Carbon prices and forest preservation over space and time in the Brazilian Amazon*†

Juliano Assunção (Climate Policy Initiative and PUC-Rio)
Lars Peter Hansen (University of Chicago)
Todd Munson (Argonne National Laboratories)
José A. Scheinkman (Columbia University)

December 12, 2023

Abstract

Some portions of land in Brazilian Amazon are forested, and other portions used in agricultural activities, principally cattle-ranching. Deforestation (reforestation) emits (captures) carbon, which has consequence for the global climate. The social and private productivities for these alternative land uses vary across locations within the Amazon region. In this research, we build and analyze a spatial/dynamic model of socially efficient land allocation to establish a benchmark for *ad-hoc* policies. We show how to incorporate the stochastic evolution of cattle prices, and we explore the consequences of ambiguity in the location-specific productivities on the socially efficient policy. Finally, we assess the consequences of imposing alternative social costs of carbon emissions on the spatial/dynamic allocation of land use. Our results indicate that with modest transfers per ton of net CO_2, Brazil would find it optimal to choose policies that produce substantial capture of greenhouse gasses in the next 30 years, suggesting that the management of tropical forests could play an important role on climate change mitigation in the near future.

*This draft is preliminary and incomplete.

†We thank Pengyu Chen, Patricio Hernandez, João Pedro Vieira, Daniel (Samuel) Zhao for their expert research assistance and to Diana Petrova and Carmen Quinn for editorial assistance.

1 Introduction

This paper investigates the potential social gains to designing prudent policies that combat deforestation in the Brazilian Amazon through the lens of a spatial and dynamic model. We build the model to capture the trade-off between agricultural production[1] and forest preservation or regeneration.

The Amazon forest contains 123 ± 31 billion tons of captured carbon that can be released into the atmosphere, equivalent to the historical cumulative emissions of the United States (Malhi et al. (2006), Friedlingstein et al. (2022)). The Brazilian Amazon occupies 60% of the 2.7 million square miles that comprise the Amazon. From 1985 to 2021, the agricultural area in the Brazilian Amazon increased from 68.6 to 240.5 thousand square miles. The associated deforestation, comprising an area the size of Texas, has resulted in high emissions, setting the Brazilian Amazon as a substantial outlier in a plot of countries' emissions per-capita *vs.* GDP per-capita. (see Figure 1.)

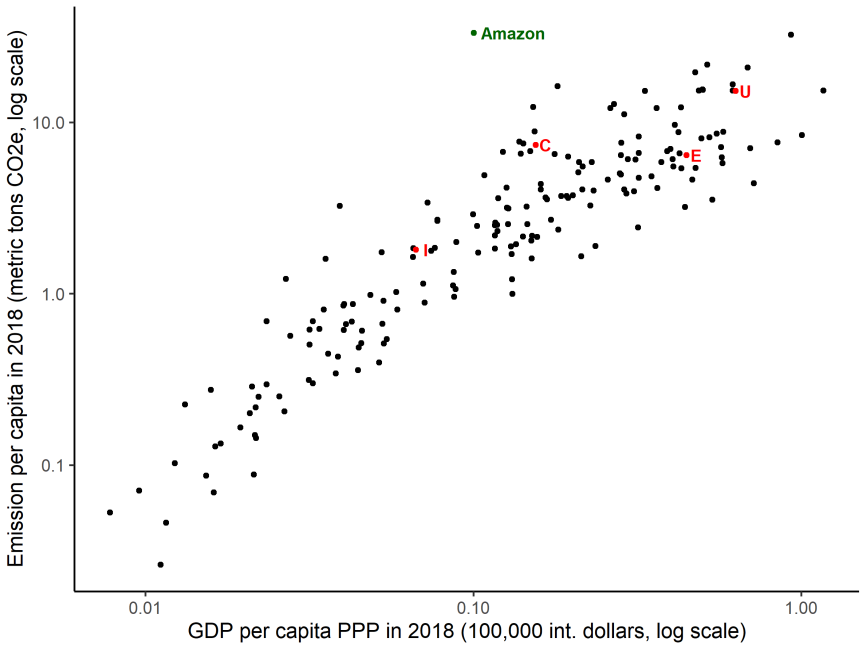


Figure 1: Each dot represents a country in 2018, except for the European Union and the Brazilian Amazon. Highlighted letters stand for (C)hina, (I)ndia, (E)uropean Union, and (U)nited States. Sources: World Bank Data, downloaded on March 2021; Fatos da Amazônia 2021 (www.amazonia2030.org).

We use the model to provide insights into structuring policy improvements that realign economic incentives for allocating land use. We make the model quantitative through the use of detailed spatial information from multiple data sets. Our data document large cross-sectional variability in cattle farming productivity and in the potential absorption of carbon in the Brazilian Amazon. To

[1]Since close to 90% of the deforested land in the Amazon Biome is currently used for pasture, we identify agriculture with cattle farming in this paper, and use the two terms interchangeably.

account for this variability, we divide the Amazon region into various subregions or sites making the spatial dimension important in our analysis. While the model has considerable cross-sectional richness, it is nevertheless highly stylized for reasons of tractability and transparency.

We pose the model in continuous time. The cross-sectional heterogeneity in productivities and the natural state constraints on the land allocation preclude standard recursive methods for solving so-called Hamilton-Jacobi-Bellman (HJB) equations. Instead, we use and extend methods from Modified Predictive Control (MPC) that were originally developed in control theory and engineering to study multi-plant production in real time. MPC methods approximate inequality constraints on the states using what is called an interior point method. They allow for uncertainty specified as a Markov process by incorporating a shorter uncertainty horizon than the overall control horizon as a means of approximation. We extend these methods to explore subjective ambiguity from the standpoint of the social planner by making a model-determined robustness adjustment to the subjective probabilities for the unknown parameters. We are once again pushed to use numerical methods, in this case a Markov chain Monte Carlo method based on Hamiltonian dynamics. Such a method is particularly valuable for high dimensional problems than the familiar Metropolis-Hastings approach. We use the Hamiltonian Markov chain approach in a novel way to confront what is sometimes referred as "deep uncertainty." Alternatively, we may interpret this treatment of subjective ambiguity as a robust Bayesian formulation of the control problem of interest.

Our analysis of the model implications proceeds in four steps. First, we use the model to elicit an estimate of the shadow price for emissions revealed by the deforestation during 1995-2008. The year 1995 is the first date at which we have reliable price data on cattle prices.[2] The year 2008 marks the beginning of the Amazon Fund, financed primarily by the Norwegian and German governments. Their funding was a pay-for-performance scheme based on an emissions price of \$5 per ton of CO2e. It generated to USD 1.2 billion in payments in 2008-2017 (Angelsen (2017), Correa et al. (2019)) and provides and example of how deforestation may be influenced by outside payments. We use this shadow price to produce simulations designed to capture "business as usual." We also use this shadow price to measure the value for Brazilians of the "forest services" provided by preserved areas. These services include climate services as well as the economic value of production that occurs without destroying the Amazon forest.[3]

We then study the impact of adding outside payments for *net* capture of CO_2 in the Amazon. We produce results at three levels of aggregation in steps two, three and four. In step two we construct the finest grid by considering 1043 sites in the Amazon biome with each measuring 67.5km$\times$ 67.5km. For this level of detail, we produce results without accounting for uncertainty in the price of agricultural output, by assuming that the price corresponds to the (stationary) average of the two-state Markov process we fit to the observed agricultural prices. We also use these deterministic solutions to ascertain whether and by how much Brazil would gain if the country

[2]Until mid-94 Brazil went through a period of very high and volatile inflation.

