Optimal Environmental Targeting in the Amazon Rainforest*

Juliano Assunção, Robert McMillan, Joshua Murphy, and Eduardo Souza-Rodrigues

July 13, 2018

*We would like to thank Sylvain Chabe-Ferret, Francisco Costa, Clarissa Gandour, Ismael Mourifié, Michel Oestreich, Alex Pfaff, Alberto Salvo, Paul Scott, and Aloysius Siow for helpful discussions, and workshop participants at the Instituto Escolhas, Brock University, the Toulouse School of Economics, and the University of Toronto for additional comments. Faisal Ibrahim provided outstanding research assistance. Financial support from the University of Toronto Mississauga is gratefully acknowledged. All remaining errors are our own. Affiliations: Juliano Assunção, PUC-Rio; Robert McMillan, University of Toronto; Joshua Murphy, Natural Resources Canada; Eduardo Souza-Rodrigues, University of Toronto.

Abstract

This paper studies the causal effects of targeted environmental regulations on deforestation and how such regulations can be optimized. We focus on the Brazilian Amazon, where the federal government issued a 'Priority List' of municipalities in 2008 that became subject to more stringent monitoring and heavier fines than elsewhere. First, we estimate the causal impact of the actual Priority List on deforestation levels using 'changes-in-changes' (Athey and Imbens, 2006), a flexible treatment effects estimation method, finding that the Priority List caused a substantial decrease in deforestation, by 40 percent, relative to the case in which no program was enacted; this led to avoided emissions of 30 million tons of carbon, with a social benefit of around $2.2 billion. Second, we develop a framework for comparing the Priority List with a targeted ex-post optimal blacklist. The framework captures how a policy maker would assign municipalities to a counterfactual list that minimized either total deforestation or total emissions, based on information about ex-post treatment effects drawn from the first part of the analysis. Accounting for spillovers, we show that the actual list resulted in carbon emissions that were at least 8 percent higher than the ex-post optimal list, while a randomly selected list of municipalities would result in emissions that were over 34 percent higher on average. The approach we develop 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, Partial Identification, Minimax Ambiguity

1 Introduction

In many developing countries, there is a growing appreciation that weak institutional arrangements undercut the effective enforcement of environmental policies.[1] Often, the unregulated and sometimes illegal activities that prevail cause severe duress to fragile ecosystems, producing outcomes that are both damaging and inefficient. As a prominent instance of this phenomenon, recent research (notably by Burgess et al., 2012) has 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 such challenging settings, targeted monitoring and enforcement policies may be advantageous, helping to focus limited resources where they can have higher-than-average impacts.

This paper examines the causal effects of a widespread form of targeting – blacklist-type government regulations – in the context of deforestation, in turn assessing how such targeted environmental regulations can be optimized. We focus on the Amazon, the world's most extensive rainforest and a vitally important ecosystem, given its fundamental roles in storing carbon, conserving biodiversity, maintaining water quality and even modulating the Earth's climate (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 25 years, some of that concern translating into increased government regulatory effort, notably on the part of Brazil's federal government. Over the past decade, this regulatory activity has coincided with a marked slowdown in deforestation.[2] As other factors may be responsible for the downward trend, among them the evolution of commodity prices, 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 existing research has not supplied the means to assess – in a systematic quantitative way – which policy configurations are likely to have most impact in limiting future deforestation: 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 deforestation rates – 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 first goal of our analysis is to estimate the causal treatment effect of the Priority List on deforestation levels in the Brazilian Amazon. To that end, we start by investigating the effective selection rule that assigned municipalities to the Priority List, given that the official criteria did not

[1] See Greenstone and Jack (2015) for a thorough review of the issues involved.
[2] Annual deforestation declined by approximately 75 percent from 2004 to 2017.

specify exactly how the list was chosen. The patterns we find in the data indicate that the federal government adhered closely to a threshold rule that essentially separated municipalities based on their deforestation levels, but not on their *trends*.[3] Indeed, we cannot reject the common trends assumption comparing municipalities on the list (versus not) leading up to the introduction of the Priority List in 2008.

In considering the short-run impact of the reform over the period 2006–2010, it might seem reasonable – based on that evidence – to estimate a standard differences-in-differences (DID) model. Yet in the current context, it is plausible to think that heterogeneous treatment effects are present in the data, with the Priority List being implemented on the group with (potentially) higher average benefits when compared to the control group.[4] Given 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 that, we need to obtain information regarding the policy impacts on the untreated.[5]

To account for this issue, we adopt the changes-in-changes (CIC) model proposed by Athey and Imbens (2006), which can be viewed as a nonlinear generalization of the DID model to the entire distribution of the potential outcomes. In a policy evaluation context with a pre- and post-policy period, Athey and Imbens 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 not to the entire distribution function, but to the average, 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 *and* across 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 (relative to the case in which no program was enacted) over the period 2009–2010. This led to avoided emissions of 30 million tons of carbon, with a social benefit of around $2.21 billion, assuming a social cost of carbon of

[3] 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 have important limitations in our setting.

[4] 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 deforest in the future, so concentrating regulatory efforts in highly deforested areas may result in more substantial reductions in total deforestation.

[5] 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 the ex-post policy calculations unreliable.

\$20/tCO_2 (Greenstone et al., 2013; Nordhaus, 2014).[6] Further, there is evidence of heterogeneous treatment effects. Although data limitations prevent us from point-identifying the treatment on the untreated, the estimated effect on the untreated is partially identified with informative identified sets: the average effect on the untreated (ATU) is between 10 and 14 percent of the estimated effect on the treated.

We also investigate the possibility that the Priority List generated spillovers. Farmers in untreated municipalities both geographically close to a Priority municipality and that had seen substantial deforestation in the past might think that monitoring could also increase there.[7] Accordingly, we split the untreated group into two (denoted the 'Spillover' and 'Control' groups), depending on whether untreated municipalities are more or less likely to react to the policy intervention. Based on the estimated CIC model, we find evidence of spillover effects: the spillover group reduced deforestation in response to the intervention, and the treatment effects for this group are smaller than the effect on the treated, but greater than the effect on the untreated.[8]

Our second goal (already referenced) 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 to explore the assignment of municipalities to a counterfactual list based on information about ex-post treatment effects drawn from the first part of the analysis; this allows us to investigate in a systematic way how (partial) knowledge of treatment effects 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. Given that some treatment effects are not point-identified, it is appropriate to analyze the policy maker's decision as a treatment choice problem under ambiguity – here, we use the minimax criterion, assuming the policy maker chooses the ex-post list in order to achieve the best of the worst outcomes (Manski, 2005). Further, to incorporate limited monitoring resources into the minimization problem, we consider two alternative constraints, one limiting the total area that can be monitored, and the other, the total number of municipalities that can be placed on the list.[9]

[6] Standard DID estimates suggest that the Priority List reduced deforestation by a half. There is no *a priori* reason to expect the DID and the CIC estimators to generate similar point estimates. In the present context, given that there are many municipalities with low levels of deforestation (not surprisingly, since deforestation is costly), the standard linear DID predicts negative deforestation for a non-negligible fraction of the observations, which may result in biased ATT estimates and affect the counterfactual optimal list.

[7] This might occur because the government may exploit economies of scale in monitoring neighbors of Priority municipalities, or because highly deforested municipalities might be at greater risk of being placed on the list in the near future. In such circumstance, neighboring farmers might react, reducing deforestation in the present in anticipation of possibly stricter monitoring in the future.

[8] We do not find evidence of increased deforestation in adjacent locations that could have resulted from farmers moving and clearing land in nearby areas with lower (perceived) monitoring.

[9] We set the constraints at the same values as those corresponding to the Priority List. Information about the resources that were effectively allocated to monitoring is difficult, if not impossible, to obtain – presumably, the larger the area or the number of municipalities monitored, the higher the monitoring costs.

Accounting for spillover effects, we show that the Priority List resulted in carbon emissions 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. Given that the lower deforestation and emissions are possible in the counterfactual because of (partial) knowledge of the treatment effects, we estimate high social returns to investments that generate information about the impacts of conservation policies.

The geographic distributions of the ex-post optimal lists reveal several interesting patterns that were not imposed in the course of the estimation. First, the overlap between protected areas and the area-constrained counterfactual list is much smaller than the overlap between the pre-existing protected areas and the original Priority List. This suggests the presence of important complementarities between these two policies that could be further utilized by the Brazilian government. Second, when 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 because placing all targeted municipalities together does not exploit the potential reduction in deforestation in adjacent locations arising from spillovers.

Beyond the current application, the approach we develop is relevant for assessing counterfactual targeted policies to reduce deforestation in other contexts. It also provides a coherent means of assessing the quantitative impacts of policy targeting more generally, using credible estimates based on a flexible treatment effects estimation approach; we develop the more general applicability of our counterfactual approach below.

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

[10] The Appendix complements 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 to four main literatures. First, a growing body of work studies the implementation of environmental policies and regulations in developing countries (Greenstone and Jack, 2015) – a complement to the vast literature studying environmental policies in a developed country setting.[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 a varying degree of success. In the case of climate issues, which are linked with the deforestation process analyzed in this paper, the available evidence is limited (Burke et al., 2016). Our analysis explores the effectiveness of a widespread form of targeting, and links the causal impacts of this to the release of carbon into the atmosphere.

Several papers examine the impact of monitoring and the role of institutions in the Amazon, notably Assunção et al. (2013), Hargrave and Kis-Katos (2013), and Burgess et al. (2017). As alternatives to monitoring, 'payments for ecological services' programs have been studied by Pattanyak et al. (2010), Alix-Garcia et al. (2012), Jayachandran et al. (2017), and Simonet et al. (2018). Compared to that literature, our paper explores the effectiveness of a policy targeting strategy as a way of overcoming institutional and political obstacles. This type of targeted strategy is applicable to other contexts encompassing a substantial portion of global rainforest cover, notably in other Amazonian countries, Africa and Southeast Asia. A number of other papers study the effects of the Priority List itself, including Assunção and Rocha (2014), Arima et al. (2014), Cisneros et al. (2015), Andrade and Chagas (2016), Harding et al. (2018), and Koch et al. (2018). Those papers use difference-in-differences and matching methods to obtain average treatment effects similar in magnitude to the corresponding estimates in our study. In addition, our estimation approach allows us to recover the effects of treatment on the untreated, and to compute an optimally targeted list.

Another branch of literature documents underlying causes of land use change and tropical deforestation, relating to population, infrastructure, agricultural prices, political economy issues or climate-related phenomena.[12] Our results indicate that monitoring policies are important drivers of land use change and deforestation, affecting not only the municipalities that are directly targeted but also generating spillovers for neighboring areas.

From an estimation standpoint, the flexible CIC model has been not been widely used to date, being implemented as a robustness check, or in supplementary analyses (Athey and Imbens, 2017). Indeed, the only empirical application we are aware of, by Havnes and Mogstad (2015), studies the

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

[12]See papers by Chomitz and Gray (1996), Stavins (1999), Pfaff (1999), Nelson and Geoghegan (2002), Andersen et al. (2002), Chomitz and Thomas (2003), Lubowski et al. (2006), Brady and Irwin (2011), Mason and Platinga (2013), and Souza-Rodrigues (2018).

effects of child care in Norway and uses the CIC model when carrying out robustness analyses. Our counterfactual approach using ex-post treatment effects is the first such usage (to the best of our knowledge) in the environmental or regulation literatures. This promising method is applicable in a variety of other settings, especially when providing estimates that can be used counterfactually, as we show.

3 Institutional Background and Regulation

In this section, we provide relevant institutional context, beginning with a brief history of the development of the Brazilian Amazon, then describing the legal environment, the introduction of satellite monitoring, and our main focus – the Priority List.

The Brazilian Amazon – Background. The Brazilian Amazon 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.[13] Forest access was open, and local economic activity was based primarily on subsistence and a few extraction activities – mainly rubber and Brazil nuts.

The occupation of the Amazon was promoted by the military dictatorship during the 1960s and 1970s, with the explicit goal of securing national borders, developing the region, and integrating it into the national economy. Hydroelectric facilities, mining, ports, and around 60,000 km of roads were constructed during that period, though an economic recession and hyperinflation in the 1980s led the government to cut investment. Ecological concerns started to shape policies in the Amazon in the late 1980s. Notably, IBAMA (the Brazilian Environmental Protection Agency) was created in 1989, given power to execute environmental policies, and operating as the national police authority concerned with the investigation and sanctioning of environmental infractions.

The Legal Environment. In terms of the legality of deforestation in the Amazon, 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 relevant ecological interest. Quite strict requirements apply to deforestation in those areas. The rest of the Amazon comprises undesignated public land, where no deforestation is allowed, or private land – approximately 20 percent of the Amazon area (according to the Agricultural Census of 2006) – where deforestation has to follow the rules of the Forest Code. This states, among other requirements, that farms in the Amazon should preserve 80 percent of their area in the form of native vegetation.

[13] This description draws on Souza-Rodrigues (2018).

While deforestation on private land can be legal if it is both authorized/licensed and accords with the Forest Code, empirical evidence suggests that compliance with the Forest Code is limited (Michalski et al., 2010; Godar et al., 2014; Borner et al., 2014). Although some deforested areas captured in our data may have been cleared legally, most deforestation in the Amazon is illegal.

In terms of law enforcement and environmental monitoring, up to the mid–2000s, IBAMA's operations in the Amazon were based largely on intelligence information collected and processed by both headquarters and regional offices.[14] Although land and air patrols were used in Amazonian policing operations 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 enforcement officials.

