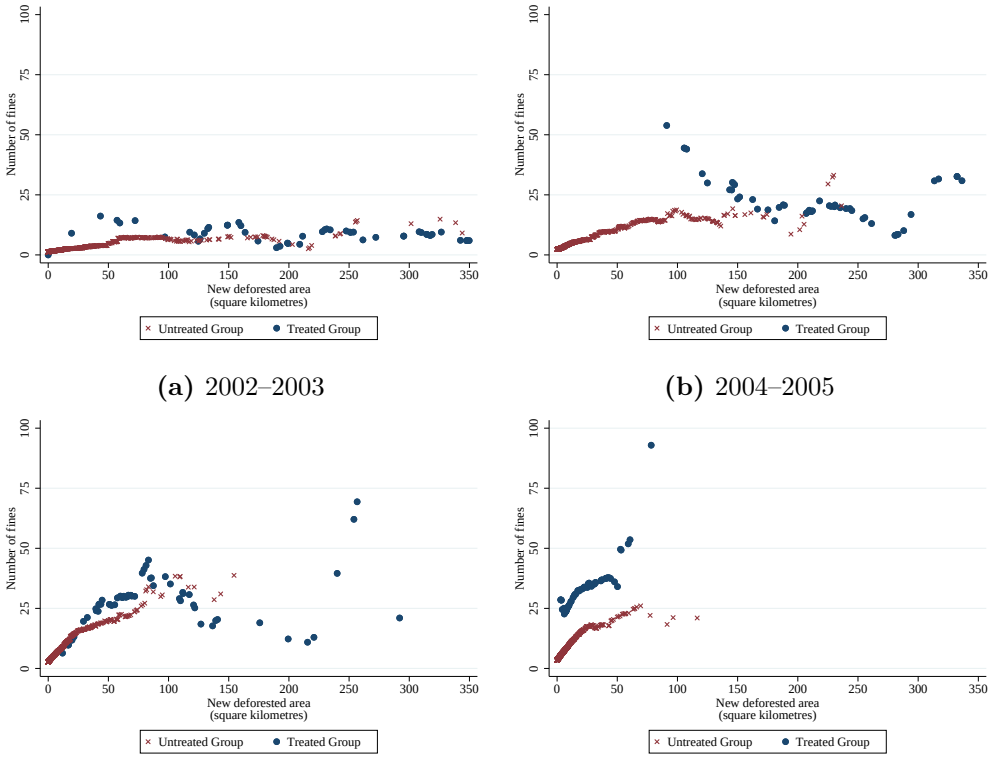
(f) Cumulative Deforestation by 2010

Figure 4: Map of Deforestation between 2006 and 2010 (with Priority Municipalities overlaid)



(a) Combinations of Z^1 and Z^2, given $Z^3 = 0$ **(b)** Combinations of Z^1 and Z^2, given $Z^3 = 1$

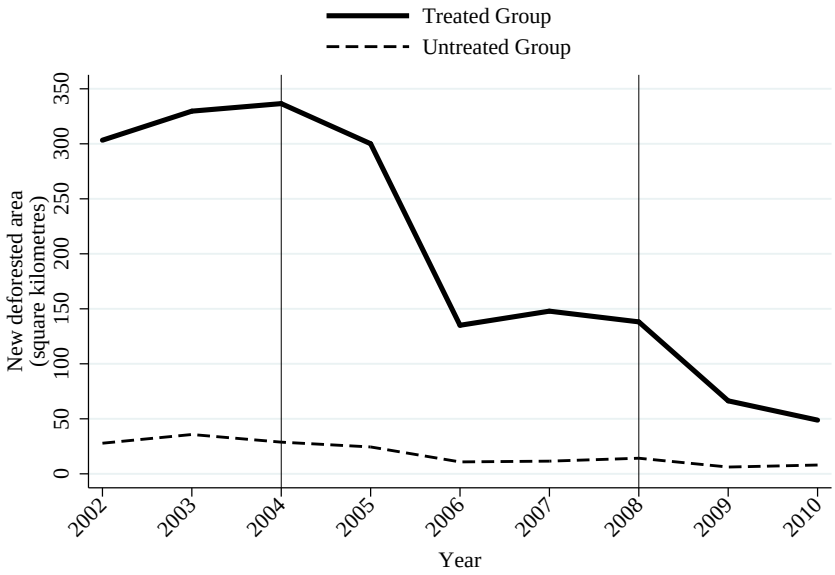
Figure 5: Selection onto Priority List in 2008 – Combinations of Criteria Variables (Z_{mt-1})



(a) 2002–2003 **(b)** 2004–2005

(c) 2006–2007 **(d)** 2009–2010

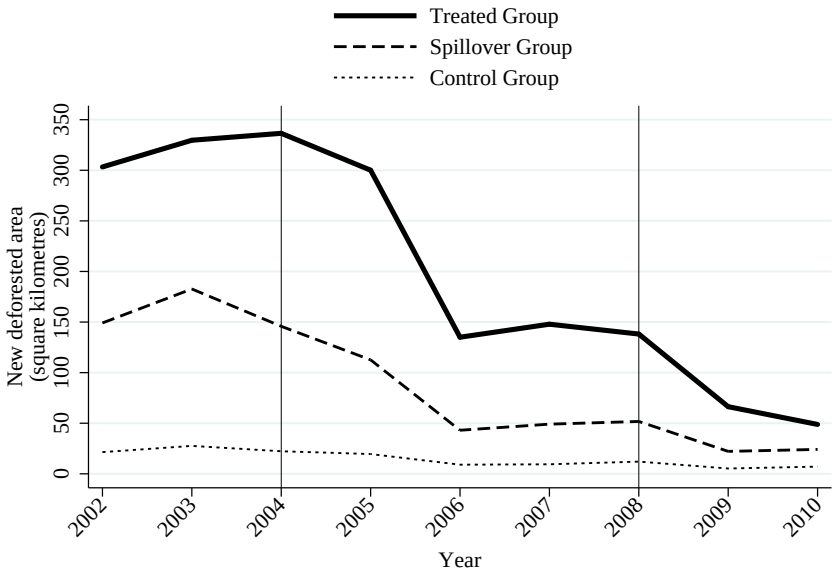
Figure 6: Fines Issued for Environmental Crimes, by Newly Deforested Area



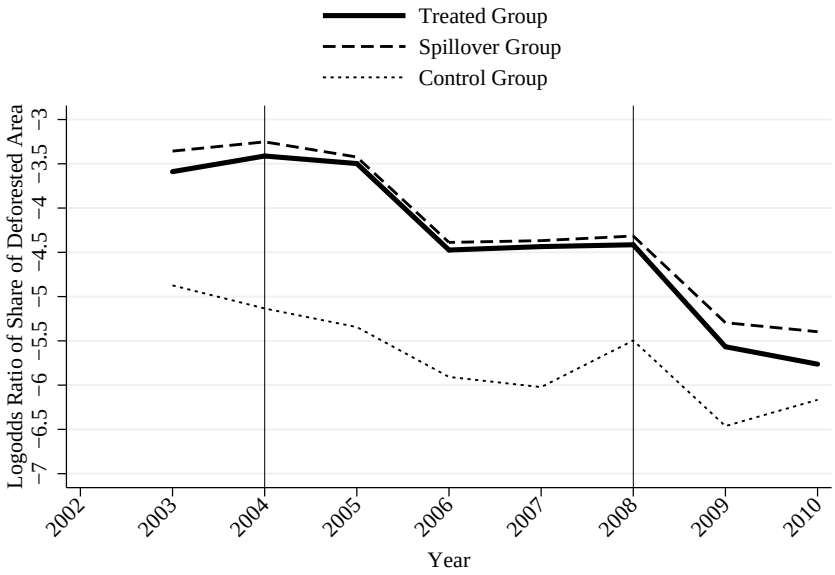
(a) Within-group average of newly deforested area, by year

(b) Within-group average of log odds ratio of deforestation shares, by year (using 2002 as the base to construct "remaining forest" in the share)

Figure 7: Evolution of Deforestation by Priority Status: level and log odds ratio of shares

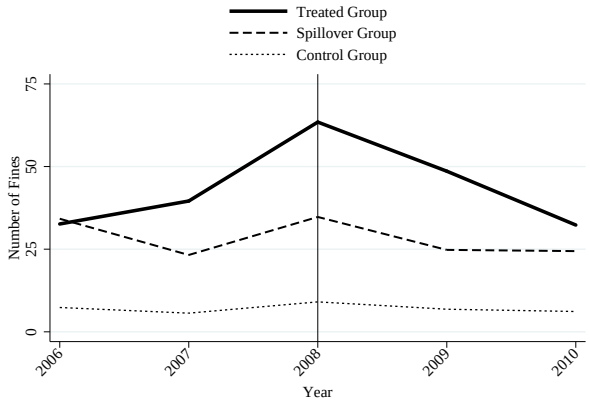


(a) Within-group average of newly deforested area, by year

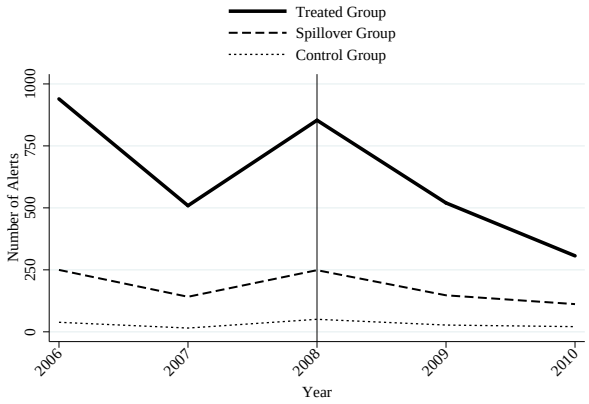


(b) Within-group average of log odds ratio of deforestation shares, by year (using 2002 as the base to construct "remaining forest" in the share)