[3]These include forest products like natural rubber, nuts, and açaí, alongside sustainable timber.

signed an agreement for a set of hypothetical dollar transfers per net ton of CO_2 captured. Our model output shows both the social gains to overall preserving or enhancing the Brazilian rain forest and to reallocating production in the cross section.

In a third step, we consider 78 sites, each of which is 270km×270km. For this level of aggregation, we produce results that take into account stochastic changes in the price of agricultural output as well as, for comparison, deterministic results for this 78-site resolution. The results are quite close, showing that our stochastic representation of externally determined agricultural prices plays a minor role in the analysis.

Estimates of crucial productivity parameters in the model are subject to non-trivial uncertainty. We allow productivities for carbon absorption and agriculture to be site specific. Moreover, while we have cross-sectional data that are informative, we use regression methods to provide the inputs needed for our analysis. Uncertainty in the regression coefficients induce uncertainty in the site specific productivity parameters. The typical approach to uncertainty quantification, explores such parameter uncertainty from the perspective of an external analyst, but we instead incorporate explicity into the decision problem. While it is appealing to address the parameter uncertainty probabilistically, there is uncertainty as to what probabilities to impose. We explore the consequences of incorporating this productivity uncertainty into the decision problem of the fictitious social planner.

In the fourth step, we explicitly consider parameter uncertainty from the perspective of the social planner. We start with a prior distribution over site-specific productivity parameters for carbon sequestration and agricultural productivity. Since the choice of prior distribution is *ad hoc* and uncertain, we engage in a prior sensitivity analysis subject to penalization. This sensitivity analysis uses the prior as a baseline distribution and the planner's objective function to ascertain the which departures from the baseline are of most concern to the planner. The magnitude of the penalty parameter limits how much sensitivity is entertained and serves as the inverse of an ambiguity aversion parameter. For the computation of the optimal solutions in this case of parameter uncertainty we use a Markov chain Monte Carlo method that is based on Hamiltonian dynamics and that is often more efficient for high dimensional problems than Metropolis-Hastings

While we assume that the planner can directly control deforestation, we also view the solution to the planner's problem as providing a benchmark for comparing the outcomes of alternative *ad hoc* policies, and suggesting improvements over current policies.[4]

The rest of the paper is organized as follows: In the next section, Section 2 we review some of the relevant literature. This is followed in Section 3 with an exposition of our theoretical model. Section 4 summarizes how we use a large collection of relevant data sets to calibrate the model. Section 5 discusses the numerical methods used to compute solutions to the social planner's

[4]Brazil has experience implementing effective policies to curb deforestation. The launch of satellite-based monitoring systems, the creation of protected areas, the enactment of conditioned-credit measures and the creation of a priority list of municipalities comprised a new approach to combat deforestation, resulted on a reduction of more than 80% in the deforestation rates between 2004 and 2012 (Gandour (2018), Assunção and Rocha (2019), Assunção et al. (2020), Assunção et al. (2022a) and Assunção et al. (2022b)).

maximization problem. In particular Section 5.1 treats the case of parameter uncertainty. Our results are presented in Section 6, which is followed by our conclusions and suggestions for further work.

2 Related Substantive Literature

Griscom et al. (2017) identify and quantify "natural climate solutions" (NCS), which include tropical forests. Heinrich et al. (2021) focuses on the potential of the Brazilian Amazon. As noted by Balboni et al. (2022), most studies on agricultural expansion and deforestation are static.

A recent branch of the literature uses discrete-choice models in order to study the link between agriculture and deforestation (Souza-Rodrigues (2019), Dominguez-Iino (2021), Araujo et al. (2022)). Souza-Rodrigues (2019) and Dominguez-Iino (2021) develop static approaches, emphasizing the role of the transportation network and trade to the design of policies, without explicitly modelling the carbon cycles associated with the forest. Araujo et al. (2022), on the other hand, presents a dynamic model along the lines of Scott (2014), allowing farmers to internalize the social value of carbon. However, the dynamics in Araujo et al. (2022) is restricted to the forward-looking behavior of farmers.

In contrast to the existing literature, our dynamic approach not only accounts for how expected future prices of agricultural goods influence optimal current land use, it also incorporates carbon emissions from deforestation and carbon capture from forest regeneration. The carbon cycle associated with abandoning agriculture and allowing forest to regrow naturally is considered explicitly in the model. In contrast, the simulation of carbon prices in Souza-Rodrigues (2019) works as a tax to the expansion agriculture into forests and it is not associated with the carbon cycle directly. Consequently, we provide a framework that integrates the impact of carbon prices on deforestation, forest restoration, and agriculture. In addition, we also take into account the uncertainty on the forest carbon measures.

Finally, our results contribute to the literature on climate policy design. Our simulation shows that deforestation in the Amazon will cross the tipping point of $20-25\%$ suggested by Lovejoy and Nobre (2018) in the scenario with carbon price equal to the shadow emission price elicited from the 1995-2008 period with no additional international payments.[5] On the other hand, additional payments of at least \$15/ton would not only safeguard the tipping point, but would also trigger forest restoration on large scale. In this sense, the carbon sink potential of secondary forests emphasized by Griscom et al. (2017) and Heinrich et al. (2021) can be realized with sufficient additional carbon payments.

[5] Notice that, while our shadow emission prices vary with the model and resolution chosen, they are all reasonably close to the \$7.26/ton estimated by Araujo et al. (2022).

3 Model

We pose the problem of a fictitious social planner who considers the trade-off between using land for agriculture and nurturing or preserving forests that function as carbon sinks. This planner internalizes the externalities resulting from deforestation. The planner's problem is dynamic with explicit heterogeneity across regions in the Amazon. Guided by empirical measurements, the regions have two important sources of heterogeneity: i) agricultural productivity and ii) ability to absorb atmospheric carbon.

