Supplemental Material for "Optimal Environmental Targeting in the Amazon Rainforest"

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

March 26, 2022

This Supplemental Material consists of the following sections: Section A provides information about the data sources and the construction of key variables. Section B explains how we inferred the selection rule that assigned municipalities to the Priority List. Section C shows relevant supporting material: estimates from regressions that partial out the covariates, the estimated actual and counterfactual distributions of the residuals, and suggestive evidence of possible channels linking the Priority List to deforestation. Section D presents further discussions of the CIC and DID models and estimated results. Section E shows how we incorporate dynamic treatment effects in our counterfactual calculations. Section F provides a detailed explanation of how the minimax optimal lists are computed in practice, Section G discusses several robustness exercises, and additional tables and figures are shown in Section H.

A Data

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

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.'[1] The data are publicly available at both pixel and municipality levels.

*Contact information: Assunção – Departamento de Economia, PUC-Rio, Rua Marques de Sao Vicente, 225/F210, 22453-900 - Rio de Janeiro/RJ, Brazil (email: juliano@econ.puc-rio.br); McMillan – Department of Economics, University of Toronto, 150 St. George Street, Toronto, ON M5S 3G7, Canada, and NBER (email: mcmillan@chass.utoronto.ca); Murphy – Natural Resources Canada, 580 Booth Street, Ottawa, ON K1A 0E4, Canada (email: joshua.murphy4@canada.ca); Souza-Rodrigues – Department of Economics, University of Toronto, 150 St. George Street, Toronto, ON M5S 3G7, Canada (email: e.souzarodrigues@utoronto.ca).

[1] Each year, a small amount of land area is not directly observed due to cloud cover. In our data, the average share of cloud cover over the municipal area is 2.5 percent (and the median is zero). Deforestation that goes undetected because of cloud cover in one year is attributed to the first subsequent year in which data permit a determination

Since 1998, Brazil's National Institute for Space Research (INPE) has been using images from LANDSAT-class satellites to produce the official statistics used by the government to track deforestation and inform public policy (INPE, 2017). The classification of land cover is performed in several steps. First, because deforestation typically occurs during the dry season, INPE selects LANDSAT images between July and September with minimal cloud coverage (the spatial resolution is 60×60 metres). Second, a linear spectral mixture model for each pixel in the data is estimated in order to obtain the pixel's fraction of different components that help predict land cover.[2] INPE then groups adjacent pixels in larger regions based on their spectral similarities. After that, it implements a cluster unsupervised classification algorithm to generate the land cover classifications. Finally, several photointerpreters verify the results (and reclassify the land cover where necessary, based on specific contexts, historical data, and their judgement). Annual deforestation is calculated taking August 1st of each year as the reference date. (A PRODES year spans the twelve months up to July 31st of each calendar year.) Deforestation is considered irreversible – i.e., once an area has been deforested, it remains classified as deforested in subsequent years.[3]

INPE's classification focuses on detecting deforestation. Yet observed remaining forests have missing observations in some years and do not always decrease monotonically over time (as they should, given that deforestation is considered irreversible). For this reason, we opt to measure 'remaining forest' simply as the remaining available area in the municipality – that is, the total municipal area minus the non-forested areas, water bodies, and previous cumulative deforested areas (2002 is used as the base year). This guarantees consistency over time, the correlation between our proxy and the forested area remaining from PRODES being 0.99. We drop observations with minimal remaining available area (less than 6 km^2). These are small municipalities mostly located at the extreme eastern and southeastern regions of the Amazon Biome, which are not especially relevant for policies intended to prevent deforestation.[4] Finally, in order to calculate the

about land use. The 'unobserved' category is a residual classification reflecting difficulties that affect visibility, such as shadows cast by clouds and smoke from forest fires.

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

[3]There is an important distinction between *incremental deforested area* and the *deforestation rate*. Incremental deforested area measures newly detected primary forest loss, while the deforestation rate adds estimates of cleared forest area that are under unobserved/clouded areas, based on local extrapolation, to that increment. The incremental deforested area is available as spatial data: this is the deforestation measure we use in the empirical analysis. In contrast, the deforestation rate is available only at the aggregate level (and is presented publicly by INPE as the official measurement of annual deforestation). This distinction is needed to reconcile the profiles in Figures 1, H5 and H6 exactly.

[4]They consist of 26 municipalities in total (21 in the state of Maranhão and 5 in the Tocantins state), with an average area approximately 6% of the average area of the municipalities in the final data set. Because of the small values used in the denominator in the calculation of deforestation shares, these small municipalities also exhibit implausibly large oscillations in shares of deforestation over time.

log odds ratio of the shares of deforestation, we assume the minimum amount of deforestation in a municipality in any year is 0.01 km^2.[5]

Priority status. The official list of Priority municipalities (with precise dates for entry and exit) comes from the Ministry of the Environment. Because there are few municipalities entering and exiting the blacklist from 2009 on, not much can be said with a high degree of accuracy about the impact of the policy in those cases. For this reason, we focus on the initial list of Priority municipalities. Specifically, the treatment group consists of municipalities that entered and remained on the list from 2008 to 2010 inclusive. This gives a total of 35 municipalities, as one exited in 2010. The untreated group is the set of municipalities that did not enter the list before 2010.

Protected areas. We calculate the total amount of protected area – whether managed by federal, state, or municipal government – using geo-referenced data from the National Register of Conservation Units, maintained by the Ministry of the Environment. Initiatives to create and expand protected areas were concentrated in the first phase of the PPCDAm (spanning 2004-07), before the first municipalities were assigned to the Priority List in 2008.

Prices. Historically, cattle ranching and crop cultivation have been important drivers of deforestation in the Amazon. To help disentangle the effects of changing commodity prices from the effects of policy interventions on deforestation, we construct beef and crop price indices for each municipality based on pre-determined cross-sectional variation in the crop mix across municipalities and time-series variation in commodity prices. Commodity prices are determined on international markets: local farmers are price takers. Data on prices of the main commodities – beef, soy, rice, corn, cassava, and cane sugar – are taken from the State Secretariat for Agriculture and Food Supply.[6] The five selected crops account for approximately 70 percent of total harvested area in the Brazilian Amazon (averaged across the 2000s). Prices are deflated to 2011 Brazilian reais. Municipality-level data on the amount and value of each form of agricultural production, which we use to weight the commodity prices, come from surveys administered by the Brazilian Institute of Geography and Statistics (IBGE) – the Municipal Crop Survey and the Municipal Livestock Survey. Specifically, for beef cattle, the weight is given by the ratio of head of cattle to municipal area (given that annual pasture area is not observable). For crops, we first calculate a weighted price for each crop by multiplying the commodity prices by the share of the municipal area used to cultivate the crop. For all agricultural products, we fix the weights in the period 2000-01 (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 struc-

[5] This is analogous to setting minimum shares in logit models to be greater than or equal to a small strictly positive number $\epsilon > 0$.

[6] The State Secretariat collects prices received by producers in the southern state of Paraná.

tural break that occurred in 2004-05, associated with the first phase of the PPCDAm. Finally, we apply principal component analysis to the individual weighted crop prices to derive a single index, capturing the price variations common to the five selected crops. The crop index is based on the first principal component; given that the first component maximizes price variance, it provides a more comprehensive measure of the agricultural output price scenario for this analysis than the individual prices (see Assunção et al., 2017; Assunção and Rocha, 2019).

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

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

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

Agricultural Suitability. Data on agricultural suitability come from the Food and Agriculture Organization's Global Agro-Ecological Zones project (FAO-GAEZ). Suitability is measured as the crop maximum attainable yields at the field level (an area of approximately 100 km^2) based on the dominant type of soil, altitude, slope, climatic conditions (among other factors), as well as the level of technology available. We follow Farrokhi and Pellegrina (2021) and consider a high-input technology, corresponding to mechanized production that uses optimum amounts of nutrients and chemical pest, disease and weed control. We select the soy and corn suitability measures (the most important crops in the Brazilian Amazon) and aggregate each of them up to the municipality level. (See Farrokhi and Pellegrina (2021) for a detailed description.)

Distance to Port. Straight-line distances from the municipal seats to the nearest port were calculated using ArcGIS, as explained in Souza-Rodrigues (2019). Data on the location of the municipal seats and ports were produced by the Brazilian Ministry of Transportation for the 2006 National Highway Plan.

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

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

Alerts. Forest clearing alert data come from the Real-Time System for the Detection of Deforestation (DETER), developed and operated by the space agency INPE. DETER makes use of satellite images from MODIS, which has a spatial resolution of 250m^2 (25 hectares), and generates alerts biweekly.[7] The data are publicly available in vector format and are aggregated up to the monthly level. Gandour (2018) has rasterized the georeferenced alerts at a 900m spatial resolution, and constructed panel data in which a cell in the raster data takes on a value of 1 if it contains an alert and a value of 0 otherwise. Gandour has also added the number of alerts per municipality per year (consistently with the PRODES year), generously sharing the aggregated data with us.

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

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

[7] After 2015, INPE upgraded the system in order to detect changes in land cover in patches larger than 1 hectare (instead of areas larger than 25 hectares), albeit less frequently (Gandour, 2018).

B The Selection Rule

In this section, we explain briefly how we inferred the selection rule that assigned municipalities to the Priority List. After investigating the scatterplots presented in Figure 3 of the main text (without the thresholds) informally, we implemented a classification tree algorithm using the criteria variables Z_{mt-1} as the explanatory variables. Specifically, we used the "*crtrees*" command in STATA, with 80% of the original data as the learning sample (split randomly), and the remaining 20% of the observations as the testing sample – all other available options were set at their default levels.

The command *crtrees* implements classification and regression trees that are composed of three sequential algorithms: 'tree-growing,' 'tree-pruning,' and 'the honest tree' (see, e.g., Breiman et al., 1984). The tree-growing algorithm splits terminal nodes of the tree recursively in order to minimize an overall 'impurity' measure (here, the Gini index) using the learning sample. The recursive partitioning stops with the largest possible tree, denoted by T_{MAX}.[8] To avoid overfitting, the tree-pruning algorithm then identifies a sequence of subtrees of T_{MAX} that minimizes a cost-complexity function. The cost-complexity function is given by the weighted sum of the impurity measure and the complexity of the tree (i.e. the number of terminal nodes): the larger the weight on the complexity, the simpler the optimal tree. A sequence of subtrees is obtained by sequentially pruning T_{MAX} based on increasing weights on complexity; pruning is done via a weakest-link algorithm (i.e., an algorithm that collapses subtrees of T_{MAX} into final nodes to minimize the cost-complexity function). The resulting sequence of trees at this stage does not necessarily coincide with the trees in the tree-growing phase. Finally, to reduce further overfitting, the honest tree algorithm chooses from the tree-pruning set of trees the simplest one (i.e., the tree with the smaller number of terminal nodes) among those that minimize the impurity measure in the *testing* sample. The procedure classifies $G = 1$ at a final node if the share of treated observations is larger than the share of untreated at that node.