Satellite-Based Monitoring. The adoption of satellite-based monitoring and other strategies implemented from the mid–2000s onwards had a significant impact on patrolling capabilities. 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), which set out new procedures for monitoring and environmental control, as well as the management of territories.[15]

Central to PPCDAm law enforcement has been the use of high-frequency remote sensing technology in the form of a satellite-based system, the Real-Time System for Detection of Deforestation (DETER), developed by the Brazilian Institute for Space Research (INPE). This has increased the capacity to monitor forest-clearing activities in the Amazon in a significant way. It processes land use images on a frequent basis, and is able to detect areas experiencing a loss of forest cover, 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. (2013) show, this satellite-based system has had an important impact on deforestation.

The Priority List. In 2008, the government launched the second phase of the PPCDAm, the centrepiece involving the creation of a blacklist to better target efforts to combat illegal deforestation.[16] 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

[14] In addition, IBAMA established a channel for anonymous voluntary reporting of illegal activity.

[15] The PPCDAm led both to the expansion of protected areas and to their being placed strategically, with new areas serving to shield advancing deforestation.

[16] The legal basis for targeting certain municipalities was set out in a Presidential Decree in December 2007.

exact rules followed are not in the public domain and so have to be inferred.[17] The Ministry of the Environment's Ordinance 28, issued in January 2008, listed 36 municipalities making up the initial Priority List (corresponding to approximately 7 percent of the number of municipalities in the Brazilian Amazon).

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. First among these, fines were increased in Priority municipalities. Private land titles were also revised in a bid to identify fraudulent documentation and illegal occupancy, and licensing requirements for rural establishments were made stricter. Municipalities on the Priority List also became subject to a series of other administrative measures that imposed additional costs to being blacklisted. They included more stringent conditions applying to the approval of subsidized credit contracts, the requirement to develop local plans for sustainable production, and more (see Brito et al., 2010; Maia et al., 2011; Arima et al., 2014).

The original list of 36 municipalities was expanded to include an additional seven municipalities in 2009. A further six municipalities 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 have been no changes in the list from 2013 to 2017 (when eight new municipalities entered the list). In total, 59 municipalities were eventually placed on the Priority List over the period 2008 to 2017, while 467 municipalities did not enter the list during the same period.

4 Data and Summary Statistics

We have assembled a municipality-year panel dataset that combines information relating to Priority status, land use, and other possible determinants of deforestation, such as prices for beef and soy, and the location of public protected areas. In this section, we describe the sample, examine aggregate trends, and highlight important differences among Priority and non-Priority municipalities.[18]

4.1 The Sample

Our analysis focuses on the time period 2006–2010. The pre-treatment period covers 2006 and 2007, and the post-treatment period is 2009–2010.[19]. The treatment group contains municipalities that

[17] Exiting the Priority List depended on reducing deforestation in a significant way (Arima et al., 2014).

[18] Appendix B provides further information about data sources and the construction of key variables.

[19] Because of the 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 would capture the combined effect of both regulatory changes.

entered and remained on the list from 2008 to 2010 inclusive (with the exception of one observation that exited the list in 2010). The control group is the set of municipalities that did not enter the list before 2010. Given few instances of missing data, we end up with a balanced panel of 490 (out of a possible 526) municipalities within the Amazon Biome. Summary statistics for the main variables of interest are presented in Table 1.

4.2 Descriptive Statistics: Aggregate Trends

To date, around 20 percent of the Brazilian Amazon has already been deforested – an area totalling over 700,000 square kilometres (larger than the state of Texas). Cleared areas are used mostly for agriculture: approximately two thirds of the deforested area comprises pasture, and approximately 8 percent is used for crops, according to Pinheiro et al. (2016).[20]

Figure 1 presents aggregate deforestation trends. This reveals two inflection points that coincide with the main policy changes associated with the first two phases of the PPCDAm, in 2004 and 2008. It is clear that deforestation fell considerably after 2004, and again in 2009, with the rate stabilizing since then. In total, annual deforestation declined by approximately 75 percent from 2004 to 2017.[21]

In terms of possible causes, Figure 2 presents 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. This is 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, which suggests that the Priority List may have helped preserve the rainforest even when international prices were rising.

The geographic location of the municipalities on the Priority List is shown in Figure 3, with Priority municipalities being found mostly in the southern and eastern regions of the Amazon – an area known as the "Arc of Deforestation." Figure 4 shows where the incremental deforestation occurred between 2006 and 2010, and also presents the cumulated deforestation by 2010 (together with Priority municipalities). These figures make clear that there is substantial persistence in terms of where new deforestation occurs; in such circumstances, a targeted policy may be effective, concentrating monitoring and enforcement in locations with persistent issues.

[20] Pinheiro 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 recently burned areas).

[21] Table 11 in Appendix G shows the total incremental deforestation by year, together with the number of fines issued, the expansion of protected areas, and the number of municipalities added to the Priority List.

4.3 Selection onto the Priority List

The Priority status of a municipality m in a given year t 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 newly deforested area at least three times in the five-year period ending with year $t-1$ (Z^3_{mt-1}). The first selection criterion relates to long-run deforestation, the second criterion considers recent new deforestation, and the third, whether deforestation accelerated in recent years.

Under the assumption that these variables alone determine Priority status, the selection equation is:

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

where G_{mt} 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, $Z_{mt-1} = \left(Z^1_{mt-1}, Z^2_{mt-1}, Z^3_{mt-1}\right)$, alone determines Priority status. 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 $Z^3_{mt-1} = 0$; the scatterplot in panel (b) holds $Z^3_{mt-1} = 1$. In both panels, municipality-year observations with $G_{mt} = 0$ and $G_{mt} = 1$ are marked with crosses and dots, respectively.

Visual inspection of panels (a) and (b) indicates 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.[22] Using only these inferred thresholds, we are able to replicate the actual 2008 assignments with 97 percent accuracy.

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

[22] The thresholds drawn in panels (a) and (b) of Figure 5 are 2,700 km^2 for Z^1_{mt-1} and 220 km^2 for Z^2_{mt-1}.

(and RD does not identify the policy treatment effect of interest in this paper). And while the criteria variables in Z_{mt-1} are natural instruments for Priority status, they are invalid instruments when the unobservables affecting deforestation decisions are serially correlated. We discuss these issues in more detail in Appendix E.

Fines and Penalties. The purpose of the Priority List was to direct monitoring and enforcement efforts to municipalities with high levels of deforestation – those, 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 amount 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 treatment and control groups.[23] 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 becomes apparent 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)). Thus, the evidence suggests that the Priority List led to more intense enforcement in Priority municipalities relative to non-Priority municipalities.[24] This is consistent with the results of Assunção and Rocha (2014), who find evidence suggesting that law enforcement is the main channel through which the policy affects deforestation.

4.4 Descriptive Statistics: Treatment versus Control

Summary statistics for a cross section of 2007 data broken out by Priority status are presented in Table 1. As expected given the selection criteria, Priority and non-Priority municipalities differ in important ways. The first row of statistics shows that, on average, the amount of incremental deforestation in Priority municipalities is 136 km^2 higher than in non-Priority municipalities. Also, total historical deforestation – the first selection criterion considered by the Ministry of the Environment when assigning Priority status – is related to treatment in a strong positive way. The average Priority municipality cleared an additional 3,385 square kilometres more forest up until 2007 than the average non-Priority municipality.

Treatment and control groups differ in other respects as well. Priority municipalities are larger and have higher levels of agricultural GDP, including larger numbers of cattle and greater crop area.

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

[24] 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.

Not surprisingly, Priority municipalities are subject to more stringent policy measures: they received more fines and accommodated larger areas for the creation of newly protected land. Although they are not very different in terms of their meteorological conditions (rainfall and temperature), they do differ in terms of the amount of carbon stock per hectare in forested and deforested areas: 111 tons of carbon per hectare (tC/ha) versus approximately 70 tC/ha, comparing Priority and non-Priority municipalities.

Figure 7 compares the evolution of deforestation among treated and untreated municipalities. The differences in deforestation levels are clear, but for both groups, new deforestation fell after 2005, and increased slightly in 2006–2008. This 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). Indeed, as we show in Section 6, we do not reject the common trends assumption before treatment. Further, the slowdown among Priority municipalities after 2008 was substantial, which provides initial evidence that deforestation may have responded to the blacklist policy.

4.5 Spillover Effects

In this subsection, 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.[25] The extent to which deforestation might be relocated geographically (a problem known as 'leakage') depends on how costly it is to move and deforest in other places. Such costs make it unlikely that such leakage would be important in the short run (although leakage may be important in the longer term). Consistent with this view, Figure 4 shows no clear evidence that new deforestation was accumulating after 2008 in non-Priority municipalities close by the municipalities placed on the list.[26]

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.[27] Indeed, Figures 4 and 7 suggest that deforestation decreased in both treated

[25]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.

[26]This suggests that such spillovers may not be 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.

[27]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).

and control municipalities following the treatment.

To investigate whether such spillover effects may be present, we split the control group in two, depending on whether untreated municipalities are more or less likely to react to the policy intervention. We consider two plausible conditions for designating 'spillover' municipalities: (a) whether a municipality shares a border with a treated municipality (i.e., adjacent locations), and (b) whether a municipality has high levels of deforestation historically. Farmers in untreated municipalities that are both geographically close to a Priority municipality and that had cleared large areas of rainforest in the past may believe that monitoring could increase there. In such circumstance, farmers may be more likely to react, reducing deforestation in the present in anticipation of possibly stricter monitoring in future.

Formally, we define our second condition to split the control group (those with 'high levels of historical deforestation') based on the threshold criteria that were (implicitly) adopted by the Brazilian government, shown in Figure 5 in Section 4.3.[28] 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, with average incremental and cumulated deforested areas, municipality sizes, numbers of cattle, and crop area that take on intermediate values.

Figure 8 compares the evolution of deforestation among the three groups. Again, the spillover group features deforestation levels between the other two groups, and the evolution profiles are similar, especially after 2005.

5 Model

In this section, we set out a framework that provides the basis for our estimation approach. Here, as a benchmark, we start with the standard differences-in-differences (DID) framework, then describe the more general changes-in-changes (CIC) model proposed by Athey and Imbens (2006). In the process, we discuss our empirical strategy and the parameters of interest, namely the average treatment effects.

Our empirical approach to studying targeted environmental regulation in the Amazon is shaped by particular data constraints. As described in the Data section, we do not observe the land use

[28]Based on that figure, the threshold criteria were: $Z^1_{mt-1} \geq 2,700$ km^2, and $Z^2_{mt-1} \geq 220$ km^2. Accordingly, 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 (specifically, results are robust to whether Z^1_{mt-1} and Z^2_{mt-1} are greater than 65 percent or 75 percent of the threshold criteria). See Appendix F.

decisions of individual farmers, but rather have land use panel data at the municipal level. Thus we focus on deforestation at the municipal level, treating that 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, and 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 at the micro-level directly is not feasible.

5.1 Empirical Framework

We make use of the standard potential outcomes notation in describing the empirical approach.[29] 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 D^0_{mt} denote the amount of deforestation in municipality m during time period t if it is not on the Priority List, and let D^1_{mt} be the outcome for the same municipality if it is on the Priority List. The observed deforestation D for municipality m at time t can then be written:

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

Standard DID Model. In the standard DID model, the usual linear regression formulation is given by:

$$D_{mt} = X'_{mt}\beta + \gamma G_m + \tau \left(G_m \times I_t\right) + \delta_t + \alpha_m + \eta_{mt} \tag{2}$$

where X_{mt} is a municipality-level vector of observed factors, including prices and agro-climatic conditions (see Section 4.4); I_t is an indicator variable that equals 0 before treatment (i.e., before 2008), and equals 1 after treatment; δ_t are time dummies; α_m is a municipality-level fixed effect; η_{mt} is a time-varying unobservable factor; and (β, γ, τ) are the parameters to be estimated. The parameter τ equals the average treatment effect on the treated; given that the Priority List should reduce deforestation, one would expect $\tau \leq 0$. The ATT can be estimated consistently based on (2) provided that the common trends assumption holds – we provide formal econometric evidence below.

5.1.1 Changes-in-Changes Model

As previously noted, the CIC model developed by Athey and Imbens (2006) is a nonlinear generalization of the DID model to the entire distribution of the potential outcomes. Formally, potential

[29] Capital letters denote random variables, and lower case letters denote corresponding realized values.

deforestation D^j_{mt} – whether in the presence or absence of the policy intervention – is given by the nonparametric specification:

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

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, climate conditions, topography, etc.) in addition to time-varying unobservables. For instance, we allow for (but are not restricted to) a decomposition of the type $U_{mt} = \alpha_m + \eta_{mt}$. The function h^j allows for very flexible time trends. Because the Priority List increases monitoring and enforcement intensity, one might expect $h^1(x, u, t) \leq h^0(x, u, t)$ for any (x, u, t).

We impose four assumptions on the model. Following Athey and Imbens (2006), we first assume

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

This assumption imposes strict monotonicity of h^j on the unobservables u. It is satisfied by the DID model, which assumes u enters the function h^0 additively. Although strict monotonicity involves a loss of generality, it allows for more flexible functional forms than an additive function, particularly, flexible interactions between the time trend and the municipality-level unobservables. 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. Because conversion costs and land quality are likely to depend on factors that are unobservable to the econometrician (such as unmeasured soil quality and topography), and because different municipalities are in different stages of the deforestation process, the impact of unobserved factors on deforestation may vary systematically over time and across municipalities.