Let i denote a site index for $i = 1, 2, ..., I$ where I is the total number of sites and $t \in [0, T]$ the point in time. We use superscripts to denote sites and subscripts to denote dates. We adopt the notational convention that uppercase letters depict the actual state and lower case letters the potential state realizations. At date t,

Z_t	$\stackrel{\text{def}}{=}$	$(Z_t^1, Z_t^2, ..., Z_t^I)$	vector of area used for agriculture expressed in hectares
X_t	$\stackrel{\text{def}}{=}$	$(X_t^1, X_t^2, ..., X_t^I)$	vector of carbon captured expressed in Mg CO2e (CO2 equivalent)
A_t	$\stackrel{\text{def}}{=}$	$(A_t^1, A_t^2, ..., A_t^I)$	vector of agricultural output

We use the notation $Z \stackrel{\text{def}}{=} \{Z_t : 0 \leqslant t \leqslant T\}$ to denote the corresponding process that evolves over time, and similarly for other states and controls. In our base model the single aggregate state variable is P_t^a, an index of cattle prices in Brazil expressed in 2017 US dollars.[6]

The state vector Z_t is subject to an instant-by-instant and coordinate-by-coordinate constraint:

$$0 \leqslant Z_t^i \leqslant \bar{z}_i$$

where $\bar{z}_i$ is the amount of land in the Amazon biome available for agriculture at site i.[7] Let $\dot{Z}_t$ be the time derivative of Z at date t.

The evolution of X^i introduces an important asymmetry into our problem. We may write a "linear" version of this problem by introducing two site specific, scalar, non-negative control variables for our fictitious planner, U_t^i and V_t^i, that distinguish positive from negative movements in the derivative of Z_t^i :

$$\dot{Z}_t^i = U_t^i - V_t^i. \tag{1}$$

The site specific state variable process X^i evolves as:

$$\dot{X}_t^i = -\gamma^i U_t^i - \alpha \left[X_t^i - \gamma^i \left(\bar{z}^i - Z_t^i \right) \right] \tag{2}$$

where the parameters satisfy: $\gamma^i > 0, \alpha > 0$, for $i = 1, 2, ..., I$. The first term in the right side of (2) connects deforestation to a loss in captured carbon. The site-specific parameter $\gamma^i > 0$ denotes

[6] We choose cattle prices because, in recent years, more than 85% of deforested land is dedicated to cattle grazing - soybean, the largest crop in the region, accounts for about 8% of the farming land (Mapbiomas - www.mapbiomas.org).

[7] For calibration of this and the other parameters see Section 4

the density of CO2e that is present in a primary forest in site i.[8] The next term expresses the growth in captured CO2e, when the size of the forest in site i is held constant. The mean-reversion coefficient α guarantees that if one lets the forest grow undisturbed in a deforested area, it would reach $[1 - \exp(-\alpha 100)]\%$ of the maximal captured CO2e in 100 years as in Heinrich et al. (2021). In our case, we choose α such that $[1 - \exp(-\alpha 100)] = 99\%$. Notice that, holding constant the deforested area, the amount of carbon in a site converges to γ^i per hectare of remaining forest. In Remark 3.3, we argue that at the optimum $U_t^i V_t^i = 0$. Thus one of the controls is always zero at the boundary, which introduces additional binding constraints into the analysis.

We model cattle output as proportional to the land allocated to cattle farming,

$$A_t^i = \theta^i Z_t^i \tag{3}$$

where θ^i is a site-specific productivity parameter.

All of the locations contribute to emissions via the capture of carbon and emissions that result because of agricultural activity with a net impact given by

$$\kappa \sum_{i=1}^{I} Z_t^i - \sum_{i=1}^{I} \dot{X}_t^i, \tag{4}$$

where the parameter κ captures the emissions that result because of cattle farming. We include a cost of adjustment to changes in the use of land with contributions from each site. It is measured by

$$\frac{\zeta}{2}\left[\sum_{i}^{i}\left(U_t^i + V_t^i\right)\right]^2.$$

The price process P^a for the agricultural output evolves exogenously as an n-state Markov chain in continuous time with time invariant transitions. This process has an infinitesimal generator represented as an intensity matrix $\mathbb{M}$ with nonnegative entries off-diagonal entries $\mathsf{m}_{\ell\ell'} \geqslant 0$ for $\ell' \neq \ell$ and diagonal entries

$$\mathsf{m}_{\ell\ell} = -\sum_{\ell'=1,\ell'\neq\ell}^{n} \mathsf{m}_{\ell\ell'}.$$

The implied transition probability matrix over an interval of time τ is $\exp(\tau\mathbb{M})$ computed using a matrix counterpart to a power series.

Since many carbon trading schemes are based on emissions, we assume that the planner takes as given a price for carbon emissions P^e, the initial price for agriculture and the Markov process that describes the future evolution of the price P_t^a for cattle and maximizes

$$\mathbb{E}\left\{\int_0^\infty \exp(-\delta t)\left[-P^e\left(\kappa\sum_{i=1}^{I} Z_t^i - \sum_{i=1}^{I}\dot{X}_t^i\right) + P_t^a \sum_i \theta^i Z_t^i - \frac{\zeta}{2}\left(\sum_i U_t^i + V_t^i\right)^2\right] dt\right\} \tag{5}$$

[8] For simplicity, equation (2) assumes that all deforestation occurs in primary forest, what is not far from what has been observed in the Brazilian Amazon.

subject to equations (1)-(2), and the control restrictions:

$$U_t^i \geqslant 0, \qquad V_t^i \geqslant 0 \qquad t \geqslant 0.$$

where δ is the subjective discount rate. The emission price P^e considered by the planner would be the sum of the (constant) marginal value attributed by the planner to emission and any monetary transfers obtained from others, such as sales in carbon emission markets. In future work we intend to make P^e a state variable.

Remark 3.1. *The objective function* (5) *values agricultural output by the value of sales, thus assuming that inputs to production have no alternative use. This choice is dictated by lack of data on the cost of attracting or redeploying agricultural inputs, but it biases the results in favor of agricultural use.*

Remark 3.2. *The only interaction across sites in objective function* (5) *occurs through the adjustment costs. These interactions are intended to be the result of a less than perfectly elastic supply of resources needed for changing land use at the level of the whole Amazon. We could have introduced instead interactions across sites via nonlinearities in the valuation of agricultural output and/or emissions.*

Remark 3.3. *To argue that the controls U_t^i and V_t^i satisfy the complementary slackness condition $U_t^i V_t^i = 0$ for each pair (i,t) it is easier to consider a discrete time model. The proof for the analogous result for the continuous time case goes through by taking limits. Suppose you take take a point where the optimal trajectory involves $\min\{U_t^i, V_t^i\} > \Delta > 0$. If the planner lowers both controls by Δ, then at time t, one obtains an increase of Δ in X_t^i and lower emissions $\gamma^i\Delta$. Equation* (2) *implies that X_t^i would have a lower drift and converge over time to the stationary solution. This in turn implies that the sum of future emissions would increase by $\gamma^i\Delta$. However since the rate of discount is positive, the value of the problem would increase. Thus an optimal solution cannot involve simultaneously positive values for U_t^i and V_t^i.*

Remark 3.4. *Optimization problem 5 does not involve the stocks of (extended) carbon in the atmosphere generated by activities in the Amazon biome. However given the optimal solution one could derive the implied impact on the evolution of carbon stocks, by appending a model that maps emissions into carbon in the atmosphere.*

3.1 Adding parameter uncertainty

We investigate a static formulation of robustness to parameter uncertainty. For each site, we consider the parameter pair $\varphi^i = (\gamma^i, \theta^i)$ for $i = 1, 2, ..., I$. Let $\boldsymbol{\varphi}$ denote full parameter vector including all sites and hence of dimension $2 \times I$.