The final classification tree we obtained from this procedure is presented in Figure H1. It contains five nodes in total, three of which are final nodes. The decision tree is given by the simple threshold rule: $G_{mt} = \{Z^2_{mt-1} \geq 222\} \times \{Z^1_{mt-1} \geq 2,137\}$, where $\{.\}$ denotes the indicator function.[9] This is the selection rule presented in Figure 3 of the main text. This rule correctly classifies 98% of the municipalities; it misclassifies just 8 municipalities out of 490: 6 untreated municipalities (which are above the thresholds) and only 2 treated units (which are below the thresholds).

Overall, the government seems to have overlooked a few municipalities (the six misclassified untreated units) that are home to relevant agricultural activity (mainly, but not just, soybean

[8] It ceases either when all observations in the nodes have the same value or when the number of observations in the nodes is smaller than some pre-established limit. (The default command in STATA sets the lower limit to one.)

[9] We rounded up the first threshold slightly, to 222 – an innocuous adjustment.

production) and did not assign them to the Priority List. At the same time, the government included two misclassified municipalities into the Priority List that have, on average, larger forested areas and less economic pressure (as measured by local GDP and agricultural activities) than the untreated municipalities. In terms of politics, only one of the misclassified municipalities had a mayor that was affiliated with the political coalition of the Brazilian president.[10]

C Regressions, Distributions, and Possible Channels

In this section, we present relevant supporting material: estimates from regressions that partial out the covariates X_{mt}, the estimated distributions of the residuals $F_{V_{gt}^j}$, and suggestive evidence of possible channels linking the Priority List to deforestation.

Partialling-out Regressions. Table H2 reports the estimated coefficients from regressions that project the log odds ratio of deforestation shares onto covariates and group statuses interacted with time dummies, as explained in Section 5 of the main text. The first column in the table ignores spillovers, while the second column incorporates spillover effects. All the coefficients have the expected sign, the exceptions not being statistically significant. For example, the impact of lagged rainfall on deforestation is significant and hump-shaped (reasonable given that both low and high levels of rainfall make agriculture unproductive); the share of protected area is strongly negatively related to the deforestation share; predetermined cropland area (as of 2001) is also negatively related, though not significantly so; and agricultural suitability (as measured by FAO-GAEZ agricultural potential for soy) is positively associated, as one would expect. These covariates have significant explanatory power: together, they increase the R^2 of the partialling-out regression from a low 6.12% (not shown in the table) to 46% when taking spillovers into account.

Distributions of the Residuals. Table H3 presents the estimated supports of the residuals V_{mt} for both the treated and untreated groups in all years 2006-10. It is clear that the supports of the treated group are strictly inside the supports of the untreated group in every year. This leads to the point identification of the ATT and to the partial identification of the ATU under Assumption 3, as explained in Section 5 of the main paper.

The estimated factual and counterfactual distributions of the residuals V_{mt} are presented in Figures H2 and H3. They are similar to Figure 1, page 442, in A&I. The bottom panels (c) and (d) of Figure H2 show the counterfactual distributions of the treated group in the absence of treatment (dotted black lines) in 2009 and 2010, respectively. They are almost everywhere to the

[10]This unique case (Tabaporã, in Mato Grosso state) is an untreated municipality that is above the threshold. While this municipality could have benefited from the political affiliations, that seems unlikely to us because it is not a particularly relevant municipality in the national political arena.

right of the corresponding actual distributions (dashed lines in the lighter shade), indicating that the deforestation shares would have been higher in the absence of the policy – consistent with the stochastic dominance tests presented in Section 6 of the main text. Further, we find larger effects at lower quantiles than in upper quantiles (though not monotonically so). A possible explanation for this pattern is that, conditional on observables, highly deforested locations (i.e., those at the upper quantiles) may be well-suited to agriculture in terms of unobservables so that monitoring would have to be substantial in order to reduce the amount of deforestation; better preserved locations (i.e., at the lower quantiles), in contrast, may be more sensitive to monitoring and enforcement, given the weaker underlying motivation to deforest there.

The bottom panels (c) and (d) of Figure H3 present the lower and upper bounds of the counterfactual distributions of the untreated group had it been treated (dotted black and dotted lighter-shaded lines) in 2009 and 2010, respectively. The reallocations of the probability masses to the endpoints of the supports are clearly visible, given by the discontinuous jumps on the left and right parts of the distributions. (In the interior of the support, the counterfactual cdf is point-identified.) As expected, the counterfactual distributions are to the left of the actual cdfs, indicating lower levels of deforestation among untreated municipalities if they had been put into the Priority List. (Point-identified impacts are larger at lower quantiles than upper quantiles.)

Possible Channels. We now discuss possible channels through which the Priority List may have affected deforestation. Membership of the list brought with it a bundle of provisions (as noted), with farmers in Priority municipalities becoming subject to more rigorous monitoring and law enforcement. They also faced more stringent conditions when seeking to obtain subsidized credit contracts, along with stricter licensing requirements. At the same time, protected areas may have been expanded by the government in a strategic way, taking into account the location of Priority municipalities.

Given this bundle, we are interested in exploring whether particular components appear to have been especially important in cutting deforestation (subject to data limitations). We do so in a suggestive way by estimating how the effects of treatment status – treated versus control – change over time for relevant outcome measures other than deforestation: a proxy for monitoring (the number of alerts given out by INPE), a proxy for enforcement (the number of fines issued by IBAMA), the total volume of rural credit concessions, and the share of protected areas.

Table H4 presents the regressions results, and Figure H4 provides visual evidence showing the evolution of the coefficients on treatment status over time, relative to untreated municipalities in 2006, represented by the horizontal dashed line at zero.[11]

[11] Specifically, we regress the four observable measures – fines, alerts, credit, and protected area share – on Priority status indicators interacted with time dummies, as well as on covariates, taking 2006 as the baseline year. Figure H4

In terms of our monitoring proxy (panel (a) of Figure H4), the average number of alerts issued in untreated municipalities is stable before and after the establishment of the Priority List, relative to the base. This evidence does not suggest any major substitution of monitoring effort away from untreated municipalities once the Priority List came into effect in 2008. In contrast, the number of alerts is far higher for treated municipalities; it oscillates in the pre-treatment period, then declines sharply after 2008, likely reflecting the deforestation slowdown there.

In terms of our enforcement proxy (panel (b)), the number of fines remains reasonably stable in untreated municipalities during the period 2006-10, while it increases substantially among treated municipalities between 2006 and 2008 and then falls back after that, again likely in response to the lower deforestation rates observed there following the policy intervention. In combination, the evidence suggests an increase in enforcement intensity focused on Priority municipalities, rather than a substitution away from untreated municipalities.

There are no clear distinguishing patterns by treatment group status in terms of total rural credit (aside from greater variability across the years among treated municipalities). In terms of shares of protected areas by treatment status, an increase in the share among untreated municipalities relative to 2006 is apparent: among treated municipalities, the changes in share are indistinguishable from zero.

Together, this suggestive evidence helps shed light on whether there was simply a reallocation of fixed resources following the Priority List's introduction or whether state enforcement of environmental regulations increased. The evidence is consistent with more focused targeting being combined with an increase in state capacity to implement environmental regulations, in turn altering municipal-level behavior (reflected in lower deforestation). This discussion complements (and is consistent with) the analysis presented in Assunção and Rocha (2019).

D A Discussion of the CIC and DID Models

In this section, we first expand on the discussions of the rank invariance condition and the identification of the CIC model in the main text. Then we set out the standard DID model and draw attention to reasons why we cannot use it in our optimal targeting analysis – reasons favoring the use of CIC. We then discuss tests of parallel trends, and present various DID estimates, including those from a specification that is made comparable to CIC (using the log odds ratio of deforestation shares and no covariates).

Rank Invariance and Rank Similarity. We note that Assumption 1 in the main text imposes a 'rank invariance' condition. Rank invariance preserves the intuitive notion that, conditional on

presents the coefficients on group status interacted with time dummies.

observables, a relatively highly deforested location in the data remains a relatively highly deforested location under alternative counterfactual policies – specifically, it preserves the rank u. The condition has an undesirable consequence, however, implying that the distribution of the potential outcomes (Y^0_{mt}, Y^1_{mt}), given G, is jointly degenerate, which may be implausible on logical grounds (as pointed out by Heckman et al. (1997)).

In order to avoid this, the CIC framework can be extended to allow for a 'rank similarity' condition, by incorporating two types of unobservables, U^0_{mt} and U^1_{mt}, one for each policy regime, with identical distributions. Doing so relaxes rank invariance by not requiring that the unobservables be identical, $U^0_{mt} = U^1_{mt} \equiv U_{mt}$. Instead, it allows for a common component that makes these unobservables correlated but not perfectly dependent. In this way, rank similarity allows a relatively highly deforested municipality in the data to be *more likely but not guaranteed* to remain a relatively highly deforested municipality under an alternative policy.[12] While noting that all CIC results extend to the (more appealing) rank similarity condition, we maintain Assumption 1 and our original notation (with a single unobservable u) in the main text, for ease of exposition.

We also note here that Assumption 2 in the main text imposes a rank similarity condition – now over time, not across policies. Rank invariance over time requires that a unit m in group G that is at a given rank u (i.e., at a given quantile) in the distribution of U at the time period t should be at the *exact* same rank (quantile) u in the distribution of U at $t+1$; this would render the distribution of (Y^j_{mt}, Y^j_{mt+1}), given G, jointly degenerate. Rank similarity, in contrast, does not require the rank of a unit to be exactly the same over time; it requires the realizations of U_{mt} must be drawn from the same distribution, allowing for serial (but not perfect) correlation. This argument also applies to the distributions of each term U^0_{mt} and U^1_{mt} over time, when allowing for different unobservables under alternative policy regimes.