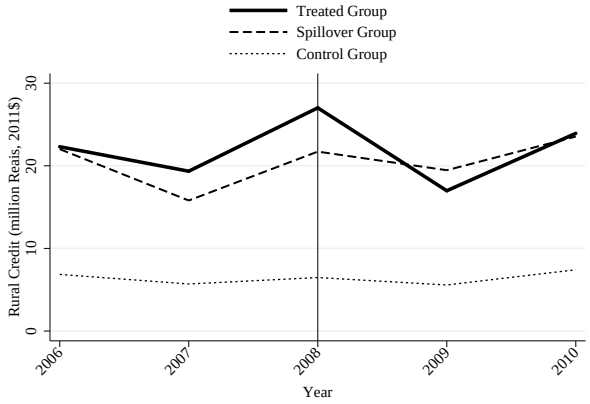
Figure 8: Evolution of Deforestation by Group: level and log odds ratio of shares



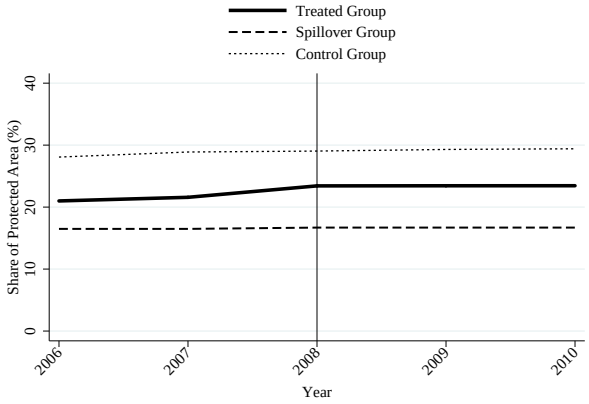
(a) Number of fines, by group and year



(b) Number of alerts, by group and year

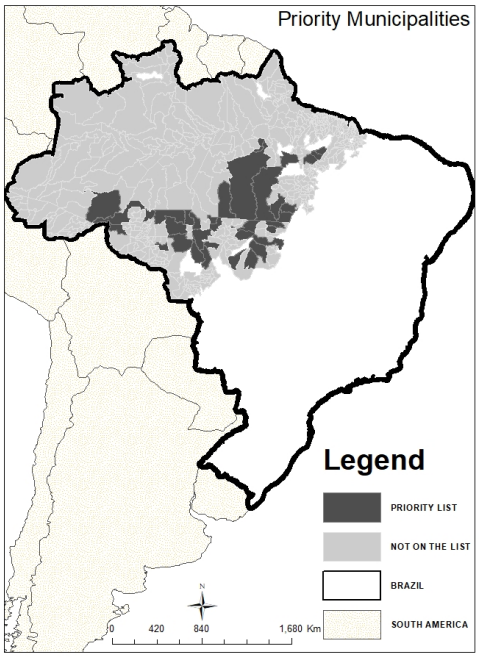


(c) Total rural credit, by group and year

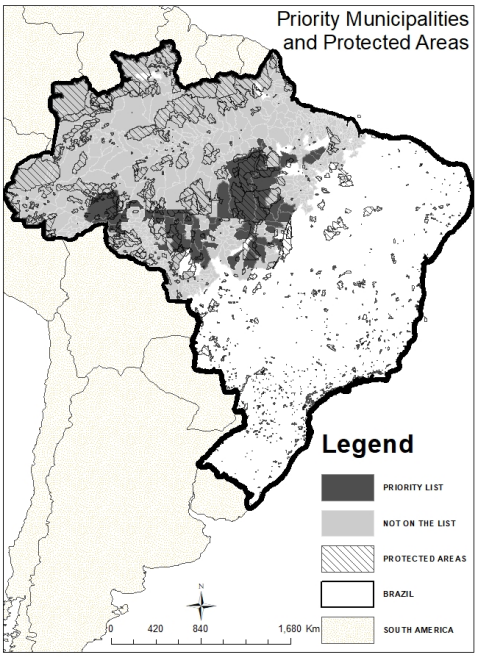


(d) Share of PAs, by group and year

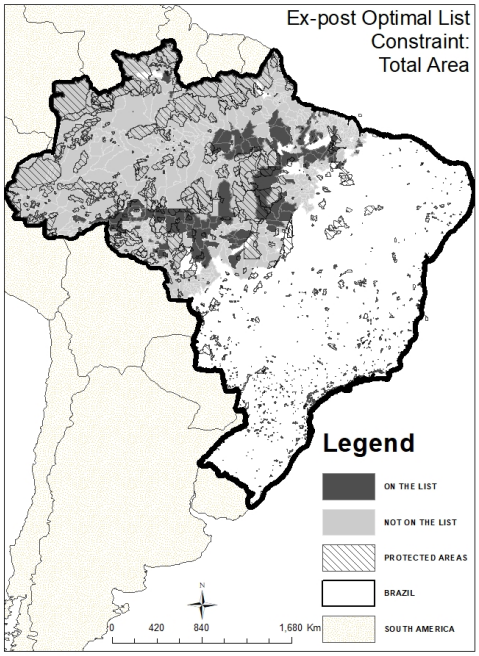
Figure 9: Evolution of Outcome Variables, by Group: Alerts, Fines, Rural Credit, and PAs



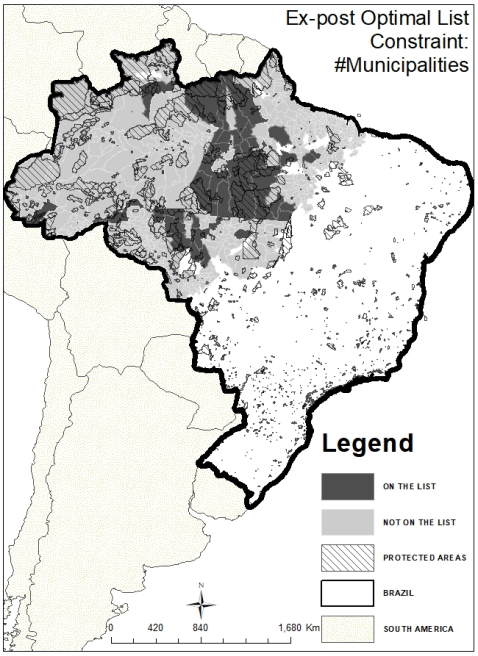
(a) Priority List



(b) Priority List and Protected Areas

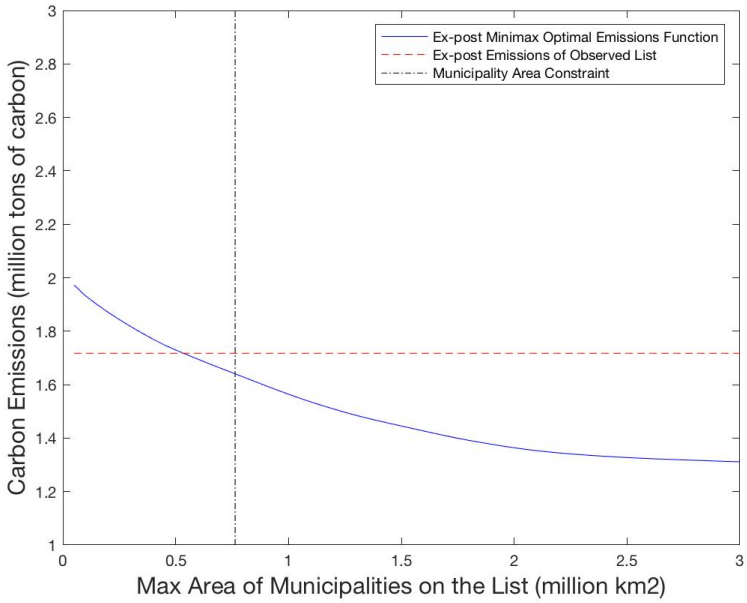


(c) Optimal List based on Total Area

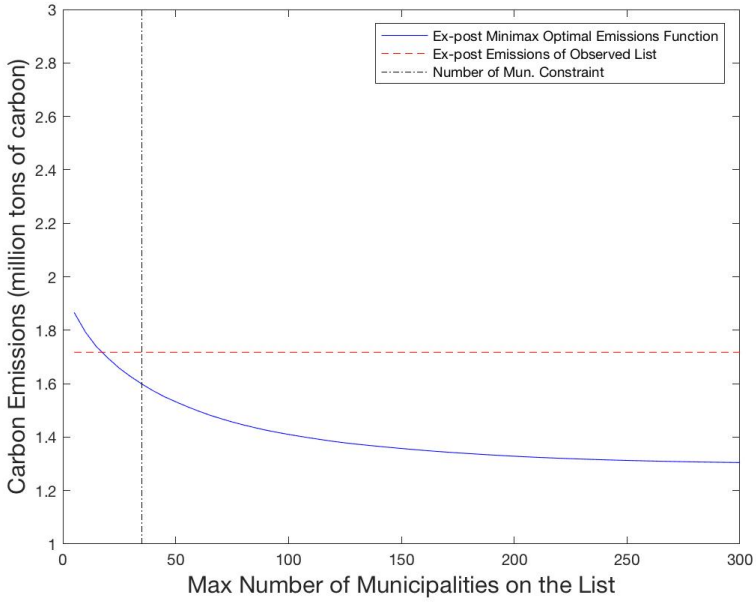


(d) Optimal List based on Number of Municipalities

Figure 10: Location of Priority List, Protected Areas, and Ex-post Optimal Lists without Spillovers