To reduce the dimension of the problem, we assume that the parameter vector $\boldsymbol{\varphi}$ that describes the full parameter vector of the pairs (γ^i, θ^i) can be written as $\boldsymbol{\varphi} = \phi(\boldsymbol{\beta}) := \exp((\beta)R^i)$, where $dim(\boldsymbol{\beta}) = 13$ and the vector R^i is a vector of geographical variables of site i. Given $\boldsymbol{\beta}$ we

write $f(\boldsymbol{d}, \boldsymbol{\beta})$ for the discounted utility obtained from a sequence of decisions $\boldsymbol{d}$ when the value of $\boldsymbol{\varphi} = \phi(\boldsymbol{\beta})$ (see A). This value depends on the choices of P^e and P^a that we omit from the notation. Our planner is uncertain about the distribution of the parameter $\boldsymbol{\beta}$ and instead has a baseline probability distribution π, which she treats as a rough approximation. We address this uncertainty by introducing ambiguity about the parameter distribution. The distribution π depends on additional parameters $\sigma_\gamma^2, \sigma_\theta^2$, and to simplify the notation we write $\boldsymbol{\rho} := (\boldsymbol{\beta}, \sigma_\gamma^2, \sigma_\theta^2)$. We use a divergence measure to capture ambiguity about the parameter distribution. For $\int g(\boldsymbol{\rho}) d\pi(\boldsymbol{\rho}) = 1$, the relative entropy (or Kullback-Leibler) divergence

$$\int \log g(\boldsymbol{\rho}) g(\boldsymbol{\rho}) d\pi(\boldsymbol{\rho}) \geqslant 0,$$

is a commonly used measure of divergence between a probability $g(\boldsymbol{\rho})d\pi(\boldsymbol{\rho})$ and the baseline $d\pi(\boldsymbol{\rho})$. To produce optimal controls that are robust to the parameter uncertainty, we solve

Problem 3.5.

$$\max_{\boldsymbol{d}} \min_{g, \int g d\pi = 1} \int f(\boldsymbol{d}, \boldsymbol{\beta}) g(\boldsymbol{\rho}) d\pi(\boldsymbol{\rho}) + \xi \int \log g(\boldsymbol{\rho}) g(\boldsymbol{\rho}) d\pi(\boldsymbol{\rho})$$

where $\xi > 0$ is penalty parameter.[9]

One may treat a full commitment to the baseline distribution by making ξ arbitrarily large. More modest settings capture a concern for robustness.

4 Calibration

We present here a summary of our calibration procedure. Appendix A describes in detail all the data used for the calibration of parameters and initial conditions. In order to calibrate our model for the Brazilian Amazon, we combine data from multiple sources. We use the year of 2017 as a reference for many variables, since this is the year of the latest Agricultural Census in Brazil. The agricultural census provides information on the value of cattle sold for slaughter per hectare of pasture land at the level of municipality. To fill in censored values and provide a smoother representation of the technology, we use the predicted value of a regression on geographical variables (see Appendix A).

To produce values for γ, we first use data from MapBiomas[10] to select pixels of $100m \times 100m$ that can be considered primary forests and data from 2017 from ESA Biomass[11] to obtain carbon per hectare. We then calculate average γ's for each municipality and, in analogy with the procedure for θ, we produce predicted values across municipalities from a regression on geographical variables to obtain a smoother representation (see Appendix A).

We project these municipal estimates into two grids of the Amazon biome. At the most detailed level we consider a regular grid of the Amazon region each comprising 2,250 high-resolution pixels

[9] Alternatively, we can think of ξ a as Lagrange multiplier on a relative entropy divergence constraint.
[10] www.mapbiomas.org (Collection 5).
[11] See Santoro and Cartus (2021).

($\tilde{3}$0m) from MapBiomas and after discarding grid elements with less than 3% of their area in the Amazon Biome are left with 1043 sites. To allow us to treat the cases of agricultural price or parameter uncertainty we will aggregate 16 sites of finer grid to produce sites of $\approx$ 268km $\times$ 268km. We obtain 78 sites after dropping sites with less than 3% in the Amazon biome.

For the case of parameter uncertainty we use conjugate prior updating (Raiffa et al. (1961)) to produce a baseline distribution of the vector $\boldsymbol{\rho}$. To derive this baseline distribution we use municipal estimates of γ and θ and regress the log of these municipal observations on a set of geographical variables (see Appendix A). We then follow Hansen and Sargent (2013), Section 5.3, to produce a posterior for $\boldsymbol{\beta}$ that we use as the baseline (unadjusted) distribution for $\boldsymbol{\beta}$.[12]

Figure 2 shows the initial land allocated to agriculture and the initial stock of absorbed carbon across the 1043-grid sites. Figure 3 shows how the carbon sequestration parameter γ^i varies across the different sites, and Figure 4 does the same for the agricultural productivity parameter, θ^i. The correlation between θ^i and γ^i is $-.3$ showing that agricultural productivity and carbon absorption capacity are negatively but imperfectly correlated.

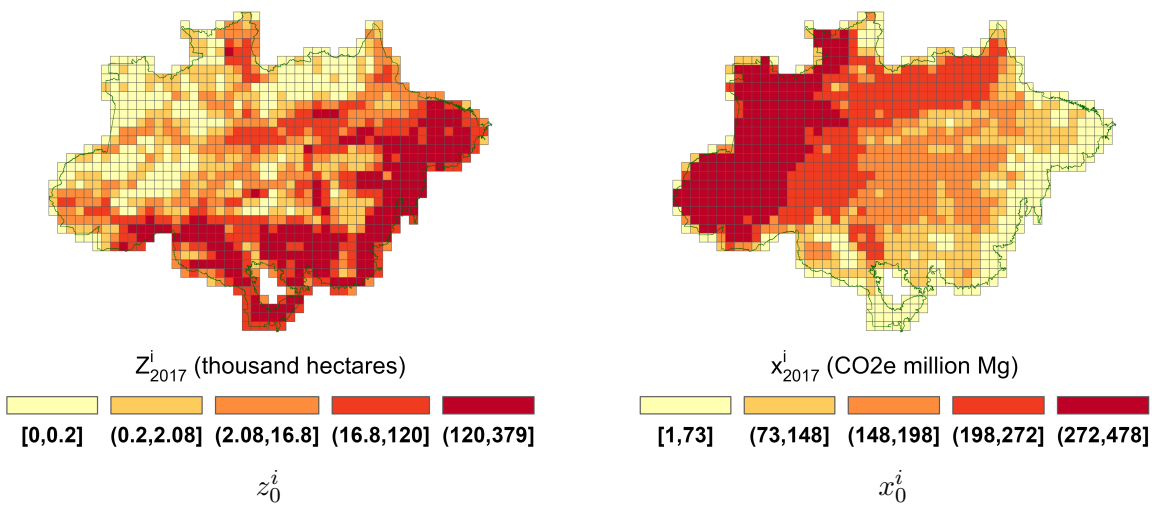


Figure 2: Initial values for agricultural area (z^i_0) and carbon stock (x^i_0)

[12]See Appendix B.