To summarize, rank similarity, together with the monotonicity assumption, allows us to match treatment and control units properly (as we discuss in the next subsection), before and after treatment, which is needed to construct counterfactual distributions while simultaneously avoiding restricting units to preserve their exact ranks over time and across policy interventions.

Identification. In Section 5, we presented a simple version of the CIC model with two consecutive periods t and $t+1$, before and after treatment. Now we provide an intuition for the identification result, and compare it with the DID identification. To recap, A&I show that under Assumptions 1

[12] Formally, we can assume that $Y^j_{mt} = h^j(X_{mt}, U^j_{mt}, t)$, for $j \in \{0,1\}$, and that U^0_{mt} and U^1_{mt} have *identical* distributions, given X and G. This allows for a common component in the unobservables: $U^0_{mt} = \mu_{mt} + \varepsilon^0_{mt}$ and $U^1_{mt} = \mu_{mt} + \varepsilon^1_{mt}$, where μ_{mt} is the common component, and ε^0_{mt} and ε^1_{mt} are idiosyncratic shocks with identical distributions. Athey and Imbens (2006) note that the CIC model can be extended in this way in their footnote 17 on page 444.

and 2, the counterfactual distribution for the treated group in the absence of treatment is given by

$$F_{Y^0_{1t+1}}(y) = F_{Y_{1t}}\left(F^{-1}_{Y_{0t}}\left(F_{Y_{0t+1}}(y)\right)\right), \tag{D1}$$

where $y \in \mathbb{Y}_{0t+1}$. (A similar expression holds for the untreated group.)

We provide intuition in the context of our study as follows. Intuitively, equation (D1) uses 'double-matching' to construct the counterfactual distribution: a treated municipality that deforested a fraction y of its forested area during period t is first matched to an untreated municipality that deforested the same fraction during the same time period. Then the untreated municipality is matched to its rank counterpart (i.e., in the same quantile) among untreated units in period $t+1$. Let y' denote the fraction deforested by this last unit during $t+1$, and define the difference $\Delta \equiv y' - y$. The difference between the shares of deforestation of the treated unit during t and during $t+1$ *in the absence of treatment* is then given by the difference between the deforestation shares of the untreated units *with the same rank* before and after treatment. That is, the counterfactual share of deforestation of the treated quantile in the absence of treatment is given by $y + \Delta$.[13] (See Figure 1, page 442, in A&I for a clear exposition of their 'double-matching' procedure.)

This is similar to the adjustment in the standard DID model, but in the DID case, the adjustment is linear and to the mean (under the parallel trends assumption), given by:

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

Importantly, because the CIC and DID models construct counterfactual outcomes in different ways, there is no *a priori* reason to expect the CIC and DID estimators will generate similar point estimates for the average treatment effects.

The Logistic DID Model. As a point of reference, we adopt a logistic regression framework for our DID model as it is common in the empirical land use literature (see, e.g., Stavins, 1999; Pfaff, 1999; Souza-Rodrigues, 2019) and is the closest specification to our semiparametric logit CIC model; see Section 5 of the main paper. (Recall that a linear model applied to deforestation levels predicts negative deforestation for a non-negligible fraction of the municipalities; this may lead to biased ATT estimates and produce misleading results when constructing the counterfactual optimal list.) In the standard DID model, the regression formulation is given by:

$$\log\left(\frac{Y_{mt}}{1-Y_{mt}}\right) = X'_{mt}\beta + \delta_t + \tau_1\left(G_m \times \delta_{2009}\right) + \tau_2\left(G_m \times \delta_{2010}\right) + \alpha_m + \eta_{mt}\ , \tag{D2}$$

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

where X_{mt} is a municipality-level vector of observed factors; δ_t are time dummies; α_m is a municipality-level fixed effect; η_{mt} is a time-varying unobservable factor; and (β, τ_1, τ_2) are the parameters to be estimated. The parameters τ_1 and τ_2 equal the average treatment effect (in terms of the log odds ratio of deforestation shares) among the treated municipalities during the first and the second years of the program, respectively, thus allowing for time-varying treatment effects. Given that the Priority List should reduce deforestation, one would expect $\tau_1 \leq 0$ and $\tau_2 \leq 0$. The parameters can be estimated consistently based on (D2) using a fixed effect estimator, provided that the parallel trends assumption holds – we provide formal econometric evidence below.

The main advantage of the fixed effects estimator in general is that, by eliminating the fixed effects term α_m from the estimated regression, it allows for potential correlations between α_m and the time-varying regressors, with no additional assumption on the distribution of the fixed effects (nor of the idiosyncratic shocks η_{mt}).

There are, however, two difficulties that prevent us from using the DID model to calculate the counterfactual optimal list – the main goal of this paper. First, converting the estimated results for the log odds ratio into average treatment effects for deforestation levels is not possible without assumptions on the distribution of unobservables – a common limitation in nonlinear panel data models. Thus, estimating the parameters in (D2) is not sufficient for our goals. Second, the DID model does not provide a framework for estimating average treatment effects on the untreated more generally (not just for nonlinear models).

In contrast, our semiparametric logit CIC model allows us to estimate deforestation ATTs under Assumptions 1 - 4 presented in the main paper. In order to mitigate concerns about possible correlations between time-varying observables and time-invariant factors affecting deforestation in the CIC model, we include several time-invariant covariates in the vector X_{mt}, such as measurements of soil quality, distance to ports, and state dummies (described in the main text). Further, while the DID model restricts group and time effects to be constant in the logistic specification, CIC allows for extra flexibility and, therefore, richer heterogeneous effects. We also note that the fixed effect estimator applied to (D2) requires covariates to be strictly exogenous for consistency. The CIC partialling-out regression, in contrast, only requires contemporaneous exogeneity; we include lagged variables when partialling covariates out to make this exogeneity condition more likely to be satisfied in the data.

Parallel Trends Test. Before presenting the main DID estimates, we assess whether the trends in the outcome variables are parallel in the pre-treatment period. First, we do so graphically. Panel (a) of Figure H5 compares aggregate trends in deforestation for treated and untreated municipalities. While differences in deforestation levels are apparent, new deforestation fell after 2005 for both groups, and increased slightly in 2006-08. (The magnitudes are more pronounced among treated municipalities possibly because they are larger in size, on average.) In panel (b), we compare the evolution across the two groups in the log odds ratio of deforestation shares – the outcome variable in our empirical framework. The same pattern emerges, with level differences before 2008 but similar movements for the two groups. There are no signs of any anticipation effects.

Figure H6 is similar to Figure H5 but splits the untreated group into the spillover and control groups. The deforestation levels for the spillover group are found between the other two groups, while the log odds ratios are slightly above those of the treated group. Again, the evolution profiles are similar, especially after 2005. Of note, while deforestation slowed down in the three groups after 2008, the slowdown among the spillover municipalities is not as pronounced as for Priority units, but is more prominent than for control units.

Next we turn to formal testing.[14] Table H5 presents the results, the first two columns ignoring potential spillover effects, the last two columns incorporating them. In columns (1) and (3), we do not include covariates, while columns (2) and (4) include them. The table indicates that the parallel trends assumption holds before treatment: the coefficients on the time dummies interacted with Priority and spillover status are not statistically significant before 2008 (aside from one interaction in the fourth column).

This evidence accords with the discussion in Section 4, where we noted that the government's actual assignment of municipalities to the Priority List was consistent with a rule that selected municipalities based primarily on the level (rather than the trend) of deforestation.

Estimated Results. Table H6 presents the coefficients from estimating variants of the DID regression model specified in equation (D2). The first column does not include covariates, while the second column does. In the third and fourth columns, we incorporate potential spillover effects in the estimation strategy, splitting the untreated municipalities into 'control' and 'spillover' groups; column (3) does not include covariates, while column (4) does.

In all specifications in the table, the Priority List appears to have reduced deforestation substantially. First, when ignoring spillovers, the coefficients on Priority status after treatment are statistically significant, and show an average reduction in the odds ratio of the deforestation share of 46% in 2009 and approximately 90% in 2010. The impacts are robust to the inclusion (or exclusion) of the covariates, with the greater impact in 2010 consistent with farmers updating their beliefs about the new policy regime. The pattern of increasing effects over time is similar to that in the CIC model.

As mentioned previously, there is no *a priori* reason to expect the DID and CIC estimators will generate similar point estimates, given that the DID and CIC models construct counterfactual out-

[14]In the main text, we present an analogous test for the CIC model – a placebo test in which we assess whether the actual distribution of deforestation shares equals the counterfactual distribution if the policy intervention were to be imposed (falsely) one year early.

comes in different ways. Nevertheless, we find estimated effects from using CIC that are somewhat smaller when computed on a comparable basis. For instance, applying the CIC estimator to the log odds ratio of deforestation shares when not conditioning on covariates implies a reduction in the odds ratio of 44% in 2009 and 69% in 2010, on average – smaller than the corresponding DID estimates, given by the coefficients on Priority status presented in column (1) of Table H6.

Next, when potential spillover effects are taken into account, the coefficients on Priority status presented in columns (3) and (4) are slightly greater than the corresponding estimates ignoring potential spillovers. (The increase is similar to that in the CIC model.) The average impact on the odds ratio is a 49% reduction in 2009 and around 96% in 2010. Overall, the estimated impacts of the Priority List on deforestation are economically and statistically significant, and robust across all specifications.

Columns (3) and (4) of the table also shed light on the estimated *spillover* effects of the Priority List following treatment. We find these effects to be negative and statistically significant: there is an average reduction in the odds ratio by approximately 38% in 2009 and by 77% in 2010. Again, the impacts are robust to whether or not the covariates are included. The evidence indicates that untreated municipalities with a treated neighbor and high levels of past deforestation reduce their deforestation rates in response to the establishment of the Priority List (perhaps following belief-updating by farmers). As expected, the (indirect) spillover effects are smaller in magnitudes than the (direct) impacts of the policy on Priority municipalities, a pattern consistent with the estimated CIC model.

E Dynamic Treatment Effects

In this section, we describe how dynamic treatment effects are incorporated into the analysis. Doing so is important when calculating counterfactual deforestation, given the evolution of the remaining forested area depends on deforestation in prior periods.

For simplicity, consider three consecutive periods: t, $t+1$, and $t+2$, where t refers to the time period before the treatment, and $t+1$ and $t+2$ refer to the first and second time periods after treatment. Here, we focus on the second period, $t+2$.