We do not restrict the way in which the functions h^j are affected by treatment status j. Municipalities that are in different stages of their deforestation process may respond differently to the policy intervention, which results 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 for more than one year – if monitoring changes based on the length of time on the list;[30] or it may take some time for potential deforesters to update their beliefs about the probability of being caught and fined.

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

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) requires X_{mt} and U_{mt} to be independent. This is stronger than the commonly adopted zero correlation assumption in linear models, but it is typically imposed in nonlinear models. Note that because the requirement conditions 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 do we need to reweight and balance them) in terms of their observable characteristics 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 holds for the non-Priority group. This is a key condition for the CIC model and plays a similar role as the common trends assumption in the standard DID model: in order to construct counterfactual predictions based on the observable distributions, it is necessary to have some form of stability over time. The assumption is less demanding than might first appear: it does not require U_{mt} to be independent over time; in fact, U_{mt} can be serially correlated (for instance, due to the presence of fixed effects). The realizations of U_{mt} may vary over time, though they must come from the same distribution. More important, 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 4.3 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 levels of 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). This in turn suggests systematic unobservable differences across groups. Assumption 2 therefore allows for policy interventions that are targeted on the group with potentially higher average benefits.

Identification. Denote by $F_{D^j_{gt}}$ the conditional distribution function of potential deforestation D^j_{mt} given $G = g$. Let the inverse distribution be given by $F^{-1}_{D^j_{gt}}(q)$ for any quantile $q \in [0, 1]$. For convenience, we also use the short-cut notation D^j_{gmt} to denote the potential outcome variable for a municipality in group $G = g$, when it is sufficiently clear from the context. 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 D^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 D_{0mt+1} (i.e., on the support of the control group at $t+1$) and is given by

$$F_{D^0_{1t+1}}(d) = F_{D_{1t}}\left(F^{-1}_{D_{0t}}\left(F_{D_{0t+1}}(d)\right)\right), \tag{3}$$

where $d \in Supp\left(D_{0mt+1}\right)$. In words, the counterfactual distribution $F_{D^0_{1t+1}}$ can be calculated based on the distribution of three observable variables: the distribution of deforestation for the same group but prior to the treatment ($F_{D_{1t}}$), and the distributions of deforestation for the control group both before and after the treatment ($F_{D_{0t}}$ and $F_{D_{0t+1}}$). Note that the distribution of D for the treated group under the treatment at $t+1$ (after treatment) is trivially identified: $F_{D^1_{1t+1}} = F_{D_{1t+1}}$. By comparing the observed $F_{D_{1t+1}}$ with the counterfactual $F_{D^0_{1t+1}}$, we can obtain 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. It uses 'double-matching' to construct the counterfactual distribution. Specifically, a treated municipality that deforested d km^2 during period t is first matched to an untreated municipality that deforested the same amount 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 d' denote the amount deforested by this last unit during $t+1$, and define $\Delta \equiv d' - d$. The difference between the 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 of the untreated units *with the same rank* before and after treatment. That is, the counterfactual deforestation of the treated unit in the absence of treatment is given by $d + \Delta$. This is similar to the adjustment in the standard DID model, though in the DID case, the adjustment is linear:

$$E\left[D^0_{mt+1}|G=1\right] = E\left[D_{mt}|G=1\right] + \left(E\left[D_{mt+1}|G=0\right] - E\left[D_{mt}|G=0\right]\right).$$

In the presence of more than one time period before treatment, there is more than one way to identify $F_{D^0_{1t+1}}$. In this case, the model becomes overidentified and the equality in (3) is testable. Note that $F_{D^0_{1t+1}}$ is identified only on the support of D_{0mt+1} for the control group at $t+1$: outside this support, $F_{D^0_{1t+1}}$ is not identified.

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

$$F_{D^1_{0t+1}}(d) = F_{D_{0t}}\left(F^{-1}_{D_{1t}}\left(F_{D_{1t+1}}(d)\right)\right), \tag{4}$$

where $d \in Supp\,(D_{1mt+1})$. Thus equation (4) provides information about treatment effects on the untreated. As before, the counterfactual distribution for the untreated $F_{D^1_{0t+1}}$ is not identified outside the support of the treated group D_{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.[31] However, we can obtain worst-case bounds in a spirit similar to Manski (2003). For instance, if $Supp\,(D_{1mt+1}) \subset Supp(D^1_{0mt+1})$, then $F_{D^1_{0t+1}}$ is identified on the subset $Supp\,(D_{1mt+1})$, and we place the remaining probability mass outside $Supp\,(D_{1mt+1})$ at the end points of $Supp(D^1_{0mt+1})$. To do so, we need prior information relating to the counterfactual support for D^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* $Supp(D^j_{gmt}) = Supp\,(D_{gmt})$ *for* $j, g = 0, 1$, and for any t.

Assumption 3 implies that while the policy intervention may affect the distribution of deforestation, it does not affect the support of the distribution. By putting all mass outside $Supp\,(D_{1mt+1})$ at the left and at the right end points of $Supp\,(D_{0mt+1})$, we obtain the lower and upper bounds for $F_{D^1_{0t+1}}$, which we denote respectively by $F^L_{D^1_{0t+1}}$ and $F^U_{D^1_{0t+1}}$ (the same reasoning applies to $F_{D^0_{1t+1}}$).[32]

Note that under Assumption 3, we cannot point identify the counterfactual distributions of both treated and untreated groups simultaneously when $Supp(D_{1mt+1}) \neq Supp(D_{0mt+1})$. Further, if $Supp(D_{1mt+1}) \subset Supp(D_{0mt+1})$, we can point identify the counterfactual distribution for the treated group $F_{D^0_{1t+1}}$, but not the control group, $F_{D^1_{0t+1}}$. In this case, we identify the average treatment on the treated, but we can only partially identify the average treatment on the untreated.[33]

Semiparametric Model. 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 Athey and Imbens (2006) suggest.

[31] We assume the respective supports of the observed variables are connected, so that if, say, $Supp\,(D_{1mt+1}) \subset Supp(D^1_{0mt+1})$, then *only* at the tails is there no information about $F_{D^1_{0t+1}}$.

[32] 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 effect of outliers, we follow the literature and trim out observations below the 3rd and above the 97th percentiles (Ginther, 2000; Lee, 2009). The empirical results are robust to the trimming. In particular, they are robust to trimming out observations below and above the percentiles $[2.5, 97.5]$ and $[3.5, 96.5]$. See Appendix F.

[33] When the opposite holds, $Supp(D_{0mt+1}) \subset Supp\,(D_{1mt+1})$, we cannot point identify the distribution of the treated in the absence of the intervention, $F_{D^0_{1t+1}}$, which implies we cannot identify the average treatment on the treated. In this case, the linear regression formulation of the DID model point identifies the ATT based on the functional form restriction.

We adopt a logit model with a single-index restriction on X_{mt}, as is common in the empirical land use literature (Stavins, 1999; Pfaff, 1999; Souza-Rodrigues, 2018). The logit model is appealing for several reasons.

First, it can be motivated by a continuum of farmers making binary choices (to deforest or not), aggregated up to the municipality level; as such, the share of deforestation flows from individual decisions – helpful in interpreting the empirical results. Second, and in contrast to the 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) – the linear model would predict negative deforestation for a non-negligible portion of the observations. Third, the logit model allows for heterogeneous effects of X_{mt} on deforestation, which is 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. Fourth, 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, which is not practical in our setting.

Formally, denote the total forested area in m at the beginning of year t by F_{mt}. The share of newly deforested area Y_{mt} is the ratio of D_{mt} and F_{mt}. 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 Assumptions 1 (i.e., strict monotonicity on u), and (b) V_{mt}^{j} satisfy the support condition in Assumption 3.

Under Assumption 4, Y_{mt}^{1} can be interpreted as the probability that a farmer deforests a plot of land in m at t when the municipality is on the Priority List (and Y_{mt}^{0} can be interpreted analogously when not).[34]

[34] A farmer's land use choice model can motivate Assumption 4. Consider a parcel of forested land i located in

By taking the log odds ratio of the share of deforestation, we obtain

$$\log\left(\frac{Y_{mt}^{j}}{1-Y_{mt}^{j}}\right)=X_{mt}'\beta+V_{mt}^{j}. \tag{6}$$

(It is easy to show that equation (6) also holds for the realized shares Y_{mt} and residuals V_{mt}.) By Assumption 2, the coefficients β are identified and can be estimated applying the ordinary least squares (OLS) estimator to equation (6). We can therefore back out, and apply the CIC model to, the residuals V_{mt}.[35]

5.2 Average Treatment Effects

We now discuss briefly how we calculate the average treatment effects. Take 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 F_{mt}) and conditional on the group $G=g$, is given by

$$E\left[D_{mt}^{j}|X_{mt},F_{mt},G_m=g\right]=\int\left[\varphi\left(X_{mt},v\right)\times F_{mt}\right]dF_{V_{gt}^{j}}\left(v\right), \tag{7}$$

where the distribution $F_{V_{gt}^{j}}$ is either observed (from the residuals of the regression (6)) or is obtained from the CIC model (i.e., from either (3) or (4) applied to the residuals V_{mt}). Given (7), 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 expectations as

$$\begin{aligned}&\int\left[\varphi\left(X_{mt},v\right)\times F_{mt}\right]dF_{V_{gt}^{j}}^{L}\left(v\right)\\ \leq\ &E\left[D_{mt}^{j}|X_{mt},F_{mt},G_m=g\right]\\ \leq\ &\int\left[\varphi\left(X_{mt},v\right)\times F_{mt}\right]dF_{V_{gt}^{j}}^{U}\left(v\right).\end{aligned} \tag{8}$$

municipality m that is under the policy regime j at time period t. Let Y_{imt}^{j} equal one if the plot is cleared and zero otherwise. The farmer deforests the plot when $Y_{imt}^{j}=1\left\{X_{mt}'\beta+v^{j}\left(U_{mt},t\right)>\varepsilon_{imt}\right\}$, where 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 U_{mt} is given by equation (5). 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.

[35] More specifically, as Athey and Imbens (2006) note, let I_{mt} be a vector of dummy variables indicating the 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 we construct the residuals with the group-time effects left in: $\log\left(\frac{Y_{mt}}{1-Y_{mt}}\right)-X_{mt}'\widehat{\beta}=I_{mt}'\widehat{\gamma}+\widehat{\nu}_{mt}$.

Bounds on average treatment effects follow naturally from (8). 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 C).

Finally, 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} is 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. To simplify, we ignore carbon decay and assume all carbon stock is immediately released into the atmosphere once a plot of land is deforested.

6 Empirical Results

In this section, we present the estimated average treatment effects of the Priority List. We first present results for the baseline case, in which possible spillovers effects are not accounted for. Then we present results when spillover effects are incorporated into the estimation strategy.

We start with the standard DID regression model specified in equation (2), and the logit model presented in equation (6); recall that the logit model is used to partial-out the covariates so that we can apply the CIC model to the residuals V_{mt}. Table 3 presents the estimated coefficients. In the first two specifications, the dependent variable is deforestation D_{mt} – the first column does not include covariates X_{mt} in the regression, while the second column does. In the third and fourth columns, the dependent variable is the log odds ratio of the share of deforestation Y_{mt} (i.e., the logit model). Column (3) does not include X_{mt}, while column (4) includes them. The model presented in the last column is our preferred specification.

We find no evidence that the common trend assumption is violated before treatment. In all specifications, the coefficients on the time dummies interacted with Priority status are not statistically significant before 2008, indicating that pre-policy trends are likely not a concern in this context. This conforms with the discussion in Section 4, where we noted that the criteria established by the Brazilian government to enter the Priority List implied a selection rule that selected municipalities based principally on the level of past deforestation, not on their trends.[36]

In all specifications considered, the Priority List appears to have reduced deforestation substantially. For the standard DID model, the coefficients on priority status after treatment are statistically significant, and show an average treatment effect on the treated of -71 km^2 in 2009 and -89 km^2 in 2010. The impacts are robust to the inclusion (and exclusion) of the covariates

[36]The pre-treatment common trend assumption is not rejected when we incorporate more years into the panel data regressions (2002–2010) – i.e., it is robust to the inclusion of a time period covering the PPCDAm structural break that occured at 2004–2005.

in the specification. Given that this group deforested about 66.4 km^2 in 2009 and 48.8 km^2 in 2010 on average, the linear DID estimates suggest that the treated group would have deforested approximately twice as much in the absence of treatment – 137 km^2 on average – both in 2009 and in 2010. The greater impact in 2010 may be the result of farmers updating their beliefs about the regime change.[37]

The logit model paints a similar picture: significant impacts of the policy intervention, and greater effects during the second year of the program. Before we turn to the CIC estimates, we discuss briefly the coefficients on the covariates in the main specification. The impact of rainfall on deforestation is significant and hump-shaped, which is reasonable given that both low and high levels of rainfall make agriculture unattractive. The estimates indicate that higher temperature induces more deforestation, but not significantly so; and larger shares of protected areas result in significantly less deforestation, as expected. Increases in the price of beef lead to more deforestation, also as one would expect.[38] Finally, municipal GDP does not have significant impacts on deforestation.