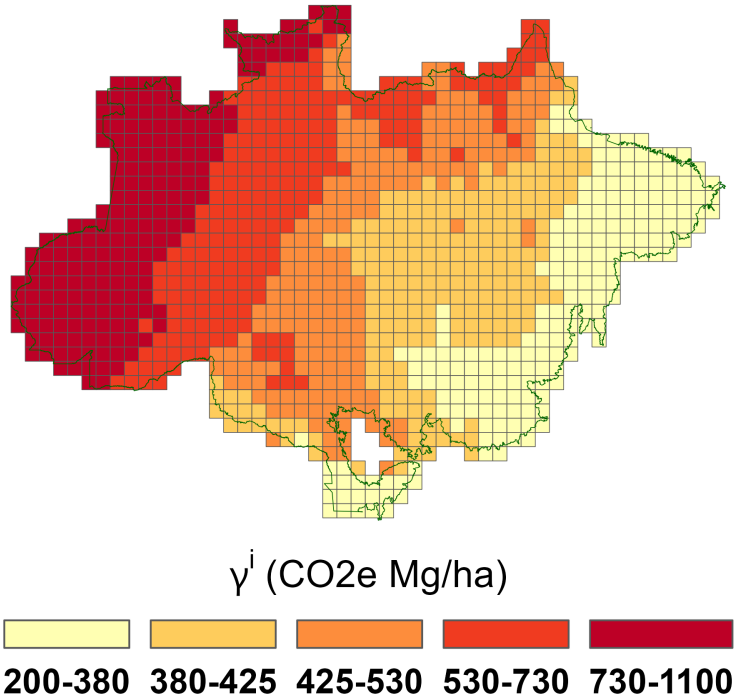


Figure 3: Carbon sequestration parameter heterogeneity.

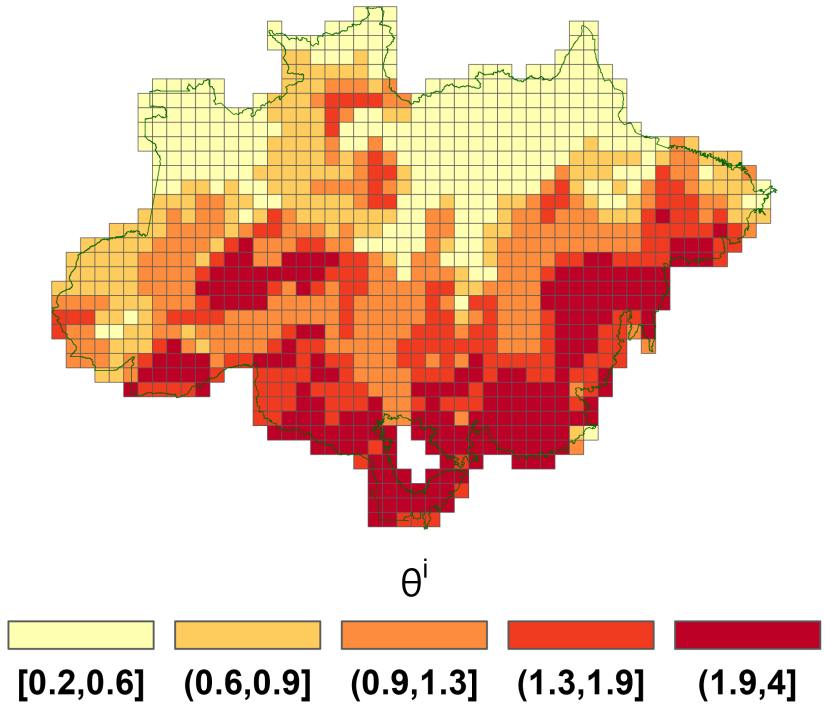


Figure 4: Agricultural productivity heterogeneity

5 Solving the maximization problem

To achieve the needed degree of economic and environmental richness, we use numerical methods to obtain model solutions. To confront this locational heterogeneity, we necessarily have a large number state variables with state-constraints that bind at the optimal solution. To confront the inequality restrictions that are central to our problem, we use a so-called interior point method. This method is implemented by imposing penalties on logarithms on variables constrained to be nonnegative. While the interior point approximation pushes solutions away from their zero boundaries, in practice the solutions will be close enough to zero to identify the binding constraints.

In the absence of price uncertainty, we are able to solve our optimization problem by simply solving a static problem over all possible trajectories over the next 200 years, using interior point methods even for 1043 sites. Given obvious uniform bounds on possible utility flows and discount rates that exceed 2% we can argue that the resulting trajectories achieve close to the infinite horizon optima.

When considering price uncertainty, we find "Modified Predictive Control" (MPC) methods (*e.g.* Scokaert and Rawlings (1998), Bemporad et al. (2002), Thangavel et al. (2018)) to be particularly suitable for solving our planner's problem. Our MPC approximation is implemented as follows.[13] Given the current period, say date zero, looking forward, we break the future into two segments: a) an uncertainty horizon of say τ time periods and b) the remaining $T - \tau$ time periods beyond this uncertainty horizon for which we abstract from uncertainty. While the cattle price distribution follows a Markov chain, to simplify our computations, we impose that the prices in periods $\tau + 1, \ldots T$ are set equal to the value that prevails at τ. We then apply the interior method to find the optimal trajectory at zero given P_0^a. We keep the optimal states for time $t = 1$ and repeat - that is we consider the problem stating at $t = 1$ with the new state vector and divide the future on two segments: an uncertainty segment of length τ and a remaining period of $T - \tau$. This step will produce an optimal state at period 2, we then repeat the procedure to produce the optimal state at period 3 and continue going forward. Thus we confront randomness in this problem by imposing appropriate "measurability" restrictions on the controls.

In practice, the dimensionality of the stochastic problem increases geometrically as a function of the uncertainty horizon, τ. This MPC method becomes tractable when the uncertainty horizon can be relatively short and still obtain good approximations. We determine an "adequate" uncertainty horizon τ^* by checking the difference in the value of the problem $V(\tau) - V(\tau - 1)$ for $\tau = 0, 1, \ldots, \tau^*$. In our example with 81 sites and price uncertainty, we chose $\tau^* = 6$.

Many of the results we will show entail projections into the future. We report results based a common randomly drawn sequence of cattle prices, P_t^a $t = 1, \ldots T$, using the observed P_0^a and the calibrated Markov chain.