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

at $t+2$ is

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

By the Law of Iterated Expectations,

$$\begin{aligned}
& E\left[D^j_{mt+2}|X_{mt+2}, X_{mt+1}, A_{mt+1}, G_m=g\right] \\
= \;& \int\left[\int\left[\varphi\left(X_{mt+2}, v'\right)\left(1-\varphi\left(X_{mt+1}, v\right)\right) A_{mt+1}\right] dF_{V^j_{gt+2}}\left(v'\right)\right] dF_{V^l_{gt+1}}(v).
\end{aligned} \tag{E3}$$

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

$$\begin{aligned}
& \int\left[\int\left[\varphi\left(X_{mt+2}, v'\right)\left(1-\varphi\left(X_{mt+1}, v\right)\right) A_{mt+1}\right] dF^L_{V^j_{gt+2}}\left(v'\right)\right] dF^L_{V^l_{gt+1}}(v) \\
\leq \;& E\left[D^j_{mt+2}|X_{mt+2}, X_{mt+1}, A_{mt+1}, G_m=g\right] \\
\leq \;& \int\left[\int\left[\varphi\left(X_{mt+2}, v'\right)\left(1-\varphi\left(X_{mt+1}, v\right)\right) A_{mt+1}\right] dF^U_{V^j_{gt+2}}\left(v'\right)\right] dF^U_{V^l_{gt+1}}(v).
\end{aligned} \tag{E4}$$

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

F Minimax Optimal Policy

In this section, we explain in more detail how the ex-post optimal lists are calculated in practice.

First, recall that we do not select a list that changes over time, as explained in the main text, because that would complicate the problem substantially. Instead, the optimal list is based on the sum of deforestation in the two years after the treatment. To simplify, let $t+1$ and $t+2$ refer to 2009 and 2010, respectively.

Denote the counterfactual assignment rule by $\phi=(\phi_1,...\phi_M)'$. This assigns the treatment to municipalities $m=1,...,M$ and can be either deterministic $\phi_m\in\{0,1\}$ or probabilistic $\phi_m\in[0,1]$. Define the sum of expected deforestation in years 2009 and 2010 of municipality m in group $G_m=g$ if it is placed on the blacklist ($\phi_m=1$) by

$$\mathcal{D}^T_{gm}\equiv E\left[D^T_{mt+1}|X_{mt+1}, A_{mt+1}, G_m=g\right]+E\left[D^T_{mt+2}|X_{mt+2}, X_{mt+1}, A_{mt+1}, G_m=g\right],$$

where the superscript T denotes ‘treated’ (note the slight change in notation to help exposition).

The first term on the right-hand side is the expected deforestation in 2009, calculated based on equation (5) in the main text, and the second term on the right-hand side is the expected deforestation in the following year, taking into account the (counterfactual) remaining forested area from the previous year, given by equation (E3) in Section E of this supplement.

Clearly, $\mathcal{D}_{gm}^{T} \in \mathbb{D}_{gm}^{T} \equiv [\underline{\mathcal{D}}_{gm}^{T}, \overline{\mathcal{D}}_{gm}^{T}]$, where $\underline{\mathcal{D}}_{gm}^{T}$ is the sum of the lower bound for the expected deforestation in 2009, given by the left-hand side of (6) in the main text, and the lower bound for the expected deforestation in 2010, calculated based on the left-hand-side of (E4) in Section E of this supplement; and $\overline{\mathcal{D}}_{gm}^{T}$ is the sum of the upper bounds for the expected deforestations in 2009 and 2010, defined similarly by the right-hand sides of (6) and (E4), respectively. Naturally, when the expected deforestation is point-identified, we obtain $\mathcal{D}_{gm}^{T} = \underline{\mathcal{D}}_{gm}^{T} = \overline{\mathcal{D}}_{gm}^{T}$.

If m is not placed on the list ($\phi_m = 0$), we have

$$\mathcal{D}_{gm}^{U} \equiv E\left[D_{mt+1}^{U}|X_{mt+1}, A_{mt+1}, G_m = g\right] + E\left[D_{mt+2}^{U}|X_{mt+2}, X_{mt+1}, A_{mt+1}, G_m = g\right],$$

where the superscript U denotes 'untreated.' As before, we have that $\mathcal{D}_{gm}^{U} \in \mathbb{D}_{gm}^{U} \equiv [\underline{\mathcal{D}}_{gm}^{U}, \overline{\mathcal{D}}_{gm}^{U}]$, where $\underline{\mathcal{D}}_{gm}^{U}$ and $\overline{\mathcal{D}}_{gm}^{U}$ are the lower and upper bounds for the untreated unit, summing over expected deforestation in 2009 and 2010. Again, under point-identification, we have $\mathcal{D}_{gm}^{U} = \underline{\mathcal{D}}_{gm}^{U} = \overline{\mathcal{D}}_{gm}^{U}$.

The expected levels of deforestation, $\mathcal{D}_{gm}^{j}$, for $j \in \{U, T\}$ (and their bounds, $\underline{\mathcal{D}}_{gm}^{j}$ and $\overline{\mathcal{D}}_{gm}^{j}$) are estimated in the data using the CIC model.

F.1 The Baseline Case

When spillover effects are not considered, the policy maker solves the following minimax problem:

$$\min_{\phi \in [0,1]^M} \max_{\mathcal{D} \in \mathbb{D}} \sum_{m=1}^{M} \phi_m \left[\{G_m = 1\} \mathcal{D}_{1m}^{T} + \{G_m = 0\} \mathcal{D}_{0m}^{T}\right] + (1 - \phi_m) \left[\{G_m = 1\} \mathcal{D}_{1m}^{U} + \{G_m = 0\} \mathcal{D}_{0m}^{U}\right],$$

where $\mathcal{D} = (\mathcal{D}_1, ..., \mathcal{D}_M)$, with $\mathcal{D}_m = (\mathcal{D}_{0m}^{U}, \mathcal{D}_{1m}^{U}, \mathcal{D}_{0m}^{T}, \mathcal{D}_{1m}^{T})$; the product set is $\mathbb{D} = \prod_{m=1,...,M} \mathbb{D}_m$, with $\mathbb{D}_m = \mathbb{D}_{0m}^{U} \times \mathbb{D}_{1m}^{U} \times \mathbb{D}_{0m}^{T} \times \mathbb{D}_{1m}^{T}$; and $\{.\}$ denotes the indicator function. Note that when municipality m does not belong to group g, i.e., when $G_m \neq g$, we have that $\{G_m = g\} \mathcal{D}_{gm}^{j} = 0$. We can therefore set $\mathcal{D}_{gm}^{j} = 0$ and $\mathbb{D}_{gm}^{j} = \{0\}$ for such cases without loss of generality.

By noting that the inner maximization in the minimax problem above is given by

$$\begin{aligned}
&\max_{\mathcal{D}\in\mathbb{D}} \sum_{m=1}^{M} \phi_m \left[\{G_m=1\}\, \mathcal{D}^T_{1m} + \{G_m=0\}\, \mathcal{D}^T_{0m}\right] + (1-\phi_m)\left[\{G_m=1\}\, \mathcal{D}^U_{1m} + \{G_m=0\}\, \mathcal{D}^U_{0m}\right] \\
= &\sum_{m=1}^{M} \max_{\mathcal{D}_m\in\mathbb{D}_m} \left[\phi_m \left[\{G_m=1\}\, \mathcal{D}^T_{1m} + \{G_m=0\}\, \mathcal{D}^T_{0m}\right] + (1-\phi_m)\left[\{G_m=1\}\, \mathcal{D}^U_{1m} + \{G_m=0\}\, \mathcal{D}^U_{0m}\right]\right] \\
= &\sum_{m=1}^{M} \phi_m \left[\{G_m=1\}\, \overline{\mathcal{D}}^T_{1m} + \{G_m=0\}\, \overline{\mathcal{D}}^T_{0m}\right] + (1-\phi_m)\left[\{G_m=1\}\, \overline{\mathcal{D}}^U_{1m} + \{G_m=0\}\, \overline{\mathcal{D}}^U_{0m}\right], \qquad \text{(F5)}
\end{aligned}$$

we define the right-hand side of (F5) as the 'social cost' function, denoted by $SC(\phi)$.[15] So the policy maker's problem under the minimax criterion simplifies to

$$\min_{\phi_t\in[0,1]^M} SC(\phi), \qquad \text{(F6)}$$

It is convenient to convert this to matrix notation. Let $\mathbf{D}^j_g$ be an $M_g \times 1$ vector with elements $\overline{\mathcal{D}}^j_{gm}$, for $j \in \{T, U\}$ and $g \in \{0,1\}$, where M_g is the number of municipalities in group g. (Recall that $\overline{\mathcal{D}}^j_{gm} = \mathcal{D}^j_{gm}$ when expected deforestation is point-identified.) Let $\mathbf{D}^j$ stack the vectors $\mathbf{D}^j_g$ for all g, so

$$\mathbf{D}^T = \begin{bmatrix} \mathbf{D}^T_0 \\ \mathbf{D}^T_1 \end{bmatrix}, \mathbf{D}^U = \begin{bmatrix} \mathbf{D}^U_0 \\ \mathbf{D}^U_1 \end{bmatrix}.$$

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 involves a straightforward linear programming problem.

F.2 Incorporating Spillovers into the Optimal List

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

There are three possibilities to consider when selecting a municipality to the optimal list: If a municipality m is placed on the blacklist ($\phi_m = 1$), the expected deforestation is given by $\mathcal{D}^T_{gm}$, as before. If it is not placed on the list ($\phi_m = 0$), there are two possible levels of deforestation: (i) if

[15] That is because the maximum of a finite sum of positive bounded numbers equals the sum of the maximum values that these numbers can take.

m has either no neighbor treated or is 'far' from the threshold criteria, expected deforestation is $\mathcal{D}_{gm}^{U}$; (ii) if it has at least one neighbor treated and is 'close' to the threshold criteria, deforestation is $\mathcal{D}_{gm}^{S}$ (where we use superscript S to denote 'spillover'). Note that the same criteria used to define the spillovers group in the data is also used when determining the optimal blacklist. I.e., we assume that spillovers operate in the counterfactual scenario in the same manner they operate in the observed data. All terms $(\mathcal{D}_{gm}^{T}, \mathcal{D}_{gm}^{U}, \mathcal{D}_{gm}^{S})$, and their corresponding upper bounds, are given by the sum of the expected deforestations in 2009 and 2010, and are calculated as explained previously.