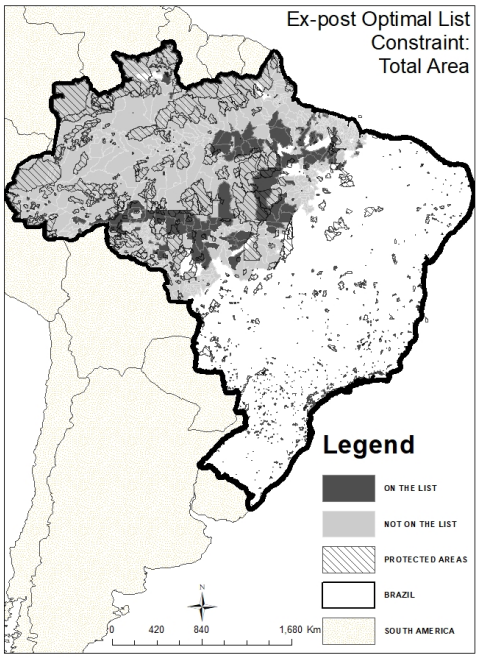


(a) Constraint: Total Area

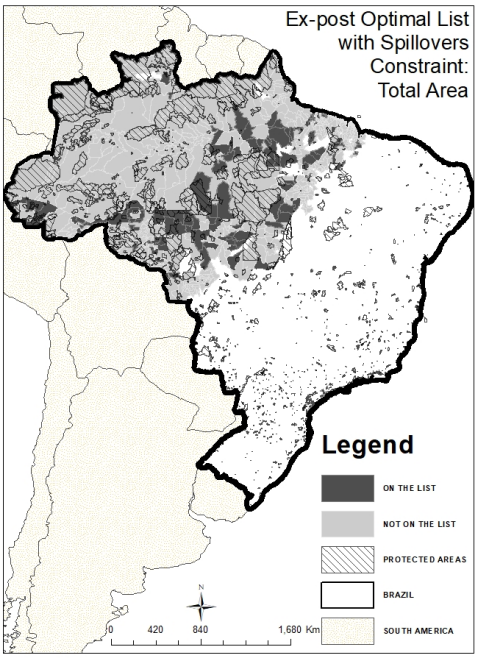


(b) Constraint: Number of Municipalities

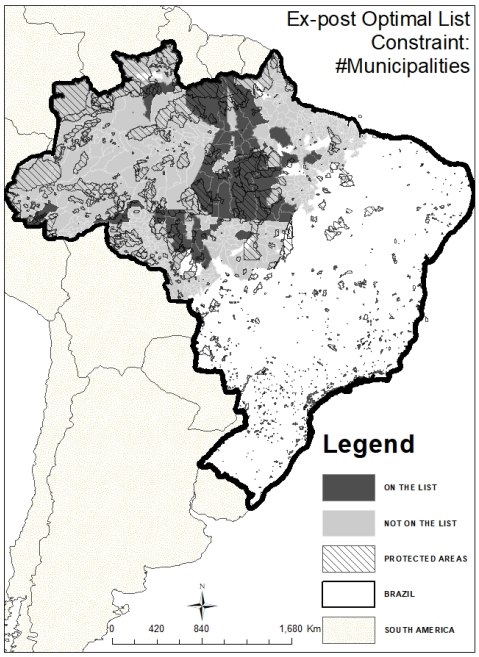
Figure 11: Ex-Post Minimax Carbon Emissions, when Varying the Constraints



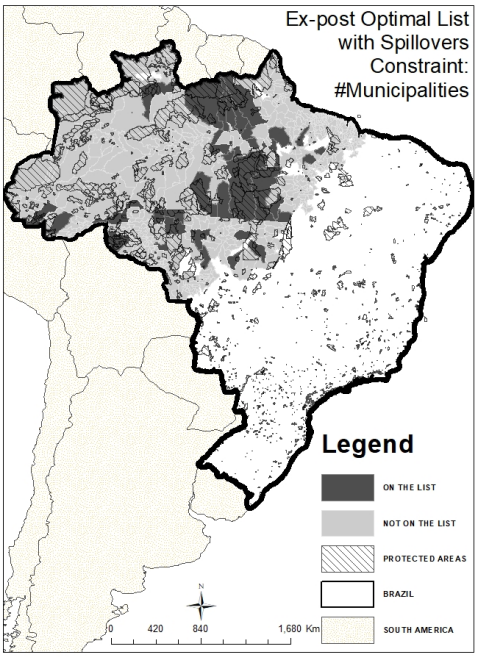
(a) No Spillover Optimal List based on Total Area



(b) Spillover Optimal List based on Total Area



(c) No Spillover Optimal List based on Number of Municipalities



(d) Spillover Optimal List based on Number of Municipalities

Figure 12: Location of Ex-post Optimal Lists without and with Spillover Effects

A Appendix: Statistical Treatment Rules

This brief appendix reviews a growing literature that analyzes statistical treatment rules in econometrics, including papers by Manski (2004, 2005a), Stoye (2009), Hirano and Porter (2009), Bhattacharya and Dupas (2012), Kasy (2016), and Kitagawa and Tetenov (2018), among others. It places the counterfactual portion of our study in that context.

In general, statistical assignment rules depend on three factors: the decision maker's objective function, the identification and estimation of conditional treatment effects conditioning on covariates, and constraints on the class of allowable policies (such as budget or capacity constraints).

Almost all papers in this area assume a utilitarian objective function, given by the sum of expected potential outcomes of the individuals in the targeted population. None of them allows for interactions (or spillovers) among individuals or treated groups in the population (that is, violations of SUTVA – the 'Stable Unit Treatment Value Assumption'). In contrast, our analysis does allow for such interactions.

One of the few empirical applications is the important study by Dehejia (2005), examining the Greater Avenues for Independence (GAIN) program, which began in California in 1986 and had the aim of increasing employment and earnings among welfare (AFDC) recipients. He considers a decision maker choosing whether to assign a welfare recipient into GAIN or AFDC, based on the individual's (predictive) distribution of future earnings under GAIN with the distribution of future earnings under AFDC. In the process, he solves for a Bayes decision rule given a prior distribution for model parameters of the treatment effects (see also Chamberlain, 2011), and finds that individualized assignment rules can indeed raise the average impact of the program by exploiting treatment-impact heterogeneity.

Several studies (including Hirano and Porter, 2009; Bhattacharya and Dupas, 2012; Armstrong and Shen, 2015) consider the statistical properties of the *plug-in* approach to empirical treatment choice. This line of work assumes that treatments in the data were assigned independently of the potential outcomes, conditional on observables (i.e., they impose the unconfoundedness, or selection-on-observables, assumption). Treatment effects are then point-identified and estimated using flexible methods.[60]

Instead of using plug-in assignment rules, Kitagawa and Tetenov (2018) propose estimating a treatment policy by maximizing the sample analog of average social welfare over a class of candidate

[60] Specifically, Hirano and Porter (2009) investigate the local asymptotic minimax optimality of the estimated plug-in rule. Bhattacharya and Dupas (2012) study the asymptotic properties of the plug-in rule under an aggregate budget constraint, and develop confidence intervals for the value of the planner's objective function; they also illustrate the method using the subsidies for bed nets experiment in Kenya. Armstrong and Shen (2015) make use of the conditional moment inequalities literature to perform statistical inference on the assignment rule with the objective of quantifying how strong the evidence is in favor of treating certain individuals.

policies. They also impose unconfoundedness in the data, so that conditional ATEs are point-identified, and they allow for budget constraints (although they assume a random allocation among the individuals selected in the unconstrained first-best case whenever the constraint is binding). Under a similar set of assumptions, Athey and Wager (2018) derive an improved welfare regret upper bound, and show how a doubly-robust machine learning approach can attain this bound asymptotically.

In contrast with studies assuming point-identified treatment effects, the research agenda initiated by Manski (2000) examines policy choices under ambiguity, where ambiguity arises when the decision maker has partial knowledge of treatment responses and is not willing to place a subjective distribution on the unknowns. Manski shows that there is no unambiguously correct decision rule when different policies are not dominated by others in all possible states. In such a scenario, he proposes making treatment choices under either a maxmin or a minimax-regret objective function. In the absence of budget constraints, Manski (2004) considers a minimax-regret function, investigating rules assigning individuals to treatments that yield the best (partially identified) outcomes, conditional on observed covariates. He derives a closed-form bound on the maximum regret of any such rule.[61] Stoye (2009) then extends this work, deriving exact minimax-regret rules for randomized experiments.

Kasy (2016) is the only study that examines partial identification of treatment effects with treatment choice under budget constraints. He considers identification (and estimation) of a partially-ordered welfare ranking over policies with a set-identified welfare criterion. Specifically, Kasy investigates the conditions under which it is possible to rank two policies given some objective function – that is, when one can identify the sign of the difference in the objective function under the two policies. This general identification problem is distinct from the task of deriving the optimal assignment rule under ambiguity based on a maxmin or minmax-regret social welfare objective.

Overall, there is no paper (either methodological or involving an empirical application) in which all the following hold simultaneously: (a) the unconfoundedness assumptions fail, so that the treatment effect and the welfare objective function are partially identified; (b) violations of SUTVA are allowed; (c) the treatment choice is made under ambiguity; and (d) the set of allowable policies must satisfy a budget (or capacity) constraint. In these combined respects, our empirical application is novel.

[61] Manski (2006) explores these ideas to investigate the choice of a profiling policy where decisions to search for evidence of crime may vary with observable covariates of individuals at risk of search. The planner chooses the (conditional) search probabilities in order to minimise the social cost of crime and search, when she has only partial knowledge of how policing and sanctions affect offence rates (see also Manski, 2005b).