As mentioned previously, the logit model constitutes the first step in estimating the CIC model. Table 4 presents the estimated treatment effects based on the CIC model, as explained in Section 5. Columns (2), (3), and (4) of Table 4 present the estimated ATT, ATU, and ATE, respectively. In the top panel, we present results for the estimated effects on deforestation for the years 2009, 2010, and for the total cumulative deforestation in those two consecutive years. We show results using 2006 and 2007 as alternative baseline years – recall that the CIC model is overidentified 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 effects on carbon emissions. It also shows 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 parentheses are 95 percent confidence intervals.[39]

We focus first on 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 -21 km^2 and -25 km^2 (depending on the baseline), and between -50 km^2 and -54 km^2 for 2010. The pattern here is similar to the standard DID regression model (i.e., increasing effects over time), but the magnitudes

[37] We do not find substantial differences in the relationship between fines issued and contemporaneous deforestation for 2009 and 2010. Fines per incremental deforestation in 2010 is slightly smaller than in 2009.

[38] The coefficient on price of crops, in contrast, is negative, though not statistically significant. (This may reflect the fact that the vast majority of deforested areas in the Amazon are converted to pasture for cattle grazing.)

[39] 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 (see discussion below). For these cases, the confidence intervals are based on Imbens and Manski (2004) procedure for the parameter of interest (not for the identified set), where the bootstrap replications are used to compute the standard errors for the lower and upper bound estimators. We implemented 500 bootstrap replications.

differ. As is well known, the DID estimator is not scale-invariant, so there is no *a priori* reason to expect the DID and the CIC estimators to generate similar point estimates. Further, the substantial fraction of the municipalities with low levels of deforestation in the data suggests that the linear DID specification may result in biased estimates in the present context.

According to the CIC estimates using 2006 as the baseline year (which provides 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 avoided the clearing of 2,540 km^2 of forested area, and the emissions of 30 million tons of carbon in the same period. The estimated social benefit of the program in terms of avoided emissions is approximately \$2,21 billion. Compared to the budget allocated to IBAMA and INPE – about \$600 million – the results suggest the program was welfare-enhancing, and that further investments in monitoring efforts would be worthwhile.

Treatment effects on the untreated are not point identified, but the identified sets are informative.[40] The average treatment effects on the untreated are statistically significant and robust to the choice of the baseline year. The 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. It suggests that the government indeed selected municipalities with potentially higher average impacts. A DID strategy could not deliver such results because it only identifies effects on the treated. Extrapolating results from the treated group to the untreated under the assumption of homogeneous effects would upwardly bias the effects on the untreated.

Spillover Effects. As discussed previously, to incorporate the potential spillover effects of the Priority List, we split the control group in two, depending on whether an untreated 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 the two criteria are referred to as the 'Spillover' group.[41]

[40] 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 12 in Appendix G). This implies that the counterfactual distribution $F_{V^1_{0t}}$ is identified only on a subset of its support, and it is not identified at the tails. This means the ATU can only be partially identified. (See Figures 12 and 13 in Appendix G for the estimated factual and counterfactual distribution functions of the residuals V^j_{gt}, for both treated and control groups in 2009 and in 2010.)

[41] Appendix F presents a robustness analysis when we change the definition of the Spillover group.

We start with the logit regression model (with and without covariates). Table 5 presents the estimated coefficients. As before, the common trends assumption is not violated in the pre-treatment period in both specifications, and for treatment and spillover groups.[42] The estimated coefficients on both Priority and Spillover status after the treatment are negative and statistically significant. This suggests the presence of spillover effects: the untreated municipalities with a treated neighbor and high levels of past deforestation appear to reduce their deforestation rates in response to the establishment of the Priority List.

Table 6 presents the estimated treatment effects based on the CIC model. The ATTs are now slightly greater than the ones estimated ignoring the potential spillover effects. This is expected: the control group now shows lower average reductions in deforestation after treatment given that it does not include the municipalities that are more likely to respond to the intervention. 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 the potential spillovers effects. The average treatment effects on the Spillover group, denoted here by ATS, are statistically significant and robust to the baseline year. While the ATSs are partially identified, the estimated sets are very informative – e.g., one of them contains a singleton in 2009 (baseline year 2007). 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 are between the estimated ATT and ATU, constituting further evidence of heterogeneous effects.

7 Optimal Policy Targeting

7.1 Optimal Policy Targeting Framework

Suppose a policy maker wants 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 (as well as the covariates). Denote the counterfactual assignment rule by $\phi_t = (\phi_{1t}, ...\phi_{Mt})$, which assigns the treatment to municipalities $m = 1, ..., M$ and which can be either probabilistic $\phi_{mt} \in [0, 1]$ or deterministic $\phi_{mt} \in \{0, 1\}$. For a given time period t, the

[42] The pre-treatment common trend assumption is rejected in the linear DID regression model when spillovers are taken into account.

policy maker wants to solve the problem

$$\min_{\phi_t \in [0,1]^M} \sum_{m=1}^{M} \left[\phi_{mt} E\left[D^1_{mt} | X_{mt}, F_{mt}, G_m \right] + (1-\phi_{mt}) E\left[D^0_{mt} | X_{mt}, F_{mt}, G_m \right] \right]. \tag{9}$$

The minimum deforestation is (trivially) achieved by a singleton rule that allocates m to the treatment when $E\left[D^1_{mt}|X_{mt}, F_{mt}, G_m\right] \leq E\left[D^0_{mt}|X_{mt}, F_{mt}, G_m\right]$.[43] The problem (9) abstracts from two important considerations, however.

The first involves constraints. The original Priority List has the intention of directing limited resources where they are expected to have the greatest impact. Given that data on the resources that were effectively allocated 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 $\overline{A}$ that can be monitored under the Priority List:

$$\sum_{m=1}^{M} a_m \times \phi_{mt} \leq \overline{A}, \tag{10}$$

where a_m is the area of municipality m. Presumably, the costs of monitoring and punishing illegal deforestation increase with the total area covered by the policy. The alternative constraint applies to the total number of municipalities $\overline{M}$ that can be placed on the list:

$$\sum_{m=1}^{M} \phi_{mt} \leq \overline{M}. \tag{11}$$

This constraint may be reasonable when the monitoring costs depend mainly on the number of districts that the inspectors must visit to issue the fines.[44]

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

[43] Any random allocation is optimal when $E\left[D^1|X, F, G\right] = E\left[D^0|X, F, G\right]$.

[44] Ideally, we would like to observe or identify the expected monitoring costs for each municipality m in each time period t both in the absence, or presence, of the treatment. Then, we could replace the constraints (10) and (11) by 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 could use $R^j = P \times D^j + MC^j$ in the social cost function (9), where P is the social cost of deforesting one parcel of land. This approach is unfeasible, however, because we do not have information about the true budget constraint K_t. Although we do know IBAMA's and INPE's total budgets, we do not know how much of the total is allocated to monitoring. In addition, we do not have data indicating how monitoring costs are distributed across municipalities.

minimize the worst-case scenario. By adopting this criterion, the policy maker can do no worse than achieve the best of the worst outcomes.

Formally, let all the feasible values that $E[D^j_{mt}|X_{mt}, F_{mt}, G_m]$ can take be indexed by $\gamma \in \Gamma$ (given by the inequality (8)). The policy maker's problem under the minimax criterion is

$$\min_{\phi_t \in [0,1]^M} \sup_{\gamma \in \Gamma} \sum_{m=1}^{M} \left[\phi_{mt} E_\gamma \left[D^1_{mt}|X_{mt}, F_{mt}, G_m \right] + (1-\phi_{mt}) E_\gamma \left[D^0_{mt}|X_{mt}, F_{mt}, G_m \right] \right] \tag{12}$$

subject either to the total area constraint (10), or to the total number of municipalities constraint (11). The minimization problem (12) subject to either constraint is a linear programming problem that is straightforward to solve numerically. In the empirical exercise, when using constraint (10), we set $\overline{A}$ 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 (11), 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.[45]

Intuition for assignment to the optimal list starts from the observation that, in the absence of any constraint, the minimum deforestation is achieved by a simple rule. For a municipality m that was originally in the control group $G_m = 0$, the policy maker assigns it to the optimal list when the observed (expected) deforestation in the absence of treatment, $E\left[D^0_{mt}|X_{mt}, F_{mt}, G_m = 0\right]$, is greater than the maximum possible amount of expected deforestation under the treatment. For a municipality m that was originally in the treatment group, $G_m = 1$, it should not be on the list when the observed (expected) deforestation under treatment, $E\left[D^1_{mt}|X_{mt}, F_{mt}, G_m = 1\right]$, is greater than the maximum possible amount of deforestation in the absence of treatment. Note that the assignment rule differs depending on the observed priority status because the objects that are partially identified differ. When constraints are taken into account, the estimated magnitudes of the treatment effects for *all* municipalities matter in the minimization problem.[46]

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

[45] 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 $\binom{490}{35} \approx 4 \times 10^{53}$.

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

To simplify the exposition, we describe how the ex-post optimal list is calculated in the presence of spillover effects in Appendix D. Here, we only emphasize that the objective function in this case 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).

7.2 Optimal Policy Targeting Results

Given the estimated treatment effects from the previous section, we now investigate the ex-post optimal assignment rule, as discussed in Section 7.1. Table 7 compares the original Priority List with the ex-post optimal list obtained by solving the relevant constrained minimization: the left panel considers the total area $\overline{A}$ that can be monitored as the constraint (see equation (10)), while the right panel restricts the number of municipalities $\overline{M}$ as the constraint (see equation (11)).[47]

Overall, the proportion of municipalities that appear on both lists is very 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 ex-post optimal lists.

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. In contrast, when the restriction is placed on 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.

Figure 9 presents the geographic distribution of municipalities on the lists. 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 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 figures in the bottom panels also include the protected areas.)

Two interesting patterns emerge from these maps – features that were not imposed by the estimation strategy. First, the overlap in the areas covered by the protected areas and by the area-constrained counterfactual list is much smaller than the overlap of the protected areas and the

[47] We present results using the baseline year 2006. Results for the baseline 2007 are similar.

original Priority List. Indeed, the former comprises an area that is just a half of the latter. This suggests the presence of important complementarities in these two policies that could be further leveraged by the Brazilian government. Second, the geographic location of the area-constrained counterfactual list suggests a protective shield acting close to the deforestation frontier. That deforestation frontier, the 'Arc of Deforestation,' is located along the southeastern edge of the Amazon Biome. The Priority List may therefore serve to complement the protected areas in impeding the deforestation process from continuing into the more pristine regions. The geographic shield may be particularly beneficial in the longer term.

We can quantify the impacts of targeted policies by comparing both the maximum possible deforestation and the carbon emissions achieved under the optimal list with the corresponding outcomes under the Priority List as well as another benchmark: a list composed of municipalities that are selected randomly.[48] Table 8 presents the results. Compared to the area-constrained optimal list, the Priority List results in about 6 percent more deforestation and 5 percent more carbon emissions in 2009–2010. The estimated avoided emissions translates into a social value of at least approximately \$562 million for that two-year span only. As the difference between the value of the carbon emissions comparing the optimal and observed list measures the minimum willingness to pay of the policy maker (or society) to obtain the ex-post information about treatment effects, we estimate high social returns in investments generating information about the effects of conservation policies.

We find that the performance of the ex-post optimal list that is restricted to consider 35 municipalities is 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 this case. Randomly selecting $\overline{M} = 35$ municipalities onto the list would result, on average, in 23–25 percent more deforestation and 26–29 percent more emissions than the number-of-municipalities-constrained optimal.

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 larger. 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 presented here, the estimated performance of the Priority List is (perhaps surprisingly) quite 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 room for improvement.

Next, we explore briefly how much the minimax solution for carbon emissions can be affected

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

by relaxing the constraints we have been imposing. Figure 10 shows the results. The left panel presents the level of emissions at the optimum for the total area constraint, while the right panel shows the results when we change the number of municipalities that can be included on the optimal list.

In each figure, the vertical lines show, respectively, the maximum $\overline{A}$ and $\overline{M}$ that correspond to the area covered by, and the number of municipalities on, the Priority List. 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 eventually level off for large $\overline{A}$. So, the benefits of including additional municipalities on the list decreases with $\overline{A}$. Because monitoring costs should increase with $\overline{A}$, concentrating efforts on a strategically selected subregion of the Amazon rainforest seems to be a suitable policy. (The same reasoning applies when we change the number of municipalities allowed in the optimal list $\overline{M}$.)

Spillover Effects. We now discuss the ex-post optimal lists when spillovers are incorporated into the minimization problem. Table 9 compares the Priority List with the ex-post optimal lists based on the two different constraints discussed above. With the constraint being the total area, the optimal list replaces a greater number of large municipalities by small municipalities when compared to the optimal list with no spillovers. Such an assignment takes advantage of the fact that fewer large municipalities placed on the list can affect deforestation in important adjacent locations.

Figure 11 presents the geographic distribution of the optimal lists. When compared to the no-spillover case in Figure 9, the area-constrained optimal list is more geographically dispersed and less contiguous. That is because of the spillover effects: putting all targeted municipalities together does not exploit the potential reduction in deforestation in adjadent locations resulted from the spillover effects. A similar pattern is observed in the number-of-municipalities constrained optimal lists.

Finally, Table 10 compares the levels of deforestation and emissions of the alternative lists. Because the optimal lists now take advantage of potential spillover effects, they can achieve lower levels of forest losses in the worst-case scenario. The Priority List now results in about 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 puts 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 allowed. Randomly selected municipalities now result in 34–44 percent more emissions than the number-of-municipalities constrained optimal.