[13] Related computational approaches have been proposed by Cai et al. (2017) and Cai and Judd (2023).

5.1 Parameter uncertainty

One nice feature of using relative entropy divergence to explore distributional sensitivity is that the minimization problem has a tractable solution:

$$-\xi \log \int \exp\left[-\frac{1}{\xi} f(\boldsymbol{d},\boldsymbol{\beta})\right] d\pi(\boldsymbol{\rho}) = \min_{g,\int g d\pi=1} \int f(\boldsymbol{d},\boldsymbol{\beta}) g(\boldsymbol{\rho}) d\pi(\boldsymbol{\rho}) + \xi \int \log g(\boldsymbol{\rho}) g(\boldsymbol{\rho}) d\pi(\boldsymbol{\rho}). \tag{6}$$

The minimizing g only depends on the projection of $\boldsymbol{\rho}$ on $\boldsymbol{\beta}$:

$$g_{\boldsymbol{d}}(\boldsymbol{\beta}) = \frac{\exp\left[-\frac{1}{\xi} f(\boldsymbol{d},\boldsymbol{\beta})\right]}{\int \exp\left[-\frac{1}{\xi} f(\boldsymbol{d},\tilde{\boldsymbol{\beta}})\right] d\pi(\tilde{\boldsymbol{\rho}})} = \frac{\exp\left[-\frac{1}{\xi} f(\boldsymbol{d},\boldsymbol{\beta})\right]}{\int_{\mathcal{B}} \exp\left[-\frac{1}{\xi} f(\boldsymbol{d},\tilde{\boldsymbol{\beta}})\right] d\pi_{\boldsymbol{\beta}}(\tilde{\boldsymbol{\beta}})}, \tag{7}$$

where $\pi_{\boldsymbol{\beta}}$ denotes the marginal distribution of $\boldsymbol{\beta}$. Notice that $g_{\boldsymbol{d}}$ tilts the distribution towards smaller objectives for each given decision $\boldsymbol{d}$. The candidate solution presumes that:

$$\int_{\mathcal{B}} \exp\left[-\frac{1}{\xi} f(\boldsymbol{d},\boldsymbol{\beta})\right] d\pi_{\boldsymbol{\beta}}(\boldsymbol{\beta}) < \infty$$

at the robustly optimal $\boldsymbol{d}$, which implicitly limits how fat the tail can be for the distribution π.

Remark 5.1. *The formula on the left-side of* (6) *is a special case of a smooth ambiguity objective, first suggested by Klibanoff et al. (2005). They deduced a rationale for an ambiguity adjustment represented using a concave function distinct from the one used for expressing risk aversion. The negative exponential in* (6) *is such a concave function. The logarithmic adjustment converts this to a certainty equivalent. While they take such a concave adjustment to be a starting point, we deduce this representation from a starting point motivated by robustness. Consistent with this difference, their axiomatic motivation is different from the distributional robustness that is of interest to us.*

For conceptual reasons, we also switch the order of the maximization and minimization. Under quite general conditions, we can invoke the min-max theorem:

Problem 5.2.

$$\min_{g,\int g d\pi=1} \max_{\boldsymbol{d}} \int_{\mathcal{B}} f(\boldsymbol{d},\boldsymbol{\beta}) g(\boldsymbol{\rho}) d\pi(\boldsymbol{\rho}) + \xi \int \log g(\boldsymbol{\rho}) g(\boldsymbol{\rho}) d\pi(\boldsymbol{\rho})$$

where $\xi > 0$ is penalty parameter.

Consider the inner maximization problem:

$$\max_{\boldsymbol{d}} \int_{\mathcal{B}} f(\boldsymbol{d},\boldsymbol{\beta}) g(\boldsymbol{\rho}) d\pi(\boldsymbol{\rho})$$

where we are free to drop the relative entropy penalty as it does not depend on the decision $\boldsymbol{d}$. Provided that this problem has a solution for the outer g minimization, the planner is maximizing against this particular (penalized) "worst case probability." This computation is of interest as a way to interpret the consequences of any given choice of the penalty parameter ξ. In practice, we

find it revealing to explore alternative choices of ξ and deduce their implications for the implied worst-case probabilities.[14]

For solving the robust problem numerically, we take an iterative approach, also supported by the min-max theorem. Specifically, we proceed as follows:

i) Given a g, we solve the maximization problem for a candidate $\boldsymbol{d}$. We ignore the relative entropy penalty term in this solution.

ii) For a given $\boldsymbol{d}$, we solve the minimization problem with the relative entropy term to obtain a new candidate for g.

iii) We repeat the steps until we achieve convergence.

For step ii), we use the quasi-analytical solution in computing (7) and a Markov chain Monte Carlo method that is based on Hamiltonian dynamics and that is often more efficient for high dimensional problems than Metropolis-Hastings (Neal et al. (2011), Carpenter et al. (2017)).[15]

6 Results

6.1 Shadow prices under business as usual

We infer a *shadow* value for the planner based on historical experience. To obtain this value, we first choose a period $[\underline{t}, \bar{t}]$ and then choose a price for emissions, that we denote by P^{ee} to match the aggregate deforestation predicted by the model to the actual aggregate deforestation from the initial period $\underline{t}$ to a final observation period $\bar{t}$. More precisely, we find the optimal trajectory from $\underline{t}$ to $\bar{t}$ given the observed initials condition $(X^o_{\underline{t}}, Z^o_{\underline{t}})$, and the realized history of agricultural prices $\{P^a_t\}^{\bar{t}}_{t=\underline{t}}$, and we choose a price of emissions $P^e(\underline{t}, \bar{t})$ so that $Z_{\bar{t}} = Z^o_{\bar{t}}$.[16] We use $\underline{t} = 1995$, the initial date for our price data and $\bar{t} = 2008$ the announcement of the Amazon fund that would pay for preservation projects in the Amazon, using funds contributed mostly by Norway and set $P^{ee} \stackrel{\text{def}}{=} P^e(1995, 2008)$.

It is not surprising that the price P^{ee} that matches observed deforestation varies with the model chosen. A model where, implicitly, a planner would act more aggressively against preservation would imply a larger P^{ee}. Note however that the larger P^{ee} when applied to future decisions would lower deforestation (or increase reforestation). Thus making P^{ee} vary with the model would bring future trajectories across models closer.[17]

[14] This follows a common practice for robust Bayesian methods.

[15] In fact, we apply the Hamiltonian Monte Carlo (HMC) sampling algorithm over the space of coefficients of the Municipal regressions. See Appendix C.

[16] We obtain a similar value if instead we minimize the norm of the vector $(\frac{X^o_{\bar{t}} - X^o_{\underline{t}}}{X^o_{\underline{t}}}, \frac{Z^o_{\bar{t}} - Z^o_{\underline{t}}}{Z^o_{\underline{t}}})$.