B Appendix: Data

In this section, we discuss data sources and the construction of variables used in our study.

Satellite-based measures of land use. Annual measures of forested area remaining, the cumulative deforested area, and incremental deforested area in each municipality are taken from the Brazilian government's satellite-based forest monitoring program known as PRODES. Other land use classifications in PRODES include 'non-forest' (mostly cerrado, which is similar to savanna), hydrography, clouds, and 'unobserved.' The data are publicly available at both pixel and municipality levels.[62]

Since 1998, Brazil's National Institute for Space Research (INPE) has been using images from LANDSAT-class satellites to produce the official statistics used by the government to track deforestation and inform public policy (INPE, 2017). The classification of land cover is performed in several steps. First, because deforestation typically occurs during the dry season, INPE selects LANDSAT images between July and September with minimal cloud coverage (the spatial resolution is 60×60 metres). Then, a linear spectral mixture model for each pixel in the data is estimated to obtain the pixel's fraction of different components that help predict land cover.[63] INPE then groups adjacent pixels in larger regions based on their spectral similarities. After image segmentation, it implements a cluster unsupervised classification algorithm to generate the land cover classifications. Finally, several photointerpreters verify the results (and reclassify the land cover when necessary, based on specific contexts, on historical data, and on their judgement). Annual deforestation are calculated taking August 1st as the reference date. (A PRODES year spans the twelve months leading up to July 31st of the current calendar year.) Deforestation is considered irreversible, i.e., once an area has been deforested, it remains classified as deforested in subsequent years.[64]

INPE's classification focuses on detecting deforestation. Yet observed remaining forests have missing observations in some years and do not always decrease monotonically over time (as it should, given that deforestation is considered irreversible). For this reason, we opt to measure

[62] In each year, a small amount of land area is unobserved due to cloud cover. In our data, the average share of cloud cover over the municipal area is 2.5 percent (and the median is zero). Deforestation that goes undetected because of cloud cover in one year is attributed to the first subsequent year in which data permit a determination about land use, as is reasonable.

[63] The pixel components considered are soil, vegetation, and shade. The image fractions that are shade and soil help in the process of identifying deforested areas. Image fraction shade is helpful for areas dominated by tropical forests due to the various strata in the forest structure and the irregularity of the canopy, which contrasts with a low amount of shade in deforested areas. Image fraction soil helps identify transition/contact areas between forest and cerrado (Camara et al., 2006).

[64] There is an important distinction between *incremental deforested area* and the *deforestation rate*. Incremental deforested area measures newly detected primary forest loss, while the deforestation rate adds to the increment estimates of cleared forest area that are under unobserved/clouded areas, based on local extrapolation. The incremental deforested area is available as spatial data. This is the deforestation measure we use in the empirical analysis. In contrast, the deforestation rate is available only at the aggregate level (and is presented publicly by INPE as the official measurement of annual deforestation).

'remaining forest' as the remaining available area in the municipality – that is, the total municipal area minus the non-forested areas, water bodies, and previous cumulative deforested areas (we use 2002 as the base year). This guarantees consistency over time (the correlation between our proxy and the PRODES remaining forested area being 0.99). We drop observations with minimal remaining available area (less than 6 km^2). These are small municipalities mostly located at the extreme eastern region of the Amazon Biome, which are not especially relevant for policies focused on preventing deforestation.[65] Finally, in order to calculate the log odds ratio of the shares of deforestation, we assume the minimum amount of deforestation in a municipality in any year is 0.01 km^2.[66]

Priority status. The official list of Priority municipalities (with precise dates for entry and exit) comes from the Ministry of the Environment. Because there are few municipalities entering and exiting the blacklist from 2009 on, not much can be said with a high degree of accuracy about the impact of the policy in these cases. For this reason, we focus on the initial list of Priority municipalities established in 2008 and consider them as our treatment group. Our control group is the set of municipalities that did not enter the list before 2010.

Protected areas. We calculate the total amount of protected area – whether managed by federal, state, or municipal government – using geo-referenced data from the National Register of Conservation Units, maintained by the Ministry of the Environment. Initiatives to create and expand protected areas were concentrated in the first phase of the PPCDAm (spanning 2004–2007), before the first municipalities were assigned to the Priority List in 2008.

Prices. Cattle ranching and crop cultivation have been important drivers of deforestation in the Amazon historically. To help disentangle the effects of changing commodity prices on deforestation from the effects of policy interventions, we construct beef and crop price indices for each municipality based on pre-determined cross-sectional variation in the crop mix across municipalities and time-series variation in commodity prices received by producers in the southern Brazilian state of Paraná; unlike prices received by producers further north in the Amazon, prices there are exogenous to the policy interventions we wish to evaluate. Data on prices received by producers of beef, soy, rice, corn, cassava, and cane sugar in the southern state of Paraná are taken from the State Secretariat for Agriculture and Food Supply. (The five selected crops account for approximately 70 percent of total crop harvested area in the Brazilian Amazon averaged across the 2000s.) Prices

[65] Because of the small values used in the denominator in the calculation of deforestation shares, these small municipalities also exhibit implausibly large oscillations in shares of deforestation over time.

[66] This is analogous to setting minimum shares in logit models to be greater than or equal to a small strictly positive number $\epsilon > 0$, for example, $\epsilon = 10^{-4}$ or $\epsilon = 10^{-5}$.

are deflated to 2011 Brazilian reais. Municipality-level data on the amount and value of each form of agricultural production, which we use to weight the Paraná prices, come from surveys administered by the Brazilian Institute of Geography and Statistics (IBGE) – the Municipal Crop Survey and the Municipal Livestock Survey. Specifically, for beef cattle, the weight is given by the ratio of head of cattle to municipal area (given that annual pasture area is not observable). For crops, we first calculate a weighted price for each crop by multiplying the Paraná prices by the share of the municipal area used to cultivate the crop. For all agricultural products, we fix the weights in the period 2000–2001 (averaged over these two years), so that they capture the relative importance of the different products within municipalities' agricultural production in years preceding the (available) sample period, and preceding the structural break that occurred in 2004–2005 with the first phase of the PPCDAm. Finally, we apply principal component analysis to the individual weighted crop prices to derive a single index, capturing the price variations common to the five selected crops. The crop index is based on the first principal component; given that the first component maximizes price variance, it provides a more comprehensive measure of the agricultural output price scenario for this analysis than the individual prices (see Assunção et al., 2017; Assunção and Rocha, 2019).

Rainfall and Temperature. Drier forests require less effort to clear and convert to pasture or cropland because they can be burnt more easily. A prolonged dry season or otherwise low annual rainfall can thus contribute to higher rates of deforestation. Our measures of annual precipitation and temperature in each municipality are taken from Matsuura and Willmott (2012), whose gridded estimates of total monthly precipitation and average temperature are based on spatial interpolation of climate data from a large number of monitoring stations operating in South America and elsewhere. We take the accumulated precipitation over the year as our rainfall variable; our annual temperature is the average across months. (Annual values are constructed based on the PRODES year.) Municipal data are obtained from the intersection of the municipal area with one or more data points on the Matsuura and Willmott's grid (we take the within-municipality average when appropriate). In cases in which the intersection is empty, we construct a buffer area around the municipality boundary and then take the intersection of the buffer area with the grid points.

Municipalities' Gross Domestic Product. Annual data on municipalities' total and agricultural GDP come from IBGEs regional account system. Agricultural GDP includes crop and livestock production.

Number of Cattle and Crop Area. Annual data on the number of cattle and crop area per municipality come from the IBGE's surveys: the Municipal Livestock Survey and the Municipal Crop

Survey.

Carbon Stock. The amount of carbon stock above ground is calculated by Baccini et al. (2012). We combined their raster data of carbon stock with the PRODES data to calculate the average carbon stock in forested and deforested areas in each municipality.[67]

Fines. Data on the number of fines issued for environmental offenses come from IBAMA. We collapse the information down to a municipality-year panel to match our deforestation data. To maintain consistency, we consider the PRODES year as the relevant unit of time in our sample – i.e., the total number of fines in a municiplity in a given year captures all fines applied in that municipality in the twelve months leading up to August of that year.

Alerts. Forest clearing alert data come from the Real-Time System for the Detection of Deforestation (DETER), developed and operated by the space agency INPE. DETER makes use of satellite images from MODIS, which has a spatial resolution of 250m (25 hectares), and generates alerts biweekly.[68] The data are publicly available in vector format and are aggregated up to the monthly level. Gandour (2018) has rasterized the georeferenced alerts at a 900m spatial resolution, and constructed panel data in which a cell in the raster data takes on a value of 1 if it contains an alert and a value of 0 otherwise. Gandour has also added the number of alerts per municipality per year (consistently with the PRODES year) and has generously shared the aggregated data with us, for which we are very grateful.

The alert system was implemented in 2004, but remained in experimental mode through mid-2005. While few months of data are available for 2004 and early-2005, consistent alert data starts in the second half of 2005 (Gandour, 2018). Consonant with the time period covered in our main data set, we make use of DETER alert data from 2006 to 2010.

Rural Credit. The Brazilian Central Bank collects detailed information covering all rural loan contracts negotiated by farmers and banks (including private and state-owned banks, as well as credit cooperatives). The microdata contain information about the amount borrowed, the interest rate, initial and maturation dates, and the category under which credit was loaned (short-term operating funds, investment, or commercialization). The values of the contract loans were aggregated up to the municipality-year level (Assunção and Rocha, 2019).

[67] There are 18 municipalities with missing carbon stock data, most of them in the Eastern Amazon.