To be specific, in order to incorporate spillovers in the social cost function, $SC(\phi)$, we first consider the 'treated neighbor' component of the criteria. The adjacency matrix indicating whether municipality m and n are neighbors is given by

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

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

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

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

Next, consider the second criterion: whether past deforestation of m is close to the threshold rule or not. Denote this by the indicator variable $R_m \in \{0,1\}$. The two criteria are satisfied only when $R_m N_m(\phi) = 1$. Specifically, if m is not placed on the blacklist ($\phi_m = 0$), we expect deforestation to be $\mathcal{D}_{gm}^{U}$ when $R_m N_m(\phi) = 0$, and $\mathcal{D}_{gm}^{S}$ when $R_m N_m(\phi) = 1$.

Following the reasoning presented in the previous subsection, the policy maker's social cost function (corresponding to the inner maximization in the minimax problem) is now given by

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

As before, let $\mathbf{D}_g^j$ be an $M_g \times 1$ vector with elements $\overline{\mathcal{D}}_{gm}^{j}$, for $j \in \{U, T, S\}$ and $g \in \{0, 1, 2\}$. Let $\mathbf{D}^j$ stack the vectors $\mathbf{D}_g^j$ for all g, giving

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

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

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

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

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

where $Diag\left(\mathbf{D}^S - \mathbf{D}^U\right)$ is the diagonal matrix with the elements of the vector $\mathbf{D}^S - \mathbf{D}^U$ in the diagonal.

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 the genetic algorithm to find the global minimum (Deep et al., 2009).

The genetic algorithm is a stochastic search algorithm – convenient in the current context because it allows for integer optimization in high-dimensional constrained minimization problems. The procedure requires an initial population matrix, in which each row represents a guess for the optimal list, ϕ – that is, each row is composed of elements taking values that are either zeros or ones specifying which of the $M = 490$ municipalities are to be included on the optimal list, subject to the constraint in question (either total number of municipalities or total municipality area).

In each step, the objective function is evaluated for each 'individual' (vector) in the population matrix, and the most 'promising' individuals (in terms of minimizing the criterion function) are selected stochastically from the population. The selected vectors are then modified – recombined and possibly randomly mutated – to form a new generation of candidate solutions. The new generation is then used in the next iteration of the procedure. The algorithm stops when the value of the criterion function cannot be further reduced (up to a pre-determined tolerance level), or when a maximum number of generations has been reached.

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

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

G Robustness Analyses

In this section, we investigate the robustness of our main results to (i) the choice of baseline year, (ii) the way observations were trimmed to reduce the impact of outliers on the estimated treatment effects, and (iii) the definition of the spillover group.

Alternative Baseline Years. Here, we consider the use of each alternative baseline year, 2006 and 2007. Table H7 shows the results for the average treatment effects. For each post-treatment year, we present the estimates based on (a) the baseline year 2006, (b) the main specification (averaging across the two baseline years), and (c) the baseline year 2007. (Table H7 is comparable to Table 2 in the main text.) The results are similar across the baseline years. The estimated ATTs are between -21 km^2 and -28 km^2 in 2009, and increase to between -53 km^2 and -63 km^2 in 2010. The estimated bounds for the ATU and ATE are similar in magnitude across the baselines years as well. (All treatment effects are significantly different from zero.)

Table H8 presents the implications of the counterfactual optimal lists for deforestation and carbon emissions (which are comparable to Table 6 in the main text). As before, we show results for each baseline year and the main specification. Again, the results are robust to the choice of baseline year.

Trimming. Recall that by placing all probability mass outside the support $\mathbb{Y}_{1t+1}$ at the left and right end points of $\mathbb{Y}_{0t+1}$, we obtain the lower and upper bounds for $F_{Y^1_{0t+1}}$; the same reasoning applies to $F_{Y^0_{1t+1}}$. In practice, we follow the literature and trim observations below the 3rd and above the 97th percentiles to minimize the influence of outliers (Ginther, 2000; Lee, 2009). We now show that the empirical results are robust to such trimming – specifically, trimming observations below and above the percentiles $[2.5, 97.5]$ and $[3.5, 96.5]$. Table H9 presents the results for the average treatment effects. The top panel shows the estimated ATT, ATU, and ATE when we trim the observations below the 2.5th and above the 97.5th percentiles, while the bottom panel presents the results when the 3.5th and 96.5th percentiles are used. (Table H9 is comparable to Table 2 in the main text.) The ATTs are unaffected by the trimming, and the estimated identified sets for the ATU and ATE differ only slightly across specifications (and all treatment effects are significantly different from zero).

Table H10 presents the implications for 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 (comparable to Table 6 in the main text). Once again, the results are robust across these different specifications.

Spillovers. As explained in the main text, one of the criteria used to define whether a municipality belongs to the spillover group or not concerns whether it has high levels of past deforestation. Formally, we imposed the following condition: $Z^1_{mt-1} \geq 0.7 \times 2,137$ km^2 and $Z^2_{mt-1} \geq 0.7 \times 222$ km^2. We now show that the results are robust to different definitions of closeness of past deforestation to these thresholds, considering Z^1_{mt-1} and Z^2_{mt-1} greater than 65 percent and 75 percent of the threshold criteria.

The top panel of Table H11 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 H11 is comparable to Table 3 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 it is to the threshold criteria, and the more likely it is that farmers there may react to the policy intervention. So, when the spillover group is composed of municipalities with lower levels of past deforestation (the 65 percent group definition), we expect the treatment effects to be smaller than when the group is composed of municipalities with higher levels of past deforestation. Still, the estimated magnitudes of the ATSs are similar across the different threshold criteria.

Table H12 presents the implications for the optimal lists (comparable to Table 6 in the main text). Again, the results are robust to alternative definitions, this time of the spillover group.

H Additional Tables and Figures

Table H1: 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 H2: Partialling-Out Regression

	(1) Log odds	(2) Log odds
Control Group x Year=2006	-1.067	-1.303
	(0.179)	(0.187)
Control Group x Year=2007	-0.946	-1.177
	(0.176)	(0.184)
Control Group x Year=2008	-0.507	-0.706
	(0.174)	(0.182)
Control Group x Year=2009	-1.436	-1.623
	(0.180)	(0.189)
Control Group x Year=2010	-1.213	-1.376
	(0.179)	(0.182)
Treated Group x Year=2006	-0.121	-0.119
	(0.112)	(0.112)
Treated Group x Year=2007	-0.0126	-0.00287
	(0.136)	(0.137)
Treated Group x Year=2009	-1.127	-1.126
	(0.163)	(0.163)
Treated Group x Year=2010	-1.297	-1.294
	(0.180)	(0.180)
Spillovers Group x Year=2006		-0.216
		(0.181)
Spillovers Group x Year=2007		-0.116
		(0.212)
Spillovers Group x Year=2008		-0.205
		(0.192)
Spillovers Group x Year=2009		-1.235
		(0.203)
Spillovers Group x Year=2010		-1.283
		(0.304)
Lagged Rainfall	0.127	0.112
	(0.0543)	(0.0546)
Lagged Rainfall Squared	-0.00388	-0.00360
	(0.00112)	(0.00112)
Lagged Temperature	0.0479	0.0472
	(0.0464)	(0.0433)
Share of Protected Areas	-2.412	-2.369
	(0.238)	(0.238)
Price of Beef Lagged	0.00278	0.00422
	(0.00227)	(0.00231)
Price of Crops Lagged	0.263	0.289
	(0.328)	(0.319)
Lagged GDP	-0.00331	-0.00459
	(0.0520)	(0.0521)
Crop Area by 2001	-0.000339	-0.000339
	(0.000249)	(0.000226)
Cattle Heads by 2001	0.000964	0.000528
	(0.000436)	(0.000458)
FAO-GAEZ Corn	-0.00661	-0.00281
	(0.00611)	(0.00629)
FAO-GAEZ Soy	0.0565	0.0510
	(0.0244)	(0.0242)
Distance to Port	-0.0000998	-0.000144
	(0.000238)	(0.000235)
R^2	0.454	0.460
Observations	2445	2445

Notes: An observation is a municipality in the Brazilian Amazon. The dependent variable is the log odds ratio of deforestation shares. Rainfall is measured in millimeters (mm) and temperature is measured in degrees Celsius ($^{\circ}C$). Price of beef is a weighted average of international beef prices weighted by the ratio of head of cattle to municipal area. Price of crops is the price index based on a principal component analysis applied to individual weighted prices of the most predominant crops in the Brazilian Amazon (the weights are given by the share of the municipal area used to cultivate the crop). For all agricultural products, the weights are fixed in the period 2000–2001. Municipal GDP is measured in million Reais. All monetary amounts are expressed in December 2011 Reais. Crop area in 2001 is measured in km^2. Cattle heads in 2001 is measrured in thousands. FAO-GAEZ Soy and FAO-GAEZ Corn consist of crops maximum attainable yields at the field level, aggregated up to the municipality level. Distance to port is measured in kilometers. The coefficients on the state dummies and on the constant term are omitted. Robust standard errors in parentheses are clustered at the municipality level.