[17] The same reasoning holds across possible discount rates. Since emissions are a low duration asset relative to cattle raising, a larger discount rate would imply less deforesting for cattle thus lowering $P^e e$. Numerical simulations show that future trajectories at each level of transfers per unit of carbon captured do not change much when we move from 2 to 3%. For this reason, we only present results for a 2% discount rate.

6.2 Results for the deterministic case

In this section we discuss results for a model with a constant price for cattle that equals the stationary price associated with the estimated 2-state Markov chain ($42.03). We first discuss results for 78 sites that can be compared to the other, more computationally demanding, specifications. The shadow price is $P^{ee} = 7.5$, in this case. We show results for simulations for a period of 50 years and $P^e = P^{ee} + \$b$ with transfers $b = 0, 10, 15, 20$, and 25 per ton of *net* captured emission - that is if net emissions total E tons of CO_2, the planner receives bE.

As Figure 5 shows, with "business as usual" ($P^e = P^{ee}$), the optimal choice involves an increase in the agricultural area from 15% to more than 25% of the Biome, a fraction that may result on tipping of the eastern, southern and central Amazon to a non-forest system (Lovejoy and Nobre (2018)). The trajectories predicted are much different starting with an additional payment to the planner of $10, $15, $20 or $25 , leading to substantial decreases in agricultural area and corresponding increases in carbon captured.

The first five rows of Table 1 show the discounted value to the planner of a commitment to receive b of net transfers for each ton captured (emitted) of CO_2, when P^a is the stationary price. Here and in what follows, "forest services" are measured at the implied Brazilian shadow price. Even transfers of $10 are enough to compensate the losses of agricultural output, but the largest contributor to the gains is the increase in forest services using our estimates for the implicit valuation by Brazilians. However transfers based on $25 per ton of net captured CO_2, which is low when compared to the prevailing price of emissions in some regions of the globe, would more than double the value for the planner - a net gain of $230 billion. This net gain is composed of a loss of $317 billion in the value of cattle output,[18] which is fully compensated by $345 billion in transfers, and $211 billion in forest services. For robustness, we also produce results for an optimistic P^a, the highest price in our 2-state Markov chain. In this case transfers of $10/ton are not enough to compensate losses in income from agriculture but transfers based on $15/ton are. The net gain when transfers equal $25/ton changes little.

Table 2 displays the total effect of transfers per ton of net CO_2 captured (b), conditional on the price of cattle, on captured emissions in the first 30 years. When $P^a = \$42.03$, the stationary price, in the business a usual case, deforestation would produce carbon emissions of 13.7 billion tons in the next 30 years. With transfers of $25/ton, optimal management induces capture of 15 billion tons. Transfers total $375 billion. The effective cost per ton of the change in $C0_2$ emissions (28.7 billion tons of CO_2) is only $13.06/ton. The leverage from transfers goes up as b decreases since the business as usual emissions are a constant. For instance with transfers of $20/ton, which as can be seen from Table 1 still produces a 75% increase in the value to the planner, the effective cost $10.34/ton, approximately one-half of the nominal payment per ton. These results in Table 2 illustrate the gains from trade in instituting a contract that pays Brazil per net ton of CO_2 captured.

Figure 6 exhibits the initial distribution of occupation, and the distribution after 50 years for

[18] Recall however that we count the full output as value added, thus exaggerating the loss of agricultural output.

transfers per ton = \$0, \$10, \$15 and \$20. For reference, this figure again shows the cross-sectional parameter heterogeneity. Specifically, Figure 6 shows that for the case of transfers that exceed \$15 per ton of net emissions, the area of the Biome that is occupied by cattle farming after 50 years would be minimal.

Figure 7 illustrates, that for transfers exceeding \$15/ton, much of the change in land occupation would occur within the first 15 years. Of course, if adjustment costs are higher than we estimated (see Appendix A), this process would slow down.

Figures 8-9 and Tables 3-4 displays characteristics of the optimal solution when we consider 1043 sites. In this case the shadow price is a bit higher \$8/ton, but the planner values achieved are quite similar. In the "business as usual" case ($b = 0$) the amount of emissions would be higher than in the 78-site case, presumably because there are subsites with a higher θ that, in contrast with the original site in the 78 grid, would be deforested. This results in a 5% lower effective cost per ton in the case of $b = 25$, but this effective cost is still 9.25/ton. Finally as Figure 9 shows when $b = 15$, in contrast with the case of 78 sites, there is some residual areas in the east and south that retain, even after 50 years, some cattle activity. This is again the result of working with a grid that allows for decisions at a finer scale.

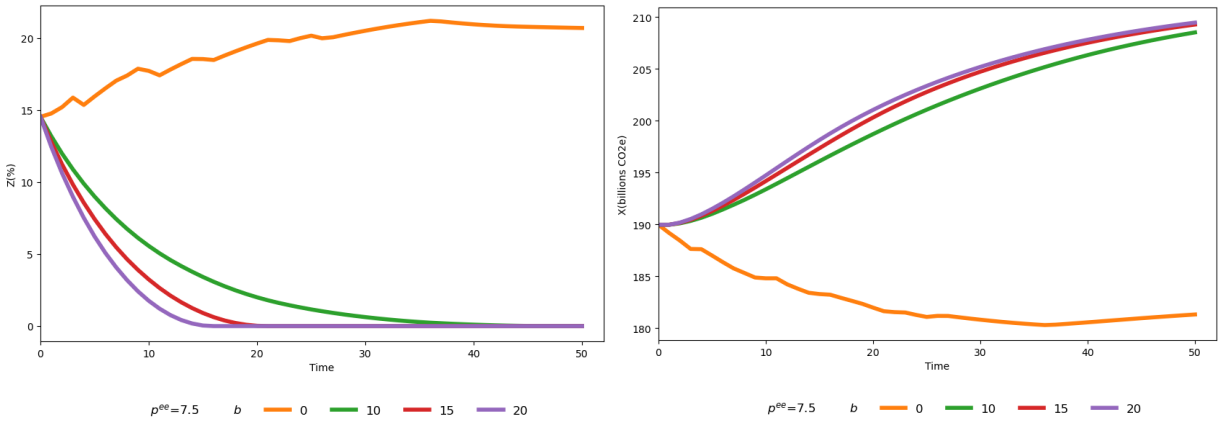


Figure 5: Agricultural Area and Carbon Stock Evolution (50 years) - deterministic case 78 sites