[68] After 2015, INPE has upgraded the system in order to detect changes in land cover in patches larger than 1 hectare (instead of areas larger than 25 hectares), albeit at a lower temporal frequency (Gandour, 2018).

C Appendix: Discussion of Alternative Identification Strategies

In this appendix, we offer a brief discussion of other identification strategies commonly employed in the literature.

Selection-on-Observables. Selection-on-observable techniques require the Conditional Independence Assumption (CIA). In the current application, that would require the independence of U_{mt} and G_m given W_{mt}, where W_{mt} can include X_{mt} and Z_{mt-1} (recalling $Z_{mt-1} = \left(Z^1_{mt-1}, Z^2_{mt-1}, Z^3_{mt-1}\right)$), and lagged variables. From (1), it is clear that conditioning on the criteria variables Z_{mt-1} may suffice to satisfy CIA. Recall that Z_{mt-1} almost completely determines Priority status, so there is little room left for G_m and U_{mt} to be correlated. However, the common support assumption required for selection-on-observable techniques fails for the same reason: it is not satisfied because there is so little overlap in the data between Priority and non-Priority groups given Z_{mt-1}.

Regression Discontinuity Designs. Given the evidence in the data for the selection rule (1), a natural candidate for estimating treatment effects is to exploit a regression discontinuity (RD) design. However, in our case, RD suffers from two important limitations. First, there are few observations close to the threshold frontier. This severely limits the value of the regression discontinuity approach. Second, and more importantly, this approach identifies average treatment effects at the cutoff frontier, but this is not the parameter we are interested in. Instead, we are interested in estimating policy treatment effects, including treatment effects other than the effects at that cutoff frontier.

Instrumental Variables Approaches. Given the triangular system of equations (1) and (5), another possibility would be to exploit instrumental variables. The criteria variables Z_{mt-1} might seem to be natural instruments. However, this approach requires Z_{mt-1} to be independent of the unobservables U_{mt}, which is not the case when U_{mt} is serially correlated. For instance, if U_{mt} incorporates a fixed effect term α_m, then Z_{mt-1} is not independent of α_m because the criteria variables involve past deforestation levels, and so are not valid instruments. We expect municipality fixed effects to be present because time-invariant unmeasured factors that differ systematically across municipalities, such as soil quality, climate conditions and topography, are likely to affect farmers' decisions to deforest. (Furthermore, the persistence in the deforestation process suggests that time-varying unobservables may be serially correlated even in the absence of fixed effects.) Finally, note that although the flexible marginal treatment effects framework developed by Heckman

and Vytlacil (2005) could be implemented to recover the ATU (as is needed to investigate optimal targeting), it also requires access to valid instruments.

D Appendix: Dynamic Treatment Effects

In this appendix, we describe how dynamic treatment effects are incorporated into the analysis. Doing so when calculating counterfactual deforestation is important because the evolution of the remaining forested area depends on deforestation in previous periods. For simplicity, consider three consecutive periods: t, $t+1$, and $t+2$, where t refers to the time period before the treatment, and $t+1$ and $t+2$ refer to the first and second time periods after the treatment. Here, we focus on the second period $t+2$.

To condition on the counterfactual deforestation in period $t+1$, first note that, for any level of deforestation d, there exists a unique v such that $d = \varphi(X, v) \times A$. Then, potential deforestation D^j at $t+2$ conditioning on the potential deforestation D^l at $t+1$ at deforestation level d, for $j, l = 0, 1$, and for group $G = g$, is

$$\begin{aligned}
& E\left[D^j_{mt+2} | D^l_{mt+1} = d, X_{mt+2}, X_{mt+1}, A_{mt+1}, G_m = g\right] \\
= \; & E\left[D^j_{mt+2} | D^l_{mt+1} = \varphi(X_{mt+1}, v) \times A_{mt+1}, X_{mt+2}, X_{mt+1}, A_{mt+1}, G_m = g\right] \\
= \; & \int \left[\varphi\left(X_{mt+2}, v'\right)(1 - \varphi(X_{mt+1}, v)) A_{mt+1}\right] dF_{V^j_{gt+2}}\left(v'\right).
\end{aligned}$$

By the Law of Iterated Expectations,

$$\begin{aligned}
& E\left[D^j_{mt+2} | X_{mt+2}, X_{mt+1}, A_{mt+1}, G_m = g\right] \\
= \; & \int \left[\int \left[\varphi\left(X_{mt+2}, v'\right)(1 - \varphi(X_{mt+1}, v)) A_{mt+1}\right] dF_{V^j_{gt+2}}\left(v'\right)\right] dF_{V^l_{gt+1}}(v). \qquad (12)
\end{aligned}$$

When the support condition does not hold, bounds for potential deforestation at $t+2$ become

$$\begin{aligned}
& \int \left[\int \left[\varphi\left(X_{mt+2}, v'\right)(1 - \varphi(X_{mt+1}, v)) A_{mt+1}\right] dF^L_{V^j_{gt+2}}\left(v'\right)\right] dF^L_{V^l_{gt+1}}(v) \\
\leq \; & E\left[D^j_{mt+2} | X_{mt+2}, X_{mt+1}, A_{mt+1}, G_m = g\right] \\
\leq \; & \int \left[\int \left[\varphi\left(X_{mt+2}, v'\right)(1 - \varphi(X_{mt+1}, v)) A_{mt+1}\right] dF^U_{V^j_{gt+2}}\left(v'\right)\right] dF^U_{V^l_{gt+1}}(v). \qquad (13)
\end{aligned}$$

Based on this reasoning, one can compute (and bound) average treatment effects for any sequence of treatments for both treated and untreated groups in time periods $t+3, t+4$, etc.

E Appendix: Minimax Optimal Policy

In this appendix, we explain in more detail how the ex-post optimal lists are calculated in practice. Denote the counterfactual assignment rule by $\phi = (\phi_1, ...\phi_M)'$. This assigns the treatment to municipalities $m = 1, ..., M$ and can be either deterministic $\phi_m \in \{0, 1\}$ or probabilistic $\phi_m \in [0, 1]$. Denote the expected deforestation of municipality m that is in group $G_m = g$ in case it is placed on the Priority List by

$$D_{gm}^T \equiv E\left[D_m^T | X_m, A_m, G_m = g\right],$$

where the superscript T denotes 'treated.' Similarly, if m is not put on the list, we have

$$D_{gm}^U \equiv E\left[D_m^U | X_m, A_m, G_m = g\right],$$

where the superscript U denotes 'untreated.' (We omit the time dimension t to simplify exposition.) When D_{gm}^j is not point-identified, we adopt the minimax criterion to select the optimal list, in which case we make use of the (estimated) upper bound on D_{gm}^j. From equation (7) in the main text, the upper bound is

$$\sup_{\gamma \in \Gamma} E_\gamma\left[D_m^j | X_m, A_m, G_m = g\right] = \int \left[\varphi\left(X_m, v\right) \times A_m\right] dF_{V_g^j}^U\left(v\right) \equiv \overline{D}_{gm}^j.$$

The expected levels of deforestation (and their upper bounds) are estimated in the data using the CIC model.

As mentioned in the main text, we do not select a list that changes over time as this complicates the problem substantially, given the combinatorics involved. Instead, the optimal list is based on the sum of deforestation in the two years after the treatment (2009 and 2010). That is, the amount of deforestation that enters the social cost function is the sum of the expected deforestation in 2009 (calculated based on equation (6) in the main text), and the expected deforestation in the following year, taking into account the counterfactual remaining forested area from the previous year, as explained in Appendix D – see equation (12). (Upper bounds are placed where appropriate.)

E.1 The Baseline Case

When spillover effects are not considered, the objective function of the policy maker is to select an assignment to minimize the social cost function $SC(\phi)$, given by

$$\begin{aligned} SC(\phi) &= \sum_{m=1}^{M} \phi_m \left[\{G_m = 1\} D_{1m}^T + \{G_m = 0\} \overline{D}_{0m}^T \right] \\ &\quad + (1-\phi_m) \left[\{G_m = 1\} \overline{D}_{1m}^U + \{G_m = 0\} D_{0m}^U \right], \end{aligned}$$

where we use $\{.\}$ to denote the indicator function. Note that the counterfactual deforestation for the treated group in the absence of the intervention is point-identified in the data – that is, $\overline{D}_{1m}^U = D_{1m}^U$.

It is convenient to convert this to matrix notation. Let $\mathbf{D}_g^j$ be an $M_g \times 1$ vector with elements D_{gm}^j, for $j = U, T$, and $g = 0, 1$, where M_g is the number of municipalities in group g. Note that when $j = T$ and $g = 0$, we need to use the upper bound, i.e., $\mathbf{D}_0^T$ is composed of the elements $\overline{D}_{0m}^T$, for $m = 1, ..., M_0$. For other combinations of j and g, we have point identification in the data. Let $\mathbf{D}^j$ stack the vectors $\mathbf{D}_g^j$ for all g, so

$$\mathbf{D}^T = \begin{bmatrix} \mathbf{D}_0^T \\ \mathbf{D}_1^T \end{bmatrix}, \mathbf{D}^U = \begin{bmatrix} \mathbf{D}_0^U \\ \mathbf{D}_1^U. \end{bmatrix}$$

Then,

$$SC(\phi) = \mathbf{D}^{U\prime}\mathbf{1} + \left(\mathbf{D}^T - \mathbf{D}^U\right)' \phi$$

where $\mathbf{1}$ is an $M \times 1$ vector of ones. Minimizing $SC(\phi)$ under the constraints specified in the main text is a simple linear programming problem.