Table H3: Support of Residuals V_{mt}, by Group and Time Period

	Support of V_{jt}				
	2006	2007	2008	2009	2010
Untreated Group	[-7.733, 3.864]	[-6.555, 3.584]	[-6.590, 3.569]	[-6.958, 2.982]	[-7.867, 2.833]
Treatment Group	[-1.668, 1.205]	[-1.810, 1.324]	[-1.784, 1.527]	[-3.326, 0.956]	[-3.441, 1.001]

Note: Authors' calculation

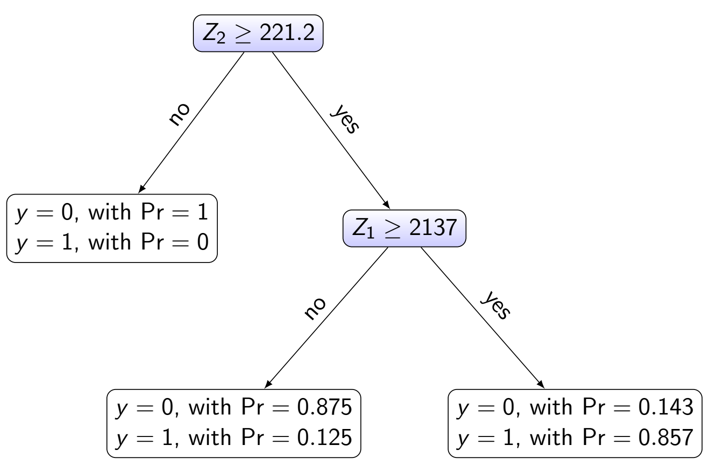


Figure H1: Selection Criterion: Classification Tree

Table H4: Regressions on Other Outcomes: Alerts, Fines, Rural Credit, and Protected Areas

	(1) Alerts	(2) Fines	(3) Total Credit	(4) PA Share
Untreated Group x Year=2007	-20.15*	1.247	-1227.7*	0.0319
	(8.388)	(1.413)	(487.2)	(0.0164)
Untreated Group x Year=2008	24.84**	5.344*	-194.6	0.0627***
	(8.561)	(2.179)	(444.7)	(0.0149)
Untreated Group x Year=2009	-0.525	3.466	-927.5	0.0593***
	(9.592)	(2.440)	(516.8)	(0.0173)
Untreated Group x Year=2010	-14.51	2.026	1201.2*	0.0545***
	(9.039)	(2.104)	(538.3)	(0.0161)
Treated Group x Year=2006	752.2***	14.99*	1476.4	-0.0413
	(118.8)	(5.964)	(2851.4)	(0.0355)
Treated Group x Year=2007	321.6***	24.28***	-922.4	-0.0195
	(70.89)	(7.025)	(2034.2)	(0.0384)
Treated Group x Year=2008	670.7***	48.04***	6614.4	0.0225
	(86.39)	(12.02)	(3374.5)	(0.0385)
Treated Group x Year=2009	342.0***	33.71***	-3426.3	0.0210
	(88.18)	(8.196)	(2626.3)	(0.0389)
Treated Group x Year=2010	125.4*	17.37**	3716.1	0.0115
	(60.82)	(5.630)	(4168.2)	(0.0401)
Lagged Rainfall	10.70**	-0.404	-991.2***	-0.000302
	(3.547)	(0.959)	(258.1)	(0.0115)
Lagged Rainfall Squared	-0.252***	-0.00540	21.21***	-0.0000217
	(0.0681)	(0.0166)	(5.233)	(0.000235)
Lagged Temperature	12.33	1.418	-144.3	0.00810
	(8.159)	(1.205)	(364.0)	(0.0112)
Price of Beef Lagged	-1.684***	-0.222**	-41.38**	-0.00463***
	(0.431)	(0.0730)	(13.50)	(0.000390)
Price of Crops Lagged	32.72	8.829	9131.3**	-0.271***
	(33.50)	(6.516)	(3141.6)	(0.0621)
Lagged GDP	5.793	4.672*	689.8*	0.00156
	(6.174)	(2.185)	(288.7)	(0.0134)
Crop Area by 2001	-0.0223	-0.00852	32.88***	0.0000914
	(0.0299)	(0.00469)	(2.142)	(0.0000497)
Cattle Heads by 2001	0.770	0.0419**	49.74***	0.0000528
	(0.397)	(0.0135)	(6.780)	(0.000110)
FAO-GAEZ Corn	-0.882	-0.244*	-20.97	0.000173
	(0.823)	(0.116)	(42.97)	(0.00106)
FAO-GAEZ Soy	2.502	-0.0393	205.7	-0.00123
	(3.507)	(0.376)	(161.0)	(0.00533)
Distance to Port	0.0800*	0.00602	1.591	0.000160**
	(0.0332)	(0.00416)	(1.650)	(0.0000604)
R^2	0.516	0.270	0.762	0.370
Observations	2445	2445	2320	2445

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

Table H5: Pre-Treatment 'Parallel Trends' Test, 2003–2007

	(1)	(2)	(3)	(4)
	Log odds	Log odds	Log odds	Log odds
Treated Group x Year=2003	-0.286	-0.286	-0.300	-0.305
	(0.184)	(0.179)	(0.188)	(0.182)
Treated Group x Year=2004	0.127	0.205	0.145	0.227
	(0.161)	(0.165)	(0.163)	(0.167)
Treated Group x Year=2005	0.249	0.288	0.272	0.314
	(0.153)	(0.157)	(0.157)	(0.161)
Treated Group x Year=2006	-0.147	-0.107	-0.158	-0.115
	(0.137)	(0.137)	(0.141)	(0.141)
Spillover Group x Year=2003			-0.195	-0.197
			(0.210)	(0.202)
Spillover Group x Year=2004			0.251	0.308
			(0.168)	(0.167)
Spillover Group x Year=2005			0.305	0.339*
			(0.160)	(0.163)
Spillover Group x Year=2006			-0.144	-0.0989
			(0.133)	(0.134)
Year=2003	1.132***	1.127***	1.147***	1.146***
	(0.0974)	(0.100)	(0.104)	(0.107)
Year=2004	0.897***	0.830***	0.878***	0.813***
	(0.0655)	(0.0991)	(0.0697)	(0.103)
Year=2005	0.689***	0.583***	0.666***	0.558***
	(0.0878)	(0.0966)	(0.0942)	(0.103)
Year=2006	0.108	0.108	0.119	0.125
	(0.0837)	(0.0955)	(0.0901)	(0.102)
Lagged Rainfall		0.0581		0.0633
		(0.0414)		(0.0417)
Lagged Rainfall Squared		-0.00166		-0.00174
		(0.000896)		(0.000903)
Lagged Temperature		-0.376*		-0.388*
		(0.158)		(0.158)
Share of Protected Areas		0.838**		0.838**
		(0.300)		(0.298)
Price of Beef Lagged		-0.00107		-0.00104
		(0.00830)		(0.00831)
Price of Crops Lagged		0.886***		0.884***
		(0.183)		(0.179)
Lagged GDP		-0.356**		-0.364**
		(0.132)		(0.131)
R^2	0.135	0.147	0.137	0.149
Observations	2454	2454	2454	2454

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

Table H6: Difference-in-Differences Results

	(1) Log odds	(2) Log odds	(3) Log odds	(4) Log odds
Treated Group x Year=2009	-0.456**	-0.457**	-0.486**	-0.488**
	(0.150)	(0.147)	(0.152)	(0.150)
Treated Group x Year=2010	-0.928***	-0.887***	-0.989***	-0.948***
	(0.160)	(0.156)	(0.162)	(0.159)
Spillover Group x Year=2009			-0.404**	-0.379**
			(0.130)	(0.130)
Spillover Group x Year=2010			-0.817***	-0.767***
			(0.213)	(0.211)
Year=2006	-0.370***	-0.569***	-0.370***	-0.584***
	(0.0661)	(0.0994)	(0.0661)	(0.0997)
Year=2007	-0.467***	-0.564***	-0.467***	-0.566***
	(0.0659)	(0.0729)	(0.0660)	(0.0729)
Year=2009	-0.948***	-0.920***	-0.918***	-0.893***
	(0.0769)	(0.0763)	(0.0816)	(0.0810)
Year=2010	-0.672***	-0.730***	-0.611***	-0.673***
	(0.0806)	(0.0830)	(0.0844)	(0.0872)
Lagged Rainfall		0.142**		0.133**
		(0.0483)		(0.0481)
Lagged Rainfall Squared		-0.00329***		-0.00316**
		(0.000971)		(0.000966)
Lagged Temperature		0.175		0.168
		(0.0908)		(0.0903)
Share of Protected Areas		0.603		0.527
		(0.671)		(0.666)
Price of Beef Lagged		-0.0231***		-0.0232***
		(0.00653)		(0.00652)
Price of Crops Lagged		-0.190		-0.269
		(0.331)		(0.336)
Lagged GDP		-0.557		-0.551
		(0.511)		(0.504)
Covariates	NO	YES	NO	YES
R^2	0.098	0.113	0.103	0.118
Observations	2450	2450	2450	2450

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

Table H7: Robustness: Deforestation Average Treatment Effects, Alternative Baseline Years

Average Treatment Effects: Deforestation (km^2)			
	ATT	ATU	ATE
2009			
Baseline 2006	-20.72	[-3.61, -3.19]	[-4.83, -4.44]
	(-23.74, -17.70)	(-3.74, -3.09)	(-4.98, -4.32)
Main Specification	-24.49	[-4.04, -3.67]	[-5.50, -5.15]
	(-27.64, -21.34)	(-4.16, -3.58)	(-5.65, -5.04)
Baseline 2007	-28.26	[-4.48, -4.14]	[-6.18, -5.86]
	(-32.17, -24.36)	(-4.60, -4.05)	(-6.33, -5.73)
2010			
Baseline 2006	-52.69	[-6.71, -5.72]	[-10.00, -9.08]
	(-57.50, -47.88)	(-6.85, -5.61)	(-10.17, -8.92)
Main Specification	-57.97	[-6.84, -5.95]	[-10.49, -9.66]
	(-62.95, -52.99)	(-6.97, -5.85)	(-10.66, -9.52)
Baseline 2007	-63.25	[-6.97, -6.18]	[-10.99, -10.25]
	(-69.13, -57.37)	(-7.10, -6.07)	(-11.17, -10.08)

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

Table H8: Robustness: Ex-Post Optimal, Alternative Baseline Years

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	-	1.20	-
Main Specification	1.06	-	1.06	-	1.22	-
Baseline 2007	1.06	-	1.07	-	1.24	-
Total Carbon Emissions						
Baseline 2006	1.05	597	1.08	920	1.23	2,793
Main Specification	1.05	622	1.08	975	1.26	3,094
Baseline 2007	1.05	646	1.09	1,031	1.28	3,383

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

Table H9: Robustness: Deforestation Average Treatment Effects, Trimming

Average Treatment Effects, Trimming: 2.5th and 97.5th Percentiles			
	ATT	ATU	ATE
2009	-24.49	[-4.05, -2.85]	[-5.51, -4.39]
	(-27.64, -21.34)	(-4.17, -2.76)	(-5.66, -4.28)
2010	-57.97	[-6.86, -5.14]	[-10.51, -8.91]
	(-62.95, -52.99)	(-6.98, -5.03)	(-10.68, -8.76)
Average Treatment Effects, Trimming: 3.5th and 96.5th Percentiles			
	ATT	ATU	ATE
2009	-24.49	[-4.04, -3.71]	[-5.50, -5.20]
	(-27.64, -21.34)	(-4.16, -3.63)	(-5.64, -5.09)
2010	-57.97	[-6.83, -6.60]	[-10.48, -10.27]
	(-62.95, -52.99)	(-6.95, -6.51)	(-10.65, -10.12)