8 Conclusion

This paper has examined the efficacy of targeted blacklist-type policies for slowing deforestation, a primary contributor to global carbon emissions and a source of considerable concern. 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 transparent program evaluation methods: deforestation was cut by 40 percent in municipalities placed on the list (relative to the case in which no policy was introduced). Then, we applied the changes-in-changes approach of Athey and Imbens (2006) to compare the actual Priority List with the ex-post optimal list. Here, we were able to show that an optimally targeted list could reduce deforestation further. Carbon emissions would be at least 8 percent higher under the Priority List, though emissions under the actual list would still be significantly lower than under a randomly chosen set of municipalities.

From a regulation perspective, our approach and findings help shed light on the gains to the environment from switching to counterfactual targeting policies. More generally, our counterfactual approach using ex-post treatment effects should be applicable in a variety of other settings, helping policy makers to assess which policy configurations are likely to have most environmental impact.

References

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.

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 Papers, Department of Economics 2016-32, 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.

Assunção, J., C. Gandour, and R. Rocha. (2013). Detering deforestation in the Brazilian Amazon: Environmental monitoring and law enforcement. *Climate Policy Initiative Working Paper Series*.

Assunção, J. and R. Rocha (2014). Getting greener by going black: The priority municipalities in Brazil. *Climate Policy Initiative Working Paper Series*.

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.

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.

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.

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*.

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.

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.

Chomitz, K. M. and D. A. Gray (1996). Roads, land use, and deforestation: A spatial model applied to Belize. *The World Bank Economic Review 10*(3), 487–512.

Chomitz, K. M. and T. S. Thomas (2003). Determinants of land use in Amazonia: A fine-scale spatial analysis. *American Journal of Agricultural Economics 85*(4), 1016–1028.

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*, 321328.

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.

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.

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.

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.

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

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.

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

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

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.

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.

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*.

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

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

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*.

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

Nelson, G. and J. Geoghegan (2002). Modeling deforestation and land use change: sparse data environments. *Agric. Econ. 27.*

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.

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*, 26.

Pinheiro, T. F., M. I. S. Escada, D. M. Valeriano, P. Hostert, F. Gollnow, and H. Müller (2016). Forest degradation associated with logging frontier expansion in the Amazon: the BR-163 region in southwestern Pará, Brazil. *Earth Interactions 20(17)*, 1–26.

Simonet, G., J. Subervie, D. Ezzine-de Blas, M. Cromberg, and A. E. Duchelle (2018). Effectiveness of a REDD+ project in reducing deforestation in the Brazilian Amazon. pp. aay028.

Souza-Rodrigues, E. (2018). Deforestation in the Amazon: A unified framework for estimation and policy analysis. *Working Paper, University of Toronto*.

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*.

A Tables and Figures

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

	Total Sample ($N = 490$)		Treated Group ($N = 35$)		Control Group ($N = 455$)	
Variable	Mean	SD	Mean	SD	Mean	SD
Satellite Images						
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	13,954
Total Area	8,726	16,716	21,815	29,409	7,719	14,899
Policy Measures						
Fines Issued	9	19	40	44	7	14
New Protected Area	333	3,317	655	2,701	308	3,361
Agriculture and Ranching						
GDP (thousands)	179	1,013	180	399	178	1,045
Agricultural GDP (thousands)	19	24	38	22	17	24
Cattle (thousands)	105	148	363	290	85	108
Crop Area	109	455	289	563	95	443
Other Variables						
Rainfall	22	6	19	2	22	6
Temperature	26	1	26	1	26	1
Carbon Stock in Forested Area (tC/ha)	189	68	212	35	187	69
Carbon Stock in Deforested Area (tC/ha)	117	48	101	23	118	49

Notes: Areas are measured in square kilometres. All monetary figures are expressed in December 2013 Reais.

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

	Treated Group ($N = 35$)		Spillover Group ($N = 24$)		Untreated Group ($N = 431$)	
Variable	Mean	SD	Mean	SD	Mean	SD
Satellite Images						
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
Policy Measures						
Fines Issued	40	44	23	20	6	13
New Protected Area	655	2,701	0	0	325	3,452
Agriculture and Ranching						
GDP (thousands)	180	399	210	380	177	1,071
Agricultural GDP (thousands)	38	22	40	46	16	21
Cattle (thousands)	363	290	248	126	76	99
Crop Area	289	563	250	396	87	444
Other Variables						
Rainfall	19	2	20	3	22	6
Temperature	26	1	26	2	26	1
Carbon Stock in Forested Area (tC/ha)	212	35	203	45	187	70
Carbon Stock in Deforested Area (tC/ha)	101	23	97	24	119	50

Notes: This table breaks the Control Group from Table 1 into Spillover and Untreated Groups. Areas are measured in square kilometres. All monetary figures are expressed in December 2013 Reais.

Table 3: Regression: Priority Status on Deforestation in the Amazon Biome

	(1) Level	(2) Level	(3) Logodds	(4) Logodds
Treated Group x Year=2006	-3.120	-5.924	-0.0596	-0.127
	(-0.09)	(-0.18)	(-0.27)	(-0.73)
Treated Group x Year=2007	9.729	10.10	-0.0208	-0.0119
	(0.26)	(0.28)	(-0.09)	(-0.07)
Treated Group x Year=2009	-71.81*	-71.67*	-1.152***	-1.132***
	(-2.36)	(-2.47)	(-4.74)	(-5.89)
Treated Group x Year=2010	-89.35**	-88.67***	-1.349***	-1.299***
	(-3.29)	(-3.43)	(-5.22)	(-6.04)
Lagged Rainfall		1.589**		0.162***
		(2.89)		(3.67)
Lagged Rainfall Squared		-0.0664***		-0.00488***
		(-6.36)		(-5.44)
Lagged Temperature		0.704		0.0425
		(1.10)		(1.60)
Share of Protected Areas		7.058		-2.414***
		(1.94)		(-17.18)
Price of Beefs Lagged		-0.167***		0.00372**
		(-6.60)		(2.58)
Price of Crops Lagged		-4.314*		-0.0144
		(-2.50)		(-0.12)
Lagged GDP		0.000000298		-2.29e-09
		(0.68)		(-0.07)
Time Dummies	YES	YES	YES	YES
State Dummies	NO	YES	NO	YES
R^2	0.330	0.383	0.042	0.447
Obs	2450	2450	2450	2450

Notes: Constant term, and coefficients on state and time dummies are omitted.
t statistics in parentheses, based on Huber-Eicker-White robust standard errors.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 4: Average Treatment Effects, CIC Model – 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)

Note: 95% confidence intervals are in parenthesis. For ATT and CTT, they were computed based on standard bootstrap. For ATU, ATE, CTU, and CTE, they were based on Imbens and Manski (2004).
Deforestation is measured in squared kilometers.
Emissions are measured in millions of tons of carbon.
Value is measured in billion U$, assuming U$ 20/tCO2.
The calculation uses the fact that 1 tC = (44/12) tCO2.

Table 5: Regression: Priority Status and Spillover Effects on Deforestation in the Amazon Biome

	(1) Logodds	(2) Logodds
Treated Group x Year=2006	-0.0596	-0.133
	(-0.27)	(-0.76)
Treated Group x Year=2007	-0.0208	-0.00672
	(-0.09)	(-0.04)
Treated Group x Year=2009	-1.152***	-1.131***
	(-4.73)	(-5.97)
Treated Group x Year=2010	-1.349***	-1.297***
	(-5.21)	(-6.11)
Spillover Group x Year=2006	0.0265	-0.291
	(0.10)	(-1.59)
Spillover Group x Year=2007	0.0463	-0.178
	(0.17)	(-0.82)
Spillover Group x Year=2008	0.0988	-0.175
	(0.37)	(-0.89)
Spillover Group x Year=2009	-0.880**	-1.163***
	(-3.22)	(-6.42)
Spillover Group x Year=2010	-0.982***	-1.216***
	(-3.37)	(-4.95)
Lagged Rainfall		0.143**
		(3.20)
Lagged Rainfall Squared		-0.00446***
		(-4.95)
Lagged Temperature		0.0435
		(1.72)
Share of Protected Areas		-2.390***
		(-17.02)
Price of Beefs Lagged		0.00413**
		(2.86)
Price of Crops Lagged		0.00424
		(0.04)
Lagged GDP		-4.55e-09
		(-0.13)
Time Dummies	YES	YES
State Dummies	NO	YES
R^2	0.059	0.452
Obs	2450	2450

Notes: Constant term, and coefficients on state and time dummies are omitted.
t statistics in parentheses, based on Huber-Eicker-White robust standard errors.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 6: Average 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)

Note: 95% confidence intervals are in parenthesis. For ATT and CTT, they were computed based on standard bootstrap. For ATU, ATS, ATE, CTU, CTS, and CTE, they were based on Imbens and Manski (2004).
Deforestation is measured in squared kilometers.
Emissions are measured in millions of tons of carbon.
Value is measured in billion U$, assuming U$ 20/tCO2.
The calculation uses the fact that 1 tC = (44/12) tCO2.

Table 7: Percent of Correctly Predicted, Ex-post Deforestation Optimal Rule

Constraint:	*Total Area*			*Number of Municipalities*		
	Optimal			*Optimal*		
	0	1		0	1	
Observed			*Percent Correct*			*Percent Correct*
0	382	73	83.96	439	16	96.48
1	7	28	80.00	16	19	54.29
		Overall	83.67		*Overall*	93.47

Note: Baseline year is 2006

Table 8: Compare Ex-Post Optimal, Priority, and Randomly Selected Lists

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

Note: "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 U$, assuming U$ 20/tCO2.

Table 9: Percent of 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	80.22	436	19	95.82
1	12	23	65.71	19	16	45.71
		Overall	79.18		*Overall*	92.24

Note: Baseline year is 2006

Table 10: Compare 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

Note: "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 U\$, assuming U\$ 20/tCO2.