E.2 Incorporating Spillovers into the Optimal List

To take spillover effects into account, we consider three groups in the data: $G_m \in \{0, 1, 2\}$. Group 1 is the treated group; group 0 is the 'pure' control; and group 2 is the 'spillover' group, which is composed of the untreated municipalities that satisfy the following two criteria: (a) they have at least one neighbor treated and (b) their previous deforestation levels were close to the threshold selection criteria.

Now there are three possibilities: If a municipality m is treated, the expected deforestation is D_{gm}^T. If it is not treated, and either has no neighbor treated or is 'far' from the threshold criteria, expected deforestation is D_{gm}^U. If it is untreated with at least one neighbor treated and is 'close' to the threshold criteria, deforestation is D_{gm}^S (where we use superscript S to denote 'spillover').

To incorporate spillovers, we first consider the geographic component of the criteria. The adjacency matrix indicating whether municipality m and n are neighbors is given by

$$W = \begin{bmatrix} 0 & w_{12} & \cdots & w_{1M} \\ w_{21} & 0 & \cdots & w_{2M} \\ \vdots & \vdots & \ddots & \vdots \\ w_{M1} & w_{M2} & \cdots & 0 \end{bmatrix},$$

where w_{mn} equals 0 if m and n are not neighbors, and it equals 1 if they are neighbors (setting $w_{mm} = 0$). Given W and a deterministic assignment rule to treatment $\phi \in \{0,1\}^M$, the number of neighbors of m that are treated is given by $\sum_{n=1}^{M} w_{mn}\phi_n$. Define the function

$$N_m(\phi) = 1\left\{\sum_{n=1}^{M} w_{mn}\phi_n > 0\right\},$$

which equals one if there is at least one neighbor of m treated, and zero if m has no neighbor treated.

The second criterion is whether past deforestation of m is close to the threshold rule or not. Denote this by the indicator variable $R_m \in \{0,1\}$. The two criteria are satisfied only when $R_m N_m(\phi) = 1$. Specifically, when m is untreated, we expect deforestation to be D_{gm}^S when $R_m N_m(\phi) = 1$, and D_{gm}^U when $R_m N_m(\phi) = 0$.

The objective function of the policy maker is now

$$\begin{aligned} SC(\phi) \quad = \quad & \sum_{m=1}^{M} \phi_m \left[\{G_m = 0\}\,\overline{D}_{0m}^T + \{G_m = 1\}\,D_{1m}^T + \{G_m = 2\}\,\overline{D}_{2m}^T\right] \\ & + (1-\phi_m)(1 - R_m N_m(\phi)) \\ & \times \left[\{G_m = 0\}\,D_{0m}^U + \{G_m = 1\}\,\overline{D}_{1m}^U + \{G_m = 2\}\,\overline{D}_{2m}^U\right] \\ & + (1-\phi_m)(R_m N_m(\phi)) \\ & \times \left[\{G_m = 0\}\,\overline{D}_{0m}^S + \{G_m = 1\}\,\overline{D}_{1m}^S + \{G_m = 2\}\,D_{2m}^S\right]. \end{aligned}$$

Again, when the counterfactual is point-identified, we have $\overline{D}_{gm}^j = D_{gm}^j$. As before, let $\mathbf{D}_g^j$ be an $M_g \times 1$ vector with elements D_{gm}^j, for $j = U, T, S$, and $g = 0, 1, 2$ (where upper bounds $\overline{D}_{gm}^j$ are

placed where appropriate). Let $\mathbf{D}^j$ stack the vectors $\mathbf{D}_g^j$ for all g, giving

$$\mathbf{D}^U = \begin{bmatrix} \mathbf{D}_0^U \\ \mathbf{D}_1^U \\ \mathbf{D}_2^U \end{bmatrix}, \mathbf{D}^T = \begin{bmatrix} \mathbf{D}_0^T \\ \mathbf{D}_1^T \\ \mathbf{D}_2^T \end{bmatrix}, \mathbf{D}^S = \begin{bmatrix} \mathbf{D}_0^S \\ \mathbf{D}_1^S \\ \mathbf{D}_2^S \end{bmatrix}.$$

Also define the vector satisfying the criteria for the spillovers effects:

$$\mathbf{NR}(\phi) = \{W\phi > 0\} \circ R,$$

where R is the $M \times 1$ vector of municipalities with elements $R_m \in \{0, 1\}$, and $\circ$ indicates Hadamard (i.e., element-by-element) multiplication. Then

$$\begin{aligned} SC(\phi) &= \left[\mathbf{D}^U + Diag\left(\mathbf{D}^S - \mathbf{D}^U\right)\mathbf{NR}(\phi)\right]'\mathbf{1} \\ &\quad + \left[\mathbf{D}^T - \mathbf{D}^U - Diag\left(\mathbf{D}^S - \mathbf{D}^U\right)\mathbf{NR}(\phi)\right]'\phi. \end{aligned}$$

Given that $SC(\phi)$ is non-linear and non-differentiable in ϕ (because of $N_m(\phi)$), we cannot solve the minimax problem using standard methods (e.g., linear programming or Newton-Raphson). Instead, we use genetic algorithm to find the global minimum (Deep et al., 2009).[69]

The genetic algorithm is a stochastic search algorithm, which is convenient in the current context because it allows for integer optimization in high-dimensional constrained minimization problems. The procedure requires an initial population matrix, in which each row represents a guess for the optimal list, ϕ – that is, each row is composed of elements taking values that are either zeros or ones specifying which of the $M = 490$ municipalities are to be included on the optimal list, subject to the constraint in question (either total number of municipalities or total municipality area). In each step, the objective function is evaluated for each 'individual' (vector) in the population matrix, and the most 'promising' individuals (in terms of minimizing the criterion function) are selected stochastically from the population. The selected vectors are then modified – recombined and possibly randomly mutated – to form a new generation of candidate solutions. The new generation is then used in the next iteration of the procedure. The algorithm stops when the value of the criterion function cannot be further reduced (up to a pre-determined tolerance level), or when a maximum number of generations has been produced.

For each minimization problem considered in the main text, we run the algorithm 20 times.

[69] We have also implemented the following procedure: for an initial ϕ^k, compute $\mathbf{NR}(\phi^k)$. Then update the list ϕ^{k+1} by solving the linear programming problem holding $\mathbf{NR}(\phi^k)$ constant. Then iterate ϕ^k until convergence. Using this procedure, convergence is not guaranteed, however. Indeed, in our experience, the procedure often ends up in cycles and does not converge to a minimum.

Each time we run the algorithm, we provide an initial population matrix with 2,000 candidate solutions. The initial population is composed of (a) the observed list, (b) the optimal list obtained by solving the linear programming problem using the worst-case deforestation for untreated municipalities, regardless of whether an untreated municipality has a neighbor treated or not, (c) the list of municipalities in ascending order of municipality area, (d) the list of municipalities in descending order of municipality area, and (e) 1,995 randomly generated lists that satisfy the constraint (and that are independently generated every time we run the algorithm). The fraction of 'promising' individuals is set to be 20 percent of the population, and the mutation rate is set at 0.01. The maximum number of generations allowed is 49,000 (which equals 100 times the number of municipalities $M = 490$), and the tolerance level for the objective function is $1e - 7$. (We note that the algorithm always stopped before hitting the maximum number of generations.) To check the reliability of the genetic algorithm in our context, we also implemented it in the no spillover case (i.e., when the linear programming solution is appropriate), and always obtained very similar results (up to numerical precision).

We implemented the algorithm in MATLAB using the command "ga," which is part of MATLAB's global optimization toolbox. For details about the creation, crossover, and mutation functions used in the integer programming version of the genetic algorithm, see Deep et al. (2009).

F Appendix: Robustness Analyses

In this appendix, we investigate the robustness of our main results to (a) the way we trimmed observations to reduce the impact of outliers in the estimated treatment effects, and (b) the definition of the spillover group.

Trimming. Recall that, by placing all probability mass outside the support $Supp\,(Y_{1mt+1})$ at the left and right end points of $Supp\,(Y_{0mt+1})$, we obtain the lower and upper bounds for $F_{Y^1_{0t+1}}$ (the same reasoning applies to $F_{Y^0_{1t+1}}$). In practice, when calculating the average treatment effects, we follow the literature and trim observations below the 3rd and above the 97th percentiles to minimize the influence of outliers (Ginther, 2000; Lee, 2009). We now show that the empirical results are robust to such trimming – specifically, to trimming observations below and above the percentiles $[2.5, 97.5]$ and $[3.5, 96.5]$. Table 17 in Appendix G presents the results for the average treatment effects. The top panel shows the estimated ATT, ATU, and ATE when we trim the observations below the 2.5th and above the 97.5th percentiles, while the bottom panel presents the results when we use the 3.5th and 96.5th percentiles. (Table 17 is comparable to Table 5 in the main text.) The ATTs are unaffected by the trimming, and the estimated identified sets for the ATU and ATE differ only slightly across specifications (and all treatment effects are significantly different from zero).

Table 18 in Appendix G presents the implications of the counterfactual optimal lists for deforestation and carbon emissions. As before, the top panel presents results for the $[2.5, 97.5]$–trimming, and the bottom panel, for the $[3.5, 96.5]$–trimming (which are comparable to Table 10 in the main text). Again, the results are robust to these different specifications.

Spillovers. As explained in the main text, one of the criteria used to define whether a municipality belongs to the Spillover group or not concerns whether it has high levels of past deforestation. Formally, we opted for the following condition: $Z^1_{mt-1} \geq 0.7 \times 2,700$ km^2 and $Z^2_{mt-1} \geq 0.7 \times 220$ km^2. We now show that the results are robust to different definitions of how close past deforestation is to these thresholds, considering Z^1_{mt-1} and Z^2_{mt-1} greater than 65 percent and 75 percent of the threshold criteria.