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

Table H10: 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	1.04	-	1.05	-	1.20	-
Total Carbon Emissions	1.03	429	1.06	746	1.24	2,906
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	1.07	-	1.07	-	1.23	-
Total Carbon Emissions	1.06	709	1.09	1,077	1.27	3,178

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

Table H11: Robustness: Deforestation Average Treatment Effects, Spillovers

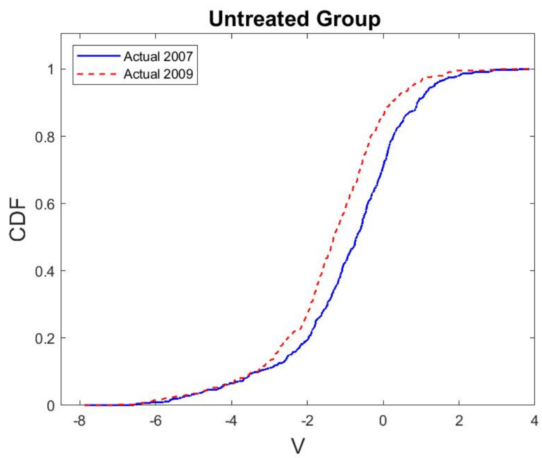
	ATT		ATU		ATS		ATE	
Average Treatment Effects, Spillovers: Above 65 Percent of the Threshold								
2009	-27.25		[-3.76,	-3.28]	[-10.91,	-10.86]	[-6.00,	-5.59]
	(-30.55,	-23.95)	(-3.85,	-3.21)	(-12.61,	-9.25)	(-6.15,	-5.47)
2010	-58.38		[-6.14,	-5.60]	[-18.92,	-16.14]	[-10.89,	-10.21]
	(-63.29,	-53.48)	(-6.25,	-5.52)	(-20.52,	-14.38)	(-11.06,	-10.06)
Average Treatment Effects, Spillovers: Above 75 Percent of the Threshold								
	ATT		ATU		ATS		ATE	
2009	-27.02		[-3.88,	-3.41]	[-11.77,	-11.74]	[-6.00,	-5.59]
	(-30.26,	-23.78)	(-3.98,	-3.34)	(-14.23,	-9.40)	(-6.14,	-5.47)
2010	-56.68		[-6.15,	-5.59]	[-21.18,	-18.28]	[-10.65,	-9.99]
	(-61.54,	-51.82)	(-6.26,	-5.50)	(-23.44,	-15.70)	(-10.81,	-9.84)

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

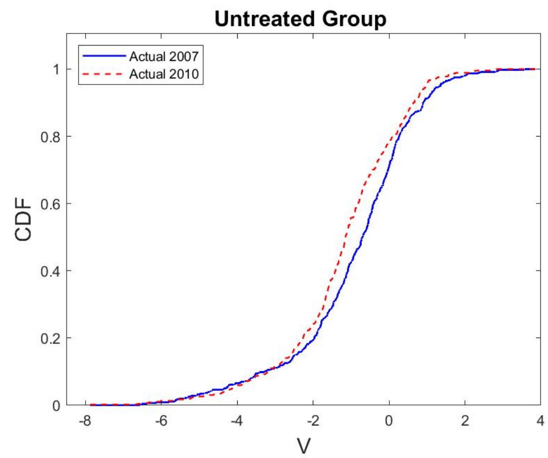
Table H12: Robustness: Ex-Post Optimal, Spillovers

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	1.11	-	1.09	-	1.26	-
Total Carbon Emissions	1.10	1,118	1.11	1,164	1.30	3,166
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	1.08	-	1.07	-	1.24	-
Total Carbon Emissions	1.07	834	1.09	991	1.28	3,174

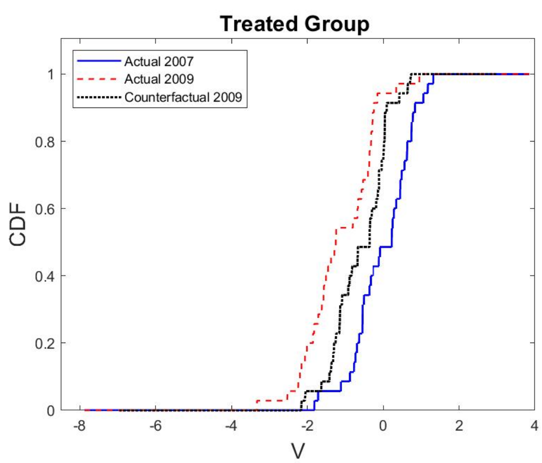
Notes: 'Ratio' divides total deforestation (total emissions) evaluated at the observed list by the ex-post optimal total deforestation (total emissions). 'Value' takes their difference. Values are measured in million US$, assuming a social cost of carbon of US$ 20/tCO2.



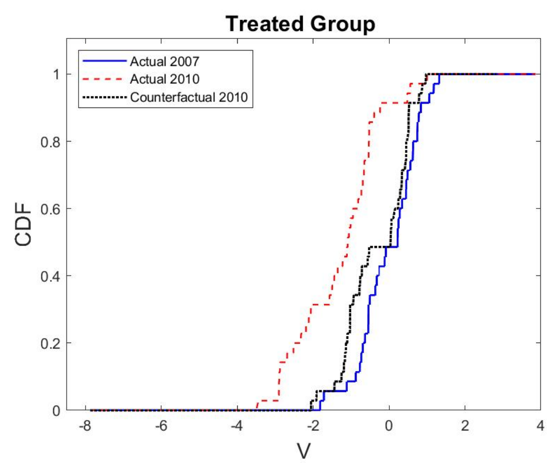
(a) Untreated, 2007 and 2009



(b) Untreated, 2007 and 2010

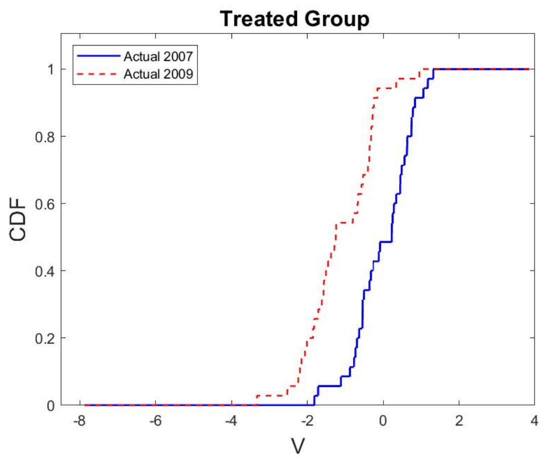


(c) Treated, 2007 and 2009

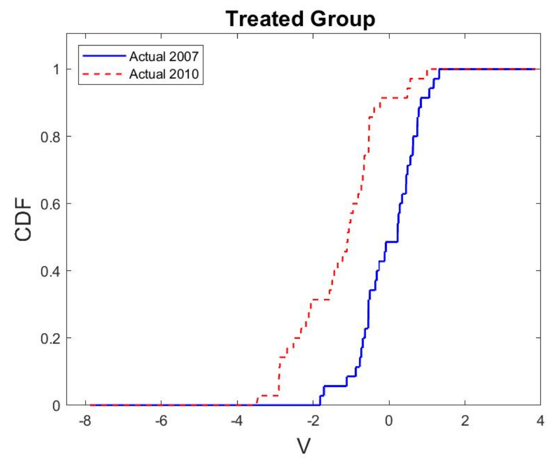


(d) Treated, 2007 and 2010

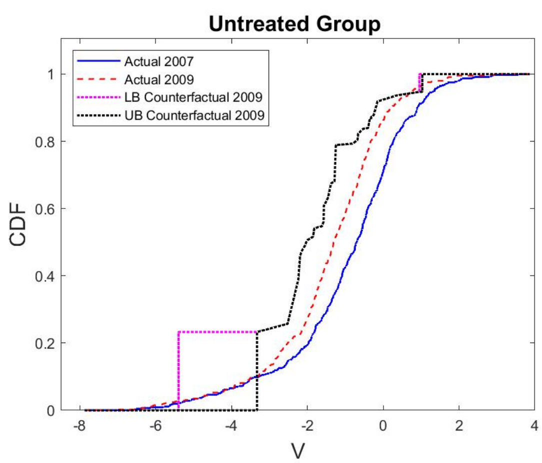
Figure H2: Factual and Counterfactual Distributions of Residuals V