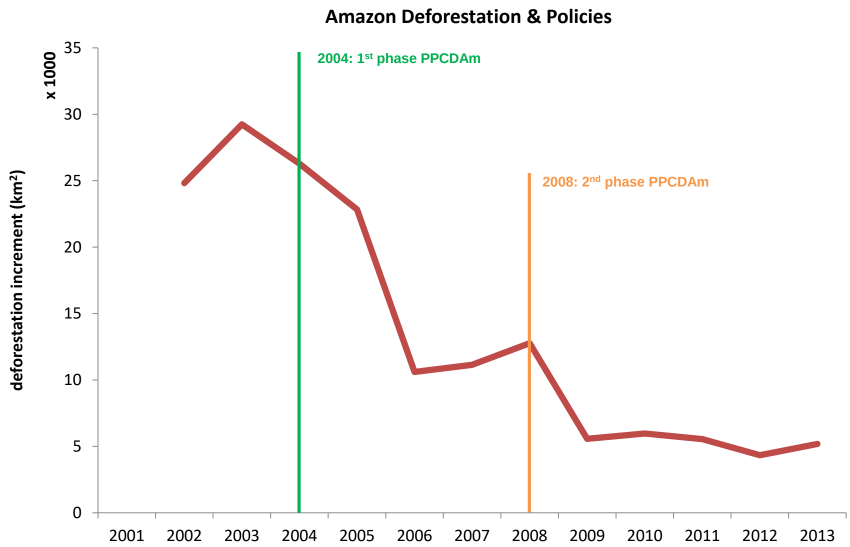


Figure 1: Deforestation Rate and Policy Changes

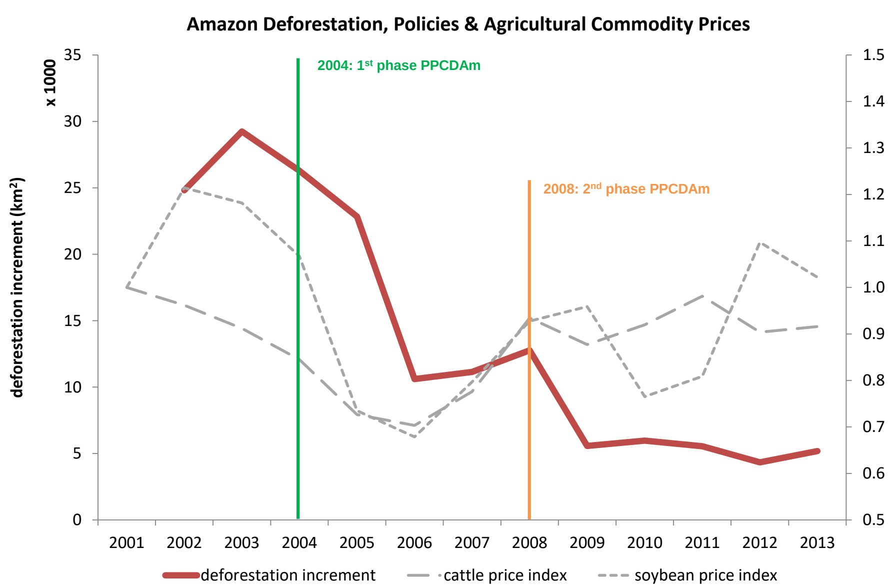


Figure 2: Deforestation Rate and Prices by year

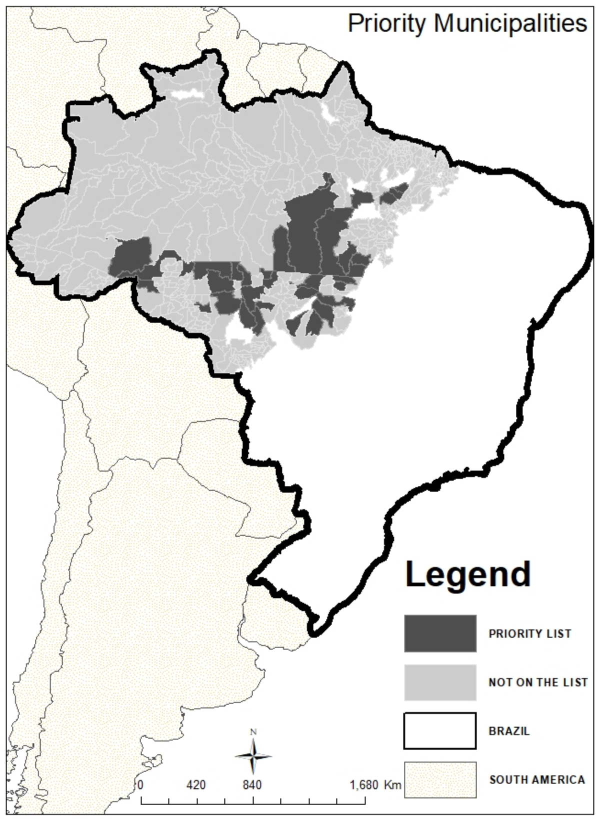
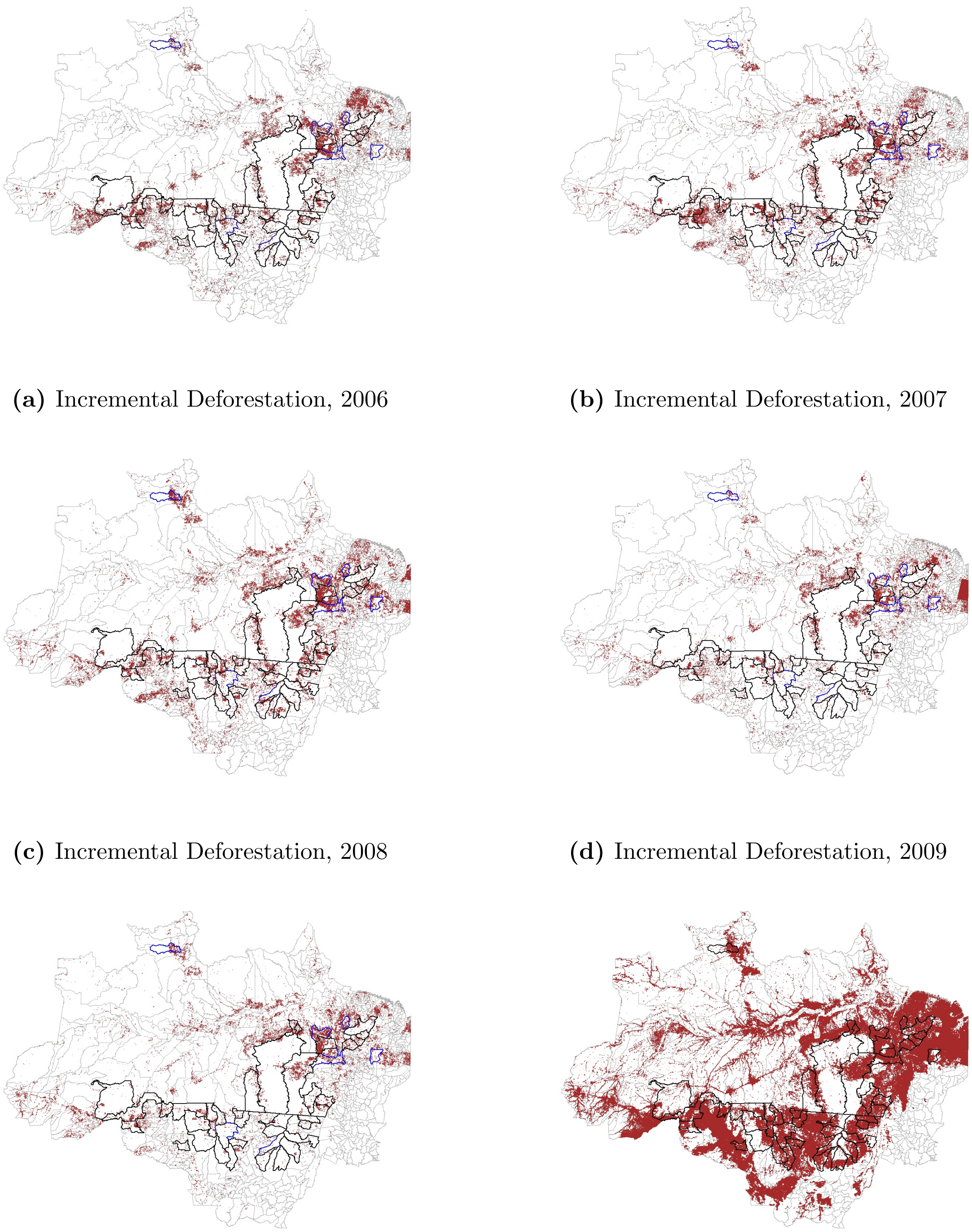


Figure 3: Map of Brazil, Amazonia, and the Location of the Priority List



(a) Incremental Deforestation, 2006

(b) Incremental Deforestation, 2007

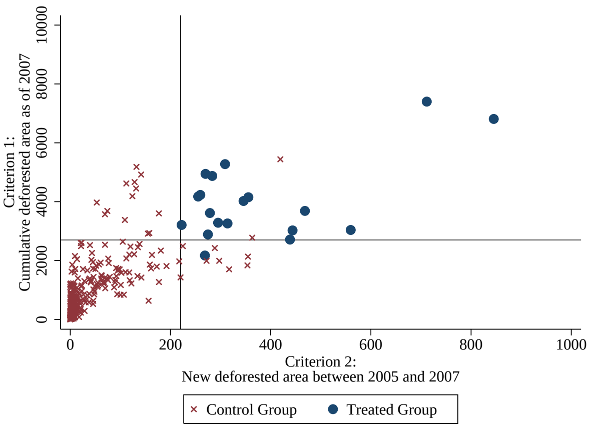
(c) Incremental Deforestation, 2008

(d) Incremental Deforestation, 2009

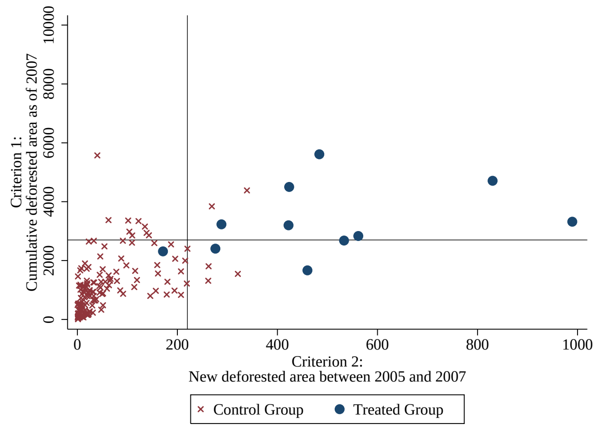
(e) Incremental Deforestation, 2010

(f) Cumulative Deforestation, 2010

Figure 4: Map of Priority Municipalities and Deforestation, 2006 and 2010

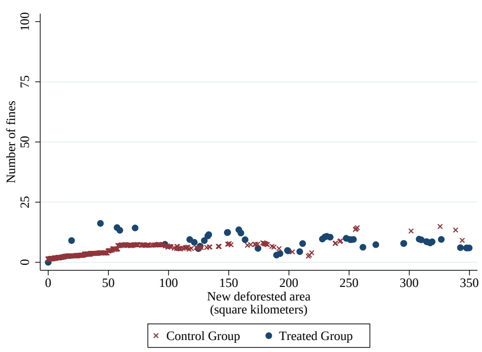


(a) Combinations of Z^1 and Z^2, given $Z^3 = 0$

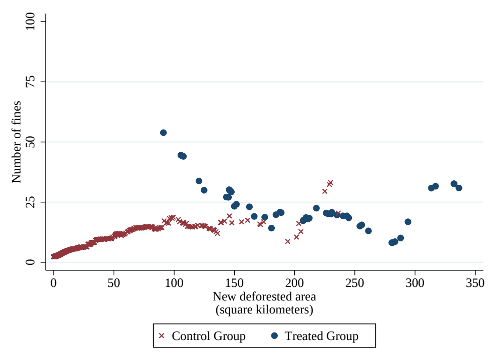


(b) Combinations of Z^1 and Z^2, given $Z^3 = 1$

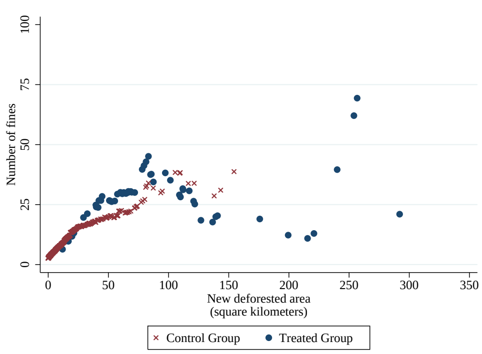
Figure 5: Selection into Priority List in 2008. Combinations of Z_{mt-1}



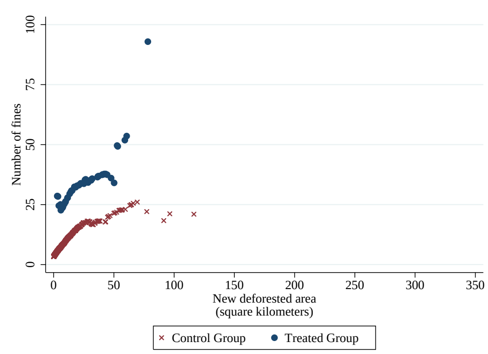
(a) 2002-2003



(b) 2004-2005



(c) 2006-2007



(d) 2009-2010

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

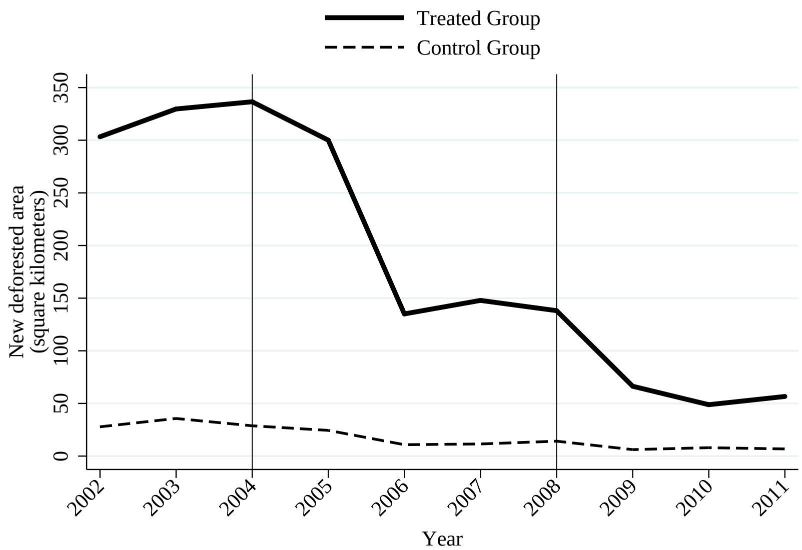


Figure 7: Within-group average of new deforested area, by year

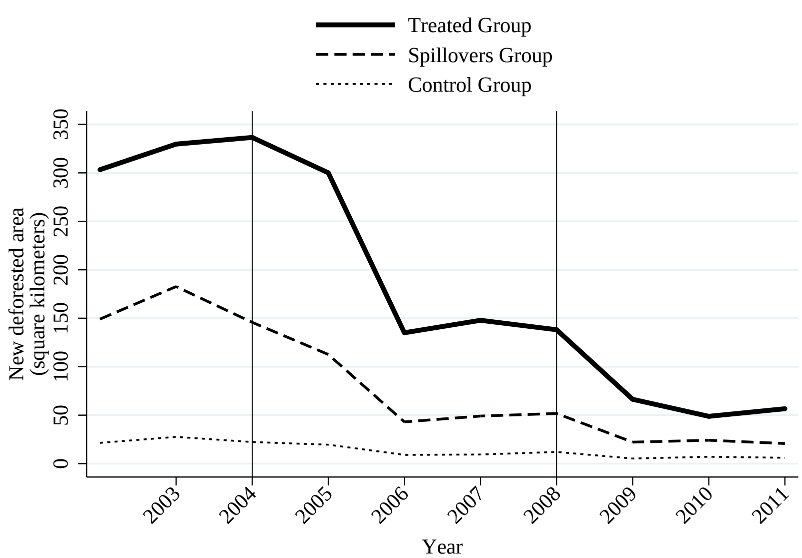
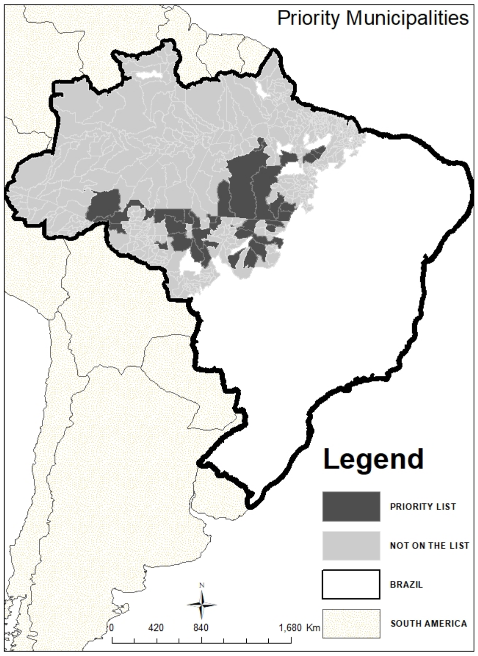
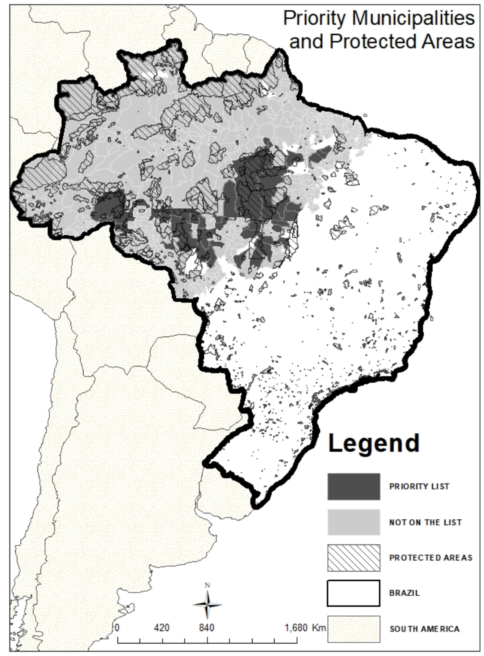