The top panel of Table 19 in Appendix G shows the ATT, ATU, ATS, and ATE when we consider the 65 percent definition for the spillover group, and the bottom panel presents the results based on the 75 percent definition. (Table 19 is comparable to Table 6 in the main text.) The ATTs and the identified sets for the ATU are essentially unaffected. The estimated ATSs increase as we move from the 65 percent to the 75 percent definitions (though not always monotonically). This is consistent with the interpretation that the greater the deforestation level in a municipality,

the closer to the threshold criteria it is, and the more likely it is that farmers there may react to the policy intervention. So, when the spillover group is composed of municipalities with lower levels of past deforestation (the 65 percent group definition), we expect the treatment effects to be smaller than when the group is composed of municipalilties with higher levels of past deforestation. Still, the estimated magnitudes of the ATSs are similar across the different threshold criteria. In addition, almost all 95 percent confidence intervals for the ATSs corresponding to the different group definitions overlap (for each combination of the baseline year, 2006–2007, and post-treatment year, 2009–2010).[70]

Table 20 in Appendix G presents the implications for the optimal lists (comparable to Table 12 in the main text). Once more, the baseline year 2006 provides more conservative estimates (as in most specifications), and the results are robust to alternative definitions of the spillover group.

[70]The only exception corresponds to the confidence intervals of the 65 percent and 75 percent groups for the baseline year 2006 and post-treatment year 2009. However, the distance between these confidence intervals is just 0.29 km^2.

G Appendix: Additional Tables and Figures

Table 13: Aggregate Time Series Data

		Policies		
Year	Total new deforested area	Municipalities on Priority List	Number of fines issued	Expansions to protected area
2002	24,812	0	1,090	–
2003	29,243	0	2,906	6,499
2004	26,283	0	3,903	5,880
2005	22,838	0	4,107	14,985
2006	10,601	0	5,568	19,209
2007	11,142	0	4,696	16,314
2008	12,773	36 (+36/-0)	7,451	6,783
2009	5,568	43 (+7/-0)	5,607	2,729
2010	5,973	42 (+0/-1)	4,737	55
2011	5,547	47 (+6/-1)	5,113	86
2012	4,335	45 (+2/-4)	–	–
2013	5,185	–	–	–

Notes: Balanced Panel of 526 municipalities in the Amazon Biome.
Areas are measured in square kilometers.

Table 14: Support of Residuals V, by Group and Time Period

Group/Year	Support of V_{jt}
Untreated Group, 2009	[-6.856, 2.927]
Treated Group, 2009	[-3.263, 0.660]
Untreated Group, 2010	[-7.658, 2.722]
Treated Group, 2010	[-3.205, 0.769]

Table 15: Difference-in-Differences Results

	(1) Log odds	(2) Log odds	(3) Log odds	(4) Log odds
Treated Group x Year=2009	-0.456**	-0.457**	-0.471**	-0.471**
	(0.150)	(0.147)	(0.152)	(0.150)
Treated Group x Year=2010	-0.928***	-0.887***	-0.964***	-0.922***
	(0.160)	(0.156)	(0.162)	(0.159)
Spillover Group x Year=2009			-0.283	-0.243
			(0.149)	(0.143)
Spillover Group x Year=2010			-0.682***	-0.624***
			(0.188)	(0.186)
Year=2006	-0.370***	-0.569***	-0.370***	-0.575***
	(0.0661)	(0.0994)	(0.0661)	(0.0997)
Year=2007	-0.467***	-0.564***	-0.467***	-0.565***
	(0.0659)	(0.0729)	(0.0660)	(0.0729)
Year=2009	-0.948***	-0.920***	-0.933***	-0.909***
	(0.0769)	(0.0763)	(0.0801)	(0.0794)
Year=2010	-0.672***	-0.730***	-0.636***	-0.698***
	(0.0806)	(0.0830)	(0.0839)	(0.0865)
Lagged Rainfall		0.142**		0.138**
		(0.0483)		(0.0482)
Lagged Rainfall Squared		-0.00329***		-0.00323***
		(0.000971)		(0.000969)
Lagged Temperature		0.175		0.169
		(0.0908)		(0.0903)
Share of Protected Areas		0.603		0.562
		(0.671)		(0.668)
Price of Beef Lagged		-0.0231***		-0.0230***
		(0.00653)		(0.00652)
Price of Crops Lagged		-0.190		-0.248
		(0.331)		(0.336)
Lagged GDP		-0.557		-0.552
		(0.511)		(0.507)
Covariates	NO	YES	NO	YES
R^2	0.098	0.113	0.101	0.115
Observations	2450	2450	2450	2450

Notes: An observation is a municipality in the Brazilian Amazon. The dependent variable is the log odds ratio of deforestation shares. Rainfall is measured in millimetres (mm) and temperature is measured in degrees Celsius ($^{\circ}C$). The price of beef is a weighted average of international beef prices weighted by the ratio of head of cattle to municipal area. The price of crops is the price index based on a principal component analysis applied to individual weighted prices of the most predominant crops in the Brazilian Amazon (the weights are given by the share of the municipal area used to cultivate the crop). For all agricultural products, the weights are fixed in the period 2000–2001. Municipal GDP is measured in million Reais. All monetary amounts are expressed in December 2011 Reais. The coefficient on the constant term is omitted. All regressions include municipality fixed effects. Robust standard errors in parentheses are clustered at the municipality level.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table 16: Pre-Treatment 'Common Trends' Test, 2003–2007

	(1) Log odds	(2) Log odds
Treated Group x Year=2003	-0.286	-0.294
	(0.179)	(0.181)
Treated Group x Year=2004	0.205	0.220
	(0.165)	(0.167)
Treated Group x Year=2005	0.288	0.305
	(0.157)	(0.160)
Treated Group x Year=2006	-0.107	-0.111
	(0.137)	(0.140)
Spillover Group x Year=2003		-0.121
		(0.222)
Spillover Group x Year=2004		0.287
		(0.181)
Spillover Group x Year=2005		0.317
		(0.181)
Spillover Group x Year=2006		-0.0866
		(0.144)
Year=2003	1.127***	1.135***
	(0.100)	(0.105)
Year=2004	0.830***	0.818***
	(0.0991)	(0.102)
Year=2005	0.583***	0.566***
	(0.0966)	(0.101)
Year=2006	0.108	0.116
	(0.0955)	(0.0995)
Lagged Rainfall	0.0581	0.0605
	(0.0414)	(0.0416)
Lagged Rainfall Squared	-0.00166	-0.00170
	(0.000896)	(0.000901)
Lagged Temperature	-0.376*	-0.382*
	(0.158)	(0.158)
Share of Protected Areas	0.838**	0.841**
	(0.300)	(0.299)
Price of Beef Lagged	-0.00107	-0.00111
	(0.00830)	(0.00833)
Price of Crops Lagged	0.886***	0.884***
	(0.183)	(0.180)
Lagged GDP	-0.356**	-0.361**
	(0.132)	(0.131)
R^2	0.147	0.148
Observations	2454	2454

Notes: An observation is a municipality in the Brazilian Amazon. The dependent variable is the log odds ratio of deforestation shares. Rainfall is measured in millimetres (mm) and temperature is measured in degrees Celsius ($^\circ C$). The price of beef is a weighted average of international beef prices weighted by the ratio of head of cattle to municipal area. The price of crops is the price index based on a principal component analysis applied to individual weighted prices of the most predominant crops in the Brazilian Amazon (the weights are given by the share of the municipal area used to cultivate the crop). For all agricultural products, the weights are fixed in the period 2000–2001. Municipal GDP is measured in million Reais. All monetary amounts are expressed in December 2011 Reais. The coefficient on the constant term is omitted. All regressions include municipality fixed effects. Robust standard errors in parentheses are clustered at the municipality level.
$^{*}\ p < 0.05, ^{**}\ p < 0.01, ^{***}\ p < 0.001.$

Table 17: Robustness: Deforestation Average Treatment Effects, Trimming

Average Treatment Effects, Trimming: 2.5th and 97.5th Percentiles			
	ATT	ATU	ATE
2009			
Baseline 2006	-21.61	[-4.75, -2.17]	[-5.95, -3.56]
	(-24.35, -18.88)	(-4.85, -2.08)	(-6.08, -3.45)
Baseline 2007	-24.91	[-4.71, -2.77]	[-6.16, -4.35]
	(-28.28, -21.55)	(-4.82, -2.67)	(-6.29, -4.22)
2010			
Baseline 2006	-50.94	[-7.41, -4.69]	[-10.52, -7.99]
	(-55.30, -46.58)	(-7.53, -4.59)	(-10.68, -7.85)
Baseline 2007	-53.93	[-7.42, -5.37]	[-10.75, -8.84]
	(-58.58, -49.29)	(-7.54, -5.26)	(-10.91, -8.69)

Average Treatment Effects, Trimming: 3.5th and 96.5th Percentiles			
	ATT	ATU	ATE
2009			
Baseline 2006	-21.61	[-4.73, -3.57]	[-5.94, -4.86]
	(-24.35, -18.88)	(-4.83, -3.49)	(-6.06, -4.75)
Baseline 2007	-24.91	[-4.70, -3.82]	[-6.14, -5.33]
	(-28.28, -21.55)	(-4.80, -3.73)	(-6.28, -5.20)
2010			
Baseline 2006	-50.94	[-7.38, -5.59]	[-10.50, -8.83]
	(-55.30, -46.58)	(-7.50, -5.49)	(-10.65, -8.69)
Baseline 2007	-53.93	[-7.40, -6.05]	[-10.72, -9.47]
	(-58.58, -49.29)	(-7.52, -5.95)	(-10.89, -9.32)

Notes: 95% confidence intervals are in parentheses. For ATT, the intervals are computed based on the standard i.i.d. nonparametric bootstrap, where the i.i.d. resampling occurs in the cross-sectional dimension. For ATU and ATE they are based on Imbens and Manski (2004). We implemented 500 bootstrap replications. Deforestation is measured in square kilometres.