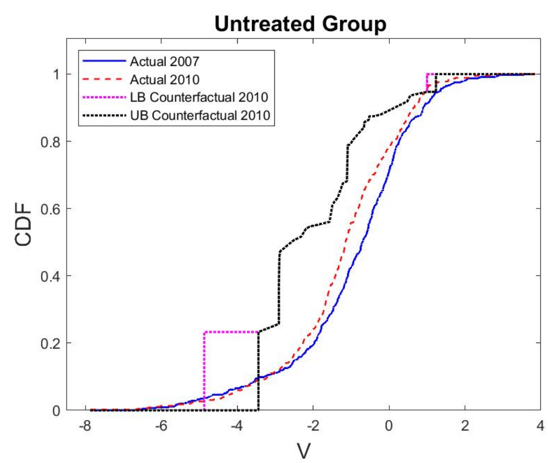
(a) Treated, 2007 and 2009



(b) Treated, 2007 and 2010

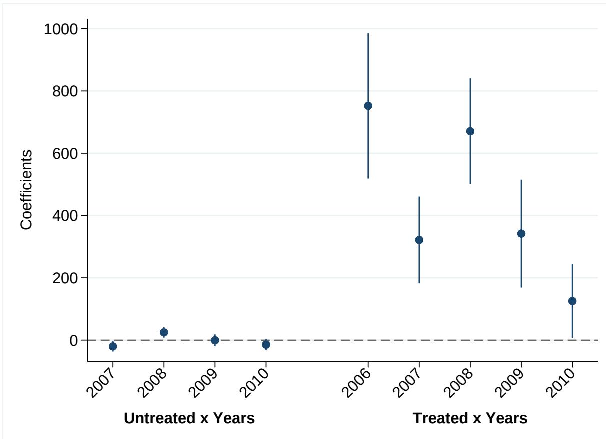


(c) Untreated, 2007 and 2009

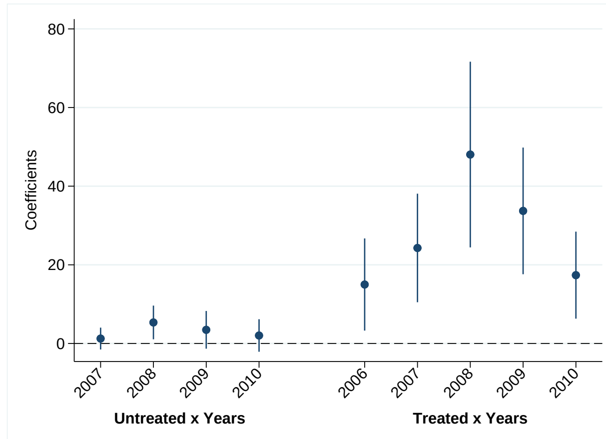


(d) Untreated, 2007 and 2010

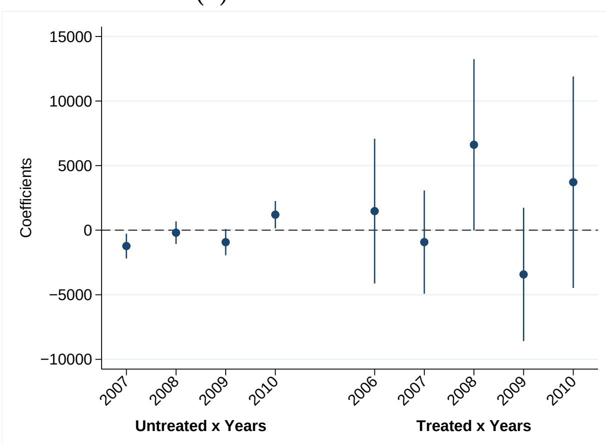
Figure H3: Factual and Counterfactual Distributions of Residuals V



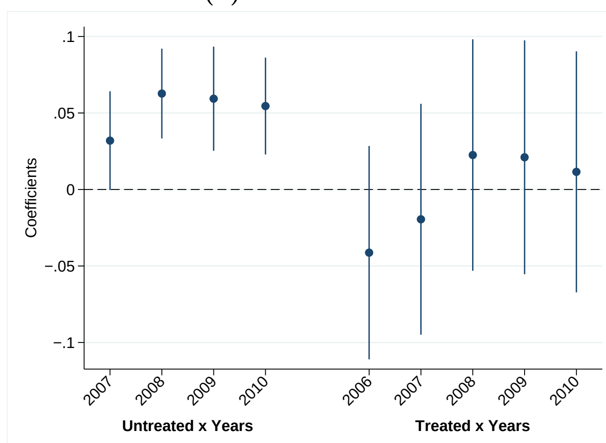
(a) Number of alerts



(b) Number of fines

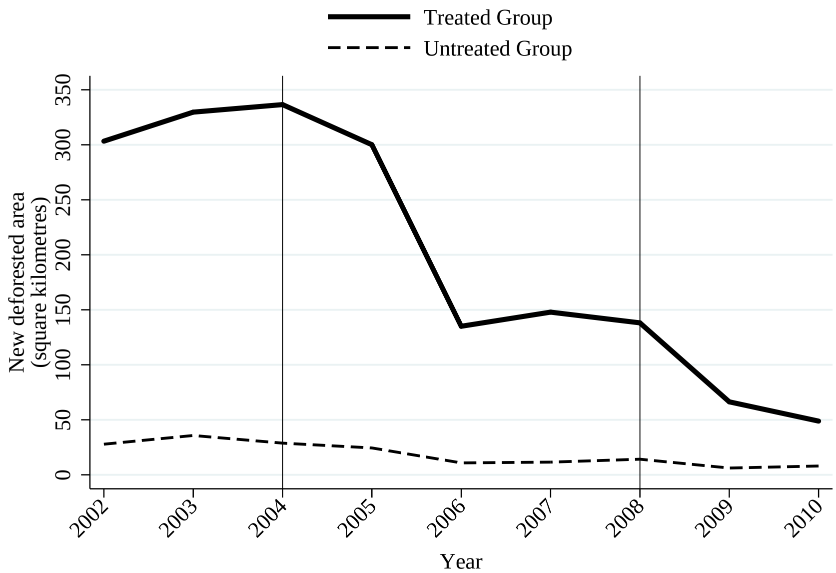


(c) Total rural credit

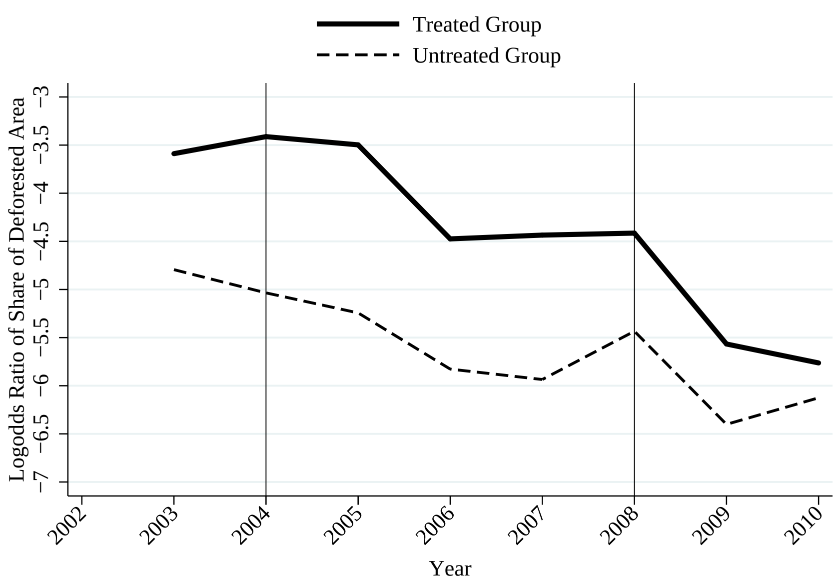


(d) Share of Protected Areas

Figure H4: Coefficients on Treatment Status, by Year: Alerts, Fines, Rural Credit, and Protected Areas

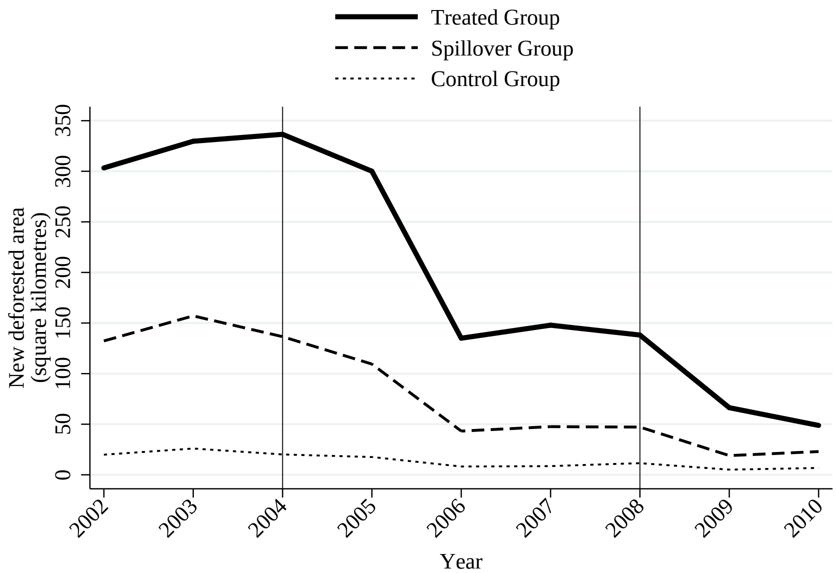


(a) Within-group average of newly deforested area, by year

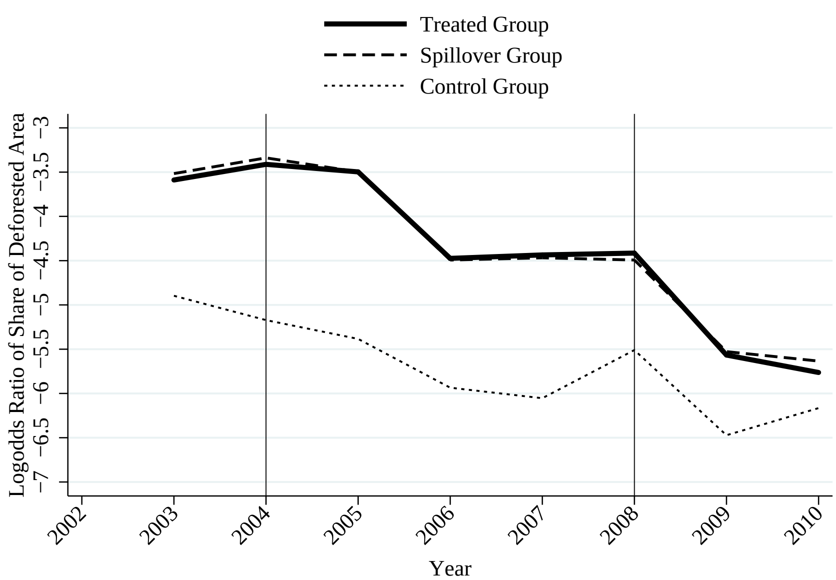


(b) Within-group average of log odds ratio of deforestation shares, by year (using 2002 as the base to construct “remaining forest” in the share)

Figure H5: Evolution of Deforestation by Priority Status: level and log odds ratio of shares



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



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

Figure H6: Evolution of Deforestation by Group: level and log odds ratio of shares

References

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

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

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.

Breiman, L., J. Friedman, R. Olshen, and C. Stone (1984). *Classification and Regression Trees* (1st ed.). Taylor & Francis.

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*.

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.

Farrokhi, F. and H. S. Pellegrina (2021). Trade, technology, and agricultural productivity. Working paper, Purdue University.

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

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

Heckman, J. J., J. Smith, and N. Clements (1997, 10). Making The Most Out Of Programme Evaluations and Social Experiments: Accounting For Heterogeneity in Programme Impacts. *The Review of Economic Studies 64*(4), 487–535.

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

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

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/*.

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

Souza-Rodrigues, E. (2019). Deforestation in the Amazon: A unified framework for estimation and policy analysis. *Review of Economic Studies 86*(6), 2713–2744.

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