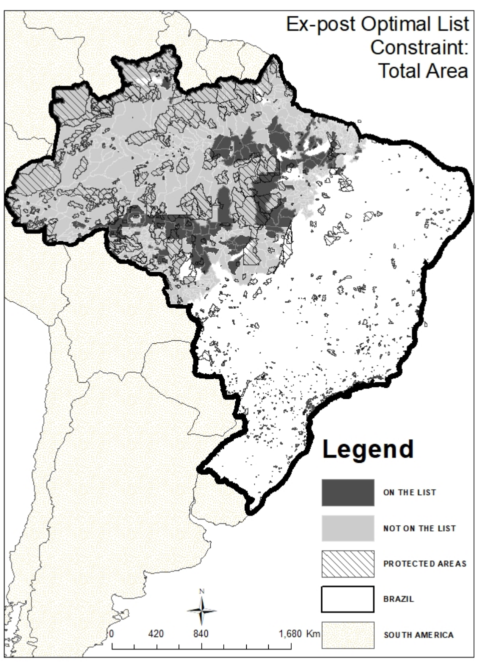
Figure 8: Within-group average of new deforested area, by year



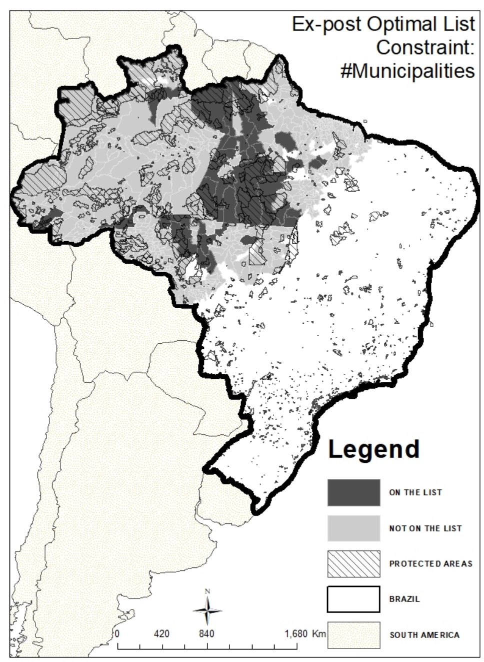
(a) Priority List



(b) Priority List and Protected Areas

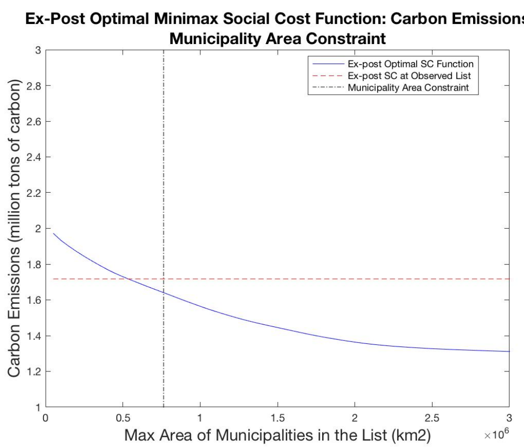


(c) Optimal List, Total Area

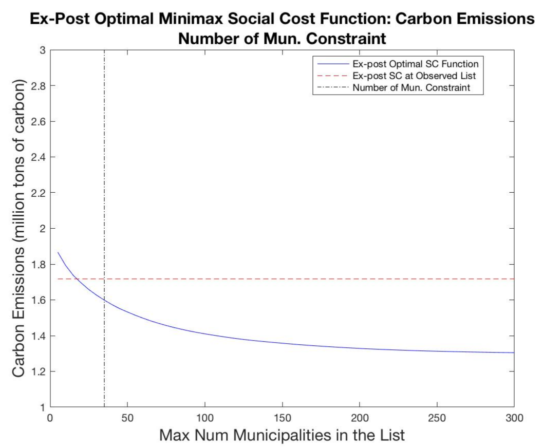


(d) Optimal List, Number of Municipalities

Figure 9: Location of Priority and Ex-post Optimal Lists

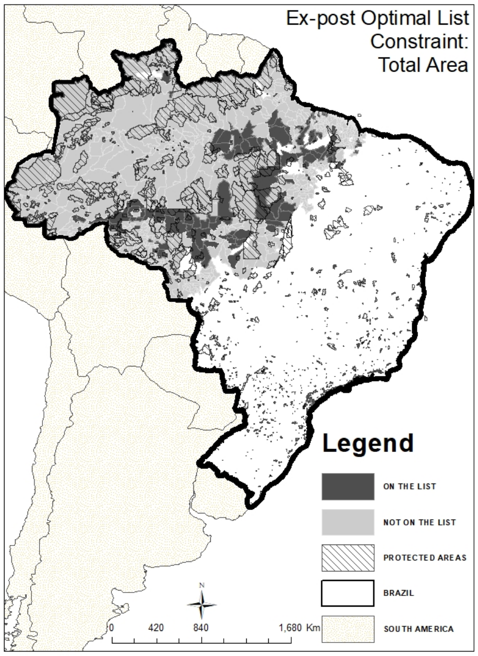


(a) Constraint: Total Area

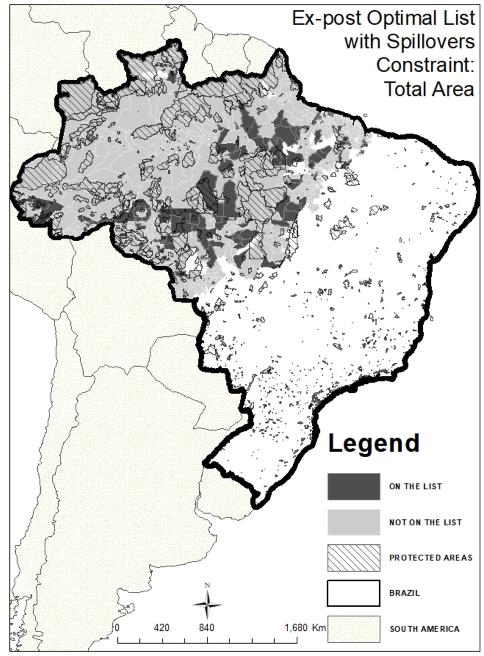


(b) Constraint: Number of Municipalities

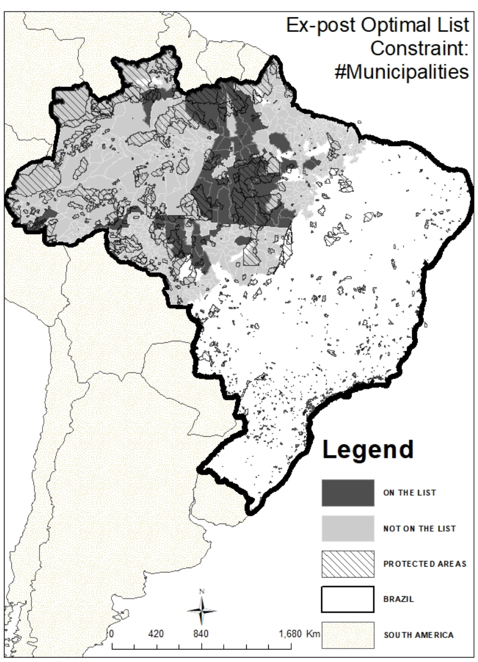
Figure 10: Ex-Post Minimax Carbon Emissions, when Varying the Constraints



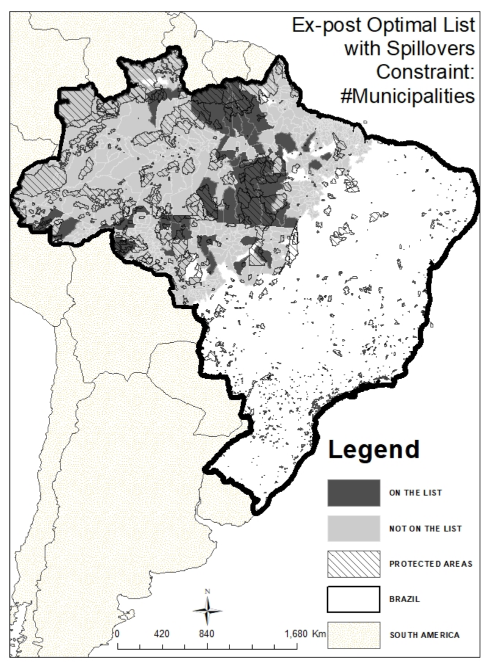
(a) No Spillover Optimal List, Total Area



(b) Spillover Optimal List, Total Area



(c) No Spillover Optimal List, Number of Municipalities



(d) Spillover Optimal List, Number of Municipalities

Figure 11: Location of Priority and Ex-post Optimal Lists with Spillover Effects

B Appendix: Data

In this section, we provide details about the data sources, and the construction of the variables used in this study. All municipalities with a positive fraction of their area in the Amazon Biome and with complete information about all covariates are included in the sample.

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 is publicly available at both pixel and municipality levels.[49]

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). Land cover classification is performed in several steps. First, because deforestation occurs typically during the dry season, INPE selects LANDSAT images between July and September with minimal clouds coverage (the spatial resolution is 60×60 meters). 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.[50] INPE then groups adjacent pixels in larger regions based on their spectral similarities. After the image segmentation, it implements a cluster unsupervised classification algorithm to generate the land cover classifications. Finally, it verifies the classifications and calculates the deforestation rate. Annual deforestation rates 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 the subsequent years.

INPE's classification centers efforts on detecting deforestation. Observed remaining forests, however, 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

[49] In each year, a small amount of land area is unobserved due to cloud cover. In our data, the average share of clouds cover on 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.

[50] The pixel components considered are soil, vegetation, and shade. The images-fraction shade and soil help in the process of identification of deforested areas; images-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; and finally the images-fraction soil helps detect transition/contact areas between forest formations and those of cerrado (Camara et al., 2006).

measure remaining forest as the remaining available area in the municipality. I.e., the total municipal area minus the non-forested areas, the water bodies, and the previous cumulative deforested areas. This guarantees consistency over time (the correlation between our proxy and the PRODES remaining forested area is 0.99). We dropped observations with small 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 substantially relevant for policies focused on preventing deforestation.[51] 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 to be 0.01 km^2 (this is analogous to setting minimum shares in logit models to be greater than or equal to a small strictly postive number $\epsilon > 0$, for example, $\epsilon = 10^{-4}$ or $\epsilon = 10^{-5}$).

Priority status. The official list of Priority municipalities, with precise dates for entry and exit from the list, comes from the Ministry of the Environment. Because there are few municipalities entering and exiting the blacklist from 2009 on, there is not much that can be said with high level 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. Efforts 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. Crop cultivation and cattle ranching have historically been important drivers of deforestation in the Amazon. To help disentangle the effects of changing commodity prices on deforestation from the effects of policy interventions, we construct a single price index for each municipality based on pre-determined, cross-sectional variation in crop mix across municipalities, and time-series variation in commodity prices received by producers in the southern Brazilian state of Paraná, which – unlike prices received by producers further north in the Amazon – are exogenous to the policy interventions we wish to evaluate. Data on prices received by producers of soy, rice, corn, cassava, cane sugar, beans, and beef in the southern state of Paraná are taken from the State Secretariat for Agriculture and Food Supply. Prices are deflated to year 2011 BRL. Municipality-level data on the amount and value of each form of agricultural production, which we use to weight the Paraná prices,

[51]In addition, these small municipalities present implausible substantial oscilations on shares of deforestation over time because of the small values used in the denominator in the calculation of the shares.

come from surveys administered by the Brazilian Institute of Geography and Statistics (IBGE), namely the Municipal Crop Survey and the Municipal Livestock Survey. Specifically, for crops, the weights are given by the shares of municipal area used as farmland for the different crops. For beef cattle, the weight is given by the ratio of heads of cattle to municipal area (given that annual pasture area is not observable). We fix the weights at 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 occurred in 2004–2005 with the first phase of PPCDAm.

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 participation are based on spatial interpolation of climate data from a large number of monitoring stations operating in South America and elsewhere.