Table 18: Robustness: Ex-Post Optimal, Trimming

Trimming: 2.5th and 97.5th Percentiles						
Constraint:	*Total Area*		*Number of Municipalities*			
	Observed vs Optimal		Observed vs Optimal		Random vs Optimal	
	Ratio	Value	Ratio	Value	Ratio	Value
Total Deforestation						
Baseline 2006	1.04	-	1.04	-	1.21	-
Baseline 2007	1.05	-	1.05	-	1.23	-
Total Carbon Emissions						
Baseline 2006	1.03	381	1.05	609	1.24	2,496
Baseline 2007	1.04	479	1.06	757	1.27	2,785

Trimming: 3.5th and 96.5th Percentiles						
Constraint:	*Total Area*		*Number of Municipalities*			
	Observed vs Optimal		Observed vs Optimal		Random vs Optimal	
	Ratio	Value	Ratio	Value	Ratio	Value
Total Deforestation						
Baseline 2006	1.06	-	1.06	-	1.23	-
Baseline 2007	1.07	-	1.07	-	1.25	-
Total Carbon Emissions						
Baseline 2006	1.05	628	1.08	942	1.27	2,770
Baseline 2007	1.06	670	1.09	1,012	1.29	2,996

Notes: 'Ratio' divides total deforestation (total emissions) evaluated at the observed list by the ex-post optimal total deforestation (total emissions). 'Value' takes their difference. Values are measured in million US$, assuming a social cost of carbon of US$ 20/tCO2.

Table 19: Robustness: Deforestation Average Treatment Effects, Spillovers

Average Treatment Effects, Spillovers: Above 65 Percent of the Threshold				
	ATT	ATU	ATS	ATE
2009				
Baseline 2006	-24.53	[-4.42, -2.82]	[-8.79, -8.77]	[-6.13, -4.75]
	(-27.47, -21.58)	(-4.51, -2.75)	(-10.48, -7.10)	(-6.26, -4.64)
Baseline 2007	-28.20	[-4.34, -3.16]	[-12.48, -12.44]	[-6.56, -5.54]
	(-31.94, -24.46)	(-4.44, -3.08)	(-14.69, -10.22)	(-6.70, -5.40)
2010				
Baseline 2006	-52.75	[-6.70, -4.97]	[-17.73, -10.19]	[-10.69, -8.71]
	(-57.28, -48.21)	(-6.81, -4.88)	(-19.56, -8.07)	(-10.85, -8.57)
Baseline 2007	-56.46	[-6.69, -5.41]	[-20.34, -15.21]	[-11.11, -9.67]
	(-61.29, -51.64)	(-6.80, -5.31)	(-22.55, -12.91)	(-11.28, -9.52)

Average Treatment Effects, Spillovers: Above 75 Percent of the Threshold				
	ATT	ATU	ATS	ATE
2009				
Baseline 2006	-23.09	[-4.37, -2.68]	[-13.64, -13.61]	[-6.05, -4.54]
	(-25.91, -20.27)	(-4.47, -2.60)	(-16.48, -10.77)	(-6.17, -4.43)
Baseline 2007	-27.52	[-4.40, -3.18]	[-12.98, -12.95]	[-6.37, -5.27]
	(-31.10, -23.93)	(-4.50, -3.09)	(-16.23, -9.69)	(-6.50, -5.14)
2010				
Baseline 2006	-51.56	[-6.74, -4.89]	[-23.34, -17.89]	[-10.55, -8.70]
	(-55.99, -47.13)	(-6.85, -4.80)	(-26.71, -14.07)	(-10.70, -8.56)
Baseline 2007	-57.59	[-6.83, -5.50]	[-23.08, -17.64]	[-11.06, -9.66]
	(-62.38, -52.81)	(-6.94, -5.40)	(-26.84, -13.19)	(-11.22, -9.51)

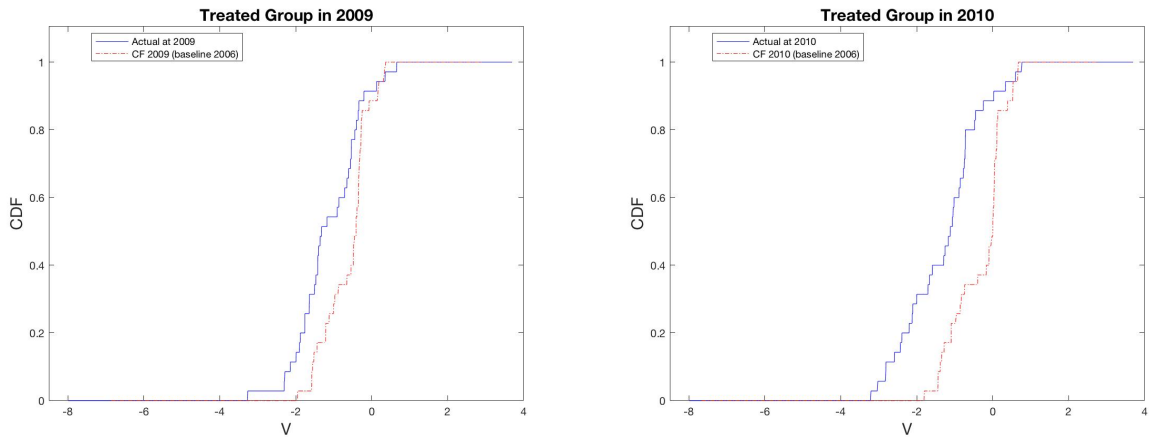
Notes: 95% confidence intervals are in parentheses. For ATT, the intervals are computed based on the standard i.i.d. nonparametric bootstrap, where the i.i.d. resampling occurs in the cross-sectional dimension. For ATU, ATS, and ATE, they are based on Imbens and Manski (2004). We implemented 500 bootstrap replications. Deforestation is measured in square kilometres.

Table 20: Robustness: Ex-Post Optimal, Spillovers

Spillovers: Above 65 Percent of the Threshold						
Constraint:	*Total Area*		*Number of Municipalities*			
	Observed vs Optimal		Observed vs Optimal		Random vs Optimal	
	Ratio	Value	Ratio	Value	Ratio	Value
Total Deforestation						
Baseline 2006	1.08	-	1.07	-	1.29	-
Baseline 2007	1.09	-	1.07	-	1.32	-
Total Carbon Emissions						
Baseline 2006	1.07	816	1.08	902	1.33	2,674
Baseline 2007	1.08	892	1.09	982	1.37	2,993

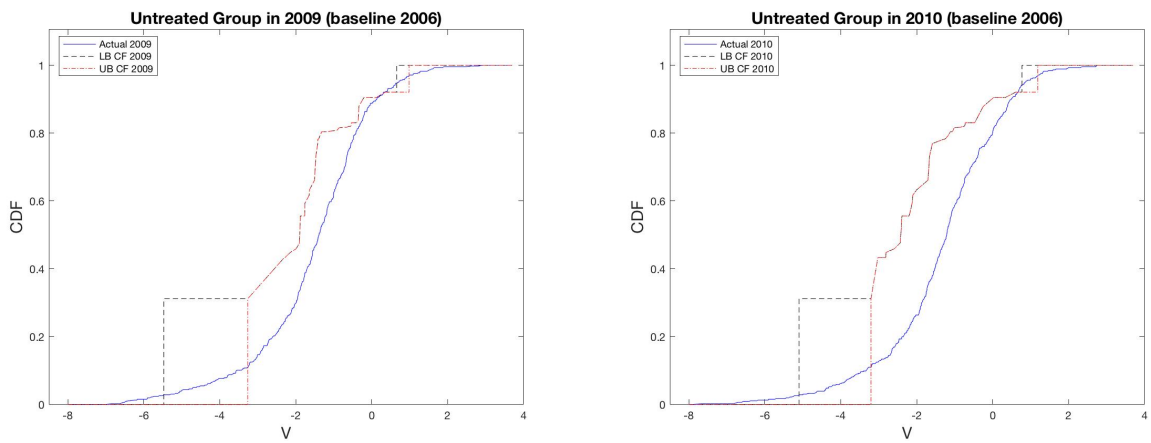
Spillovers: Above 75 Percent of the Threshold						
Constraint:	*Total Area*		*Number of Municipalities*			
	Observed vs Optimal		Observed vs Optimal		Random vs Optimal	
	Ratio	Value	Ratio	Value	Ratio	Value
Total Deforestation						
Baseline 2006	1.04	-	1.04	-	1.24	-
Baseline 2007	1.08	-	1.07	-	1.31	-
Total Carbon Emissions						
Baseline 2006	1.04	419	1.05	606	1.27	2,926
Baseline 2007	1.08	884	1.09	964	1.35	2,865

Notes: 'Ratio' divides total deforestation (total emissions) evaluated at the observed list by the ex-post optimal total deforestation (total emissions). 'Value' takes their difference. Values are measured in million US$, assuming a social cost of carbon of US$ 20/tCO2.



(a) 2009 **(b)** 2010

Figure 13: Factual and Counterfactual Distributions of Residuals V, Treated Group



(a) 2009 **(b)** 2010

Figure 14: Factual and Counterfactual Distributions of Residuals V, Untreated Group