Municipalities' Gross Domestic Product. Annual data on municipalities' total and agricultural GDPs 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 IBGE's surveys: the Municipal Livestock Survey and the Municipal Crop Survey.

Carbon Stock. The amount of carbon stock aboveground is calculated by Baccini et al. (2012). We combined their master data of carbon stock with the PRODES data to calculate the average carbon stock in forested and deforested areas in each municipality.[52]

Fines. Data on the number of fines issued for environmental offenses come from IBAMA. The information was collapsed 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.

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

C Appendix: Dynamic Treatment Effects

In this section, we discuss briefly how dynamic treatment effects are incorporated into the analysis. Incorporating dynamics 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, while $t+1$ 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 the previous period $t+1$, first note that, for any level of deforestation d, there exists a unique v such that $d = \varphi(X, v) \times F$. Then, potential deforestation D^j at $t+2$ conditioning on the potential deforestation D^l at $t+1$ at level d, for $j, l = 0, 1$, and for group $G = g$, is

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

By the Law of Iterated Expectations,

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

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

$$\begin{aligned}
& \int \left[\int \left[\varphi\left(X_{mt+2}, v'\right)\left(1 - \varphi\left(X_{mt+1}, v\right)\right) F_{mt+1}\right] dF^L_{V^j_{gt+2}}\left(v'\right)\right] dF^L_{V^l_{gt+1}}(v) \\
\leq \; & E\left[D^j_{mt+2} | X_{mt+2}, X_{mt+1}, F_{mt+1}, G_m = g\right] \\
\leq \; & \int \left[\int \left[\varphi\left(X_{mt+2}, v'\right)\left(1 - \varphi\left(X_{mt+1}, v\right)\right) F_{mt+1}\right] dF^U_{V^j_{gt+2}}\left(v'\right)\right] dF^U_{V^l_{gt+1}}(v).
\end{aligned} \tag{14}$$

Based on the reasoning leading to equations (13) and (14), 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.

D Appendix: Minimax Optimal Policy

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

$$D^T_{gm} = E\left[D^T_m | X_m, F_m, G_m = g\right],$$

where the superscript T denotes "treated." And similarly, if m is not put on the list, we have

$$D^U_{gm} = E\left[D^U_m | X_m, F_m, G_m = g\right],$$

where the superscript U denotes 'untreated.' (We omit the time dimension t to simplify). When D^j_{gm} 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^j_{gm}. From equation (8) in the main text, the upper bound is

$$\sup_{\gamma \in \Gamma} E_\gamma\left[D^j_m | X_m, F_m, G_m = g\right] = \int \left[\varphi\left(X_m, v\right) \times F_m\right] dF^U_{V^j_g}\left(v\right) \equiv \overline{D}^j_{gm}.$$

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

D.1 The Baseline Case

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

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

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. I.e., $\overline{D}^U_{1m}$=D^U_{1m}.

We now convert this to matrix notation. Let $\mathbf{D}^j_g$ be an $M_g \times 1$ vector with elements D^j_{gm}, 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., $\mathbf{D}^T_0$ is composed of the elements $\overline{D}^T_{0m}$, 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, i.e.,

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

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.

D.2 Incorporating Spillovers in 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 two criteria: (a) they have at leat one neighbor treated and (b) their previous deforestation levels were close to the threshold selection criteria.

Now we have 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 S to denote "spillover" effect).

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 (naturally, $w_{mm} = 0$). 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 it equals 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 $R_m N_m(\phi) = 1$. I.e., when m is untreated, we expect deforestation to be D^S_{gm} when $R_m N_m(\phi) = 1$, and D^U_{gm} when $R_m N_m(\phi) = 0$.

The objective function of the policy maker is now

$$\begin{aligned} SC(\phi) \;=\; & \sum_{m=1}^{M} \phi_m \left[\{G_m = 0\} \overline{D}^T_{0m} + \{G_m = 1\} D^T_{1m} + \{G_m = 2\} \overline{D}^T_{2m} \right] \\ & + (1-\phi_m)(1 - R_m N_m(\phi)) \\ & \times \left[\{G_m = 0\} D^U_{0m} + \{G_m = 1\} \overline{D}^U_{1m} + \{G_m = 2\} \overline{D}^U_{2m} \right] \\ & + (1-\phi_m)(R_m N_m(\phi)) \\ & \times \left[\{G_m = 0\} \overline{D}^S_{0m} + \{G_m = 1\} \overline{D}^S_{1m} + \{G_m = 2\} D^S_{2m} \right]. \end{aligned}$$

Again, when the counterfactual is point-identified, we have $\overline{D}^j_{gm} = D^j_{gm}$. In matrix notation, as before, let $\mathbf{D}^j_g$ be an $M_g \times 1$ vector with elements D^j_{gm}, for $j = U, T, S$, and $g = 0, 1, 2$ (where upper bounds $\overline{D}^j_{gm}$ are put where appropriate). Let $\mathbf{D}^j$ stack the vectors $\mathbf{D}^j_g$ for all g, i.e.,

$$\mathbf{D}^U = \begin{bmatrix} \mathbf{D}^U_0 \\ \mathbf{D}^U_1 \\ \mathbf{D}^U_2 \end{bmatrix}, \mathbf{D}^T = \begin{bmatrix} \mathbf{D}^T_0 \\ \mathbf{D}^T_1 \\ \mathbf{D}^T_2 \end{bmatrix}, \mathbf{D}^S = \begin{bmatrix} \mathbf{D}^S_0 \\ \mathbf{D}^S_1 \\ \mathbf{D}^S_2 \end{bmatrix}.$$

Define also 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 $\circ$ is the Hadamard (i.e., element-by-element) multiplication. Then

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

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).[53]

The genetic algorithm is a stochastic search algorithm that is convenient in the present case because it allows for integer optimization in high-dimensional constrained minimization problems. The procedure requires an initial population matrix, in which each row of the population matrix represents a guess for the optimal list, ϕ. I.e., 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 stochastically selected 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 either a maximum number of generations has been produced, or when the value of the criterion function cannot be further reduced (up to a pre-determined tolerance level).

For each minimization problem considered in the main text, we run the algorithm 20 times. Every 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 $1e-7$. The algorithm always stopped before hitting the maximum number of generations.

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).

[53]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 we iterate ϕ^k until convergence. However, convergence is not guaranteed. Indeed, in our experience, the procedure often ends up in cycles and does not converge to a minima.

E Discussion of Alternative Identification Strategies

Next, we discuss briefly other identification strategies commonly employed in the literature.

Selection-on-Observables. Selection-on-observable techniques require the Conditional Independence Assumption (CIA). In the present case, 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}, where $Z_{mt-1} = \left(Z^1_{mt-1}, Z^2_{mt-1}, Z^3_{mt-1}\right)$, and lagged variables. From (1), it is clear that conditioning on the criteria variables Z_{mt-1} may suffice to satisfy the 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 little overlap in the data between Priority and non-Priority groups given Z_{mt-1}.

Regression Discontinuity Design. 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 the present case, RD suffers from two important limitations. First, there are few observations close to the threshold frontier. This severely limits the accuracy 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, as previously discussed, we are interested in estimating policy treatment effects, which includes treatment effects other than the effects at the cutoff frontier.

Instrumental Variables Approach. Given the triangular system of equations (1) and (6), another possibility would be to exploit instrumental variables. Natural instruments are the criteria variable Z_{mt-1}. 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, as previously mentioned, 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 on 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 needed to investigate optimal targeting), it also requires access to valid instruments.

F Appendix: Robustness Analyses

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

Trimming. Recall that, by putting all probability mass outside the support $Supp\left(D_{1mt+1}\right)$ at the left and at the right end points of $Supp\left(D_{0mt+1}\right)$, we obtain the lower and upper bounds for $F_{D^1_{0t+1}}$ (the same reasoning applies to $F_{D^0_{1t+1}}$). In practice, when calculating the average treatment effects, we follow the literature and trim out 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. In particular, they are robust to trimming out observations below and above the percentiles $[2.5, 97.5]$ and $[3.5, 96.5]$. Table 13 in Appendix G presents the results for the average treatment effects. The top panel shows the estimated ATT, ATU, and ATE when we trimm out 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 13 is comparable to Table 4 in the main text.) The ATTs are unaffected by the trimming, and the estimated identified sets for the ATU and ATE differ slightly across specifications (and all treatment effects are significantly different from zero).

Table 14 in Appendix G presents the implications for the deforestation and carbon emissions of the counterfactual optimal lists. 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 8 in the main text). Again, the results are robust to these different specifications.

Spillover. As explained in the main text, one of the criteria to define whether a municipality belongs to the Spillover group or not depends on whether it has high levels of past deforestation. Formally, we opted for the following condition: $Z^1_{mt-1} \geq 0.7 \times 2,700\ \text{km}^2$ and $Z^2_{mt-1} \geq 0.7 \times 220\ \text{km}^2$. We now show that the results are robust to diffferent definitions of how close past deforestation is to these thresholds. We consider 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 15 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 15 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 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 criteria. In addition, note that 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).[54]

Table 16 in Appendix G presents the implications for the optimal lists (which is comparable to Table 10 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.

[54] 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 11: 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 12: Support of V by Group and Time Period

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

Table 13: 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)

Note: 95% confidence intervals are in parenthesis. For ATT, they were computed based on standard bootstrap. For ATU and ATE, they were based on Imbens and Manski (2004).
Deforestation is measured in squared kilometers.

Table 14: 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

Note: "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 U$, assuming U$ 20/tCO2.

Table 15: 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)

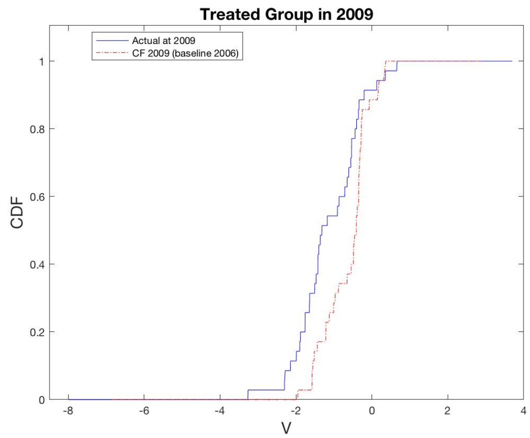
Note: 95% confidence intervals are in parenthesis. For ATT, they were computed based on standard bootstrap. For ATU, ATS, and ATE, they were based on Imbens and Manski (2004). Deforestation is measured in squared kilometers.

Table 16: Robustness: Ex-Post Optimal, Spillovers

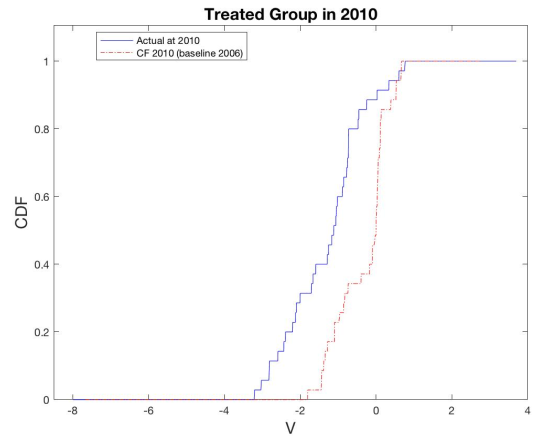
Spillovers: Above 65 Percent of the Threshold						
Constraint:	*Total Area*		*Number of Municipalities*			
	Observed vs Optimal		Observed vs Optimal		Random vs Optimal	
	Ratio	Value	Ratio	Value	Ratio	Value
Total Deforestation						
Baseline 2006	1.08	-	1.07	-	1.29	-
Baseline 2007	1.09	-	1.07	-	1.32	-
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

Note: "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 U$, assuming U$ 20/tCO2.

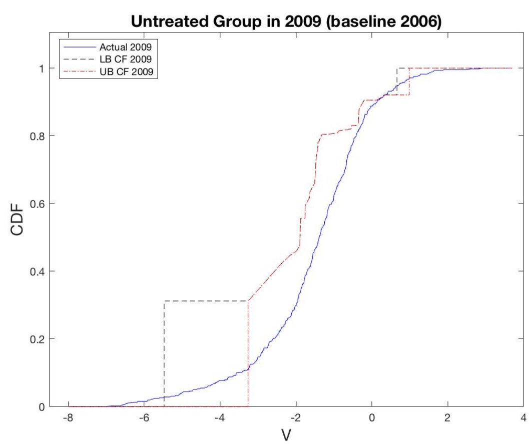


(a) 2009

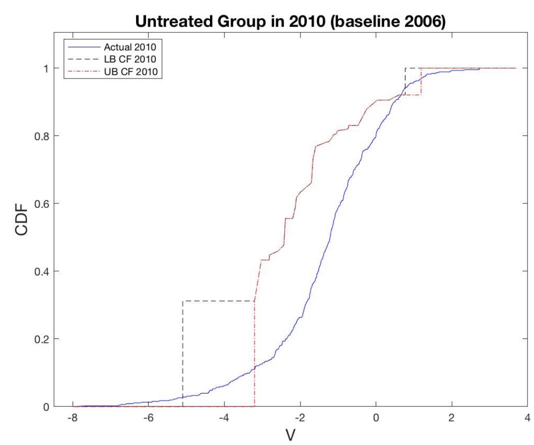


(b) 2010

Figure 12: Factual and Counterfactual Distributions of Residuals V, Treated Group



(a) 2009



(b) 2010

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