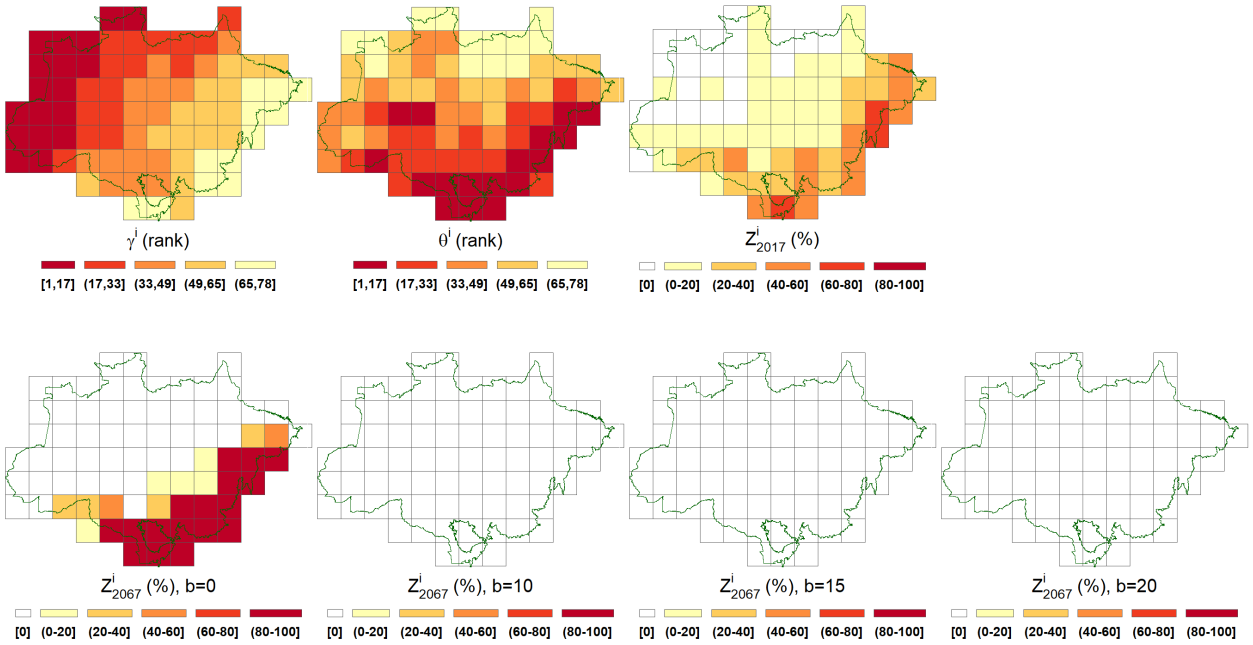


Figure 6: Agricultura Area Evolution by Agricultural Price - deterministic case 78 sites

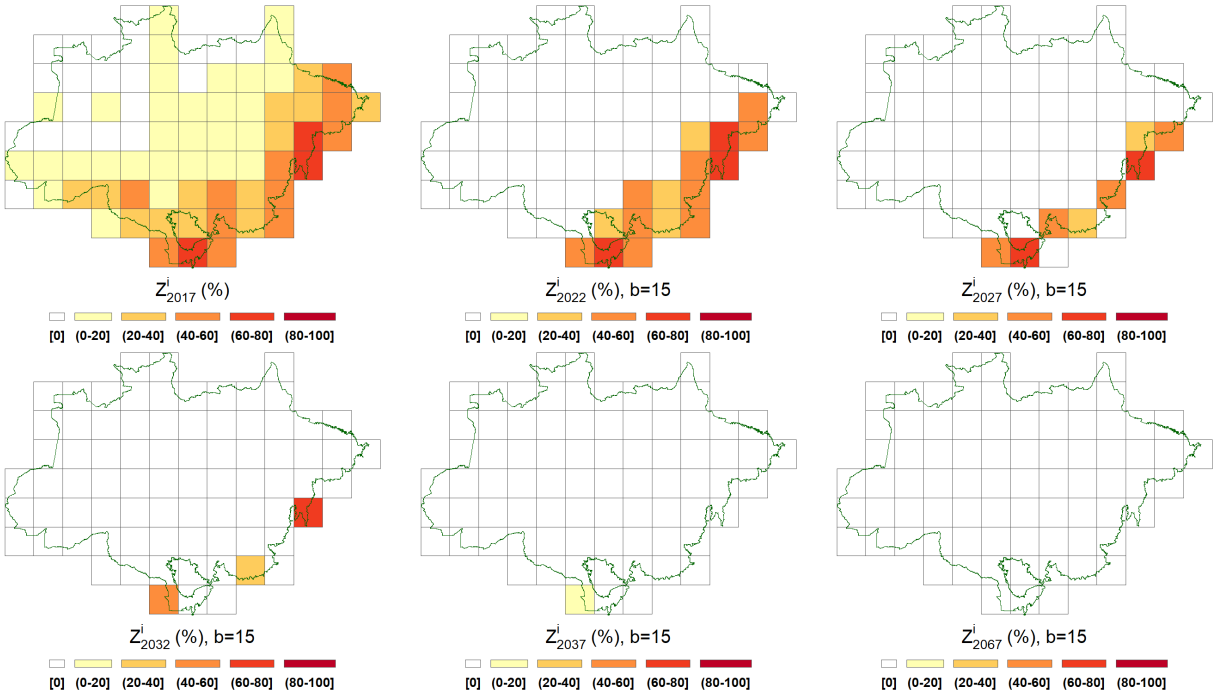


Figure 7: Agricultura Area Evolution by Time - deterministic case 78 sites

Table 1: Planner Value Decomposition (200 years) - deterministic case 78 sites

P^a (\$)	P^e (\$)	b (\$)	Agricultural Output (\$ 10^{11})	Net Transfers (\$ 10^{11})	Forest Services (\$ 10^{11})	Adjustment Costs (\$ 10^{11})	Planner Value (\$ 10^{11})
42.03	7.5	0	3.24	0.00	-1.08	0.05	2.10
42.03	17.5	10	0.40	1.21	0.90	0.11	2.41
42.03	22.5	15	0.26	1.97	0.98	0.17	3.04
42.03	27.5	20	0.20	2.71	1.01	0.22	3.71
42.03	32.5	25	0.17	3.45	1.03	0.26	4.40
44.76	7.5	0	3.54	0.00	-1.16	0.06	2.31
44.76	17.5	10	0.58	1.12	0.84	0.10	2.44
44.76	22.5	15	0.29	1.96	0.98	0.17	3.06
44.76	27.5	20	0.22	2.71	1.01	0.22	3.73
44.76	32.5	25	0.19	3.44	1.03	0.25	4.41

Notes: For P^a, 42.03 is the stationary price and 44.76 is the high price (75th percentile of the series). Climate services are calculated using baseline shadow price ($b = 0$).

Table 2: Transfer Cost (30 years) - deterministic case 78 sites

P^a (\$)	P^e (\$)	b (\$)	Net Captured Emissions (billion tons of CO2e)	Net Transfers (\$ 10^{11})	Effective cost (\$ per ton of CO2e)
42.03	7.5	0	-13.70	0.00	NaN
42.03	17.5	10	12.03	1.20	4.67
42.03	22.5	15	14.07	2.11	7.60
42.03	27.5	20	14.69	2.93	10.34
42.03	32.5	25	15.00	3.75	13.06

Notes: For P^a, 42.03 is the stationary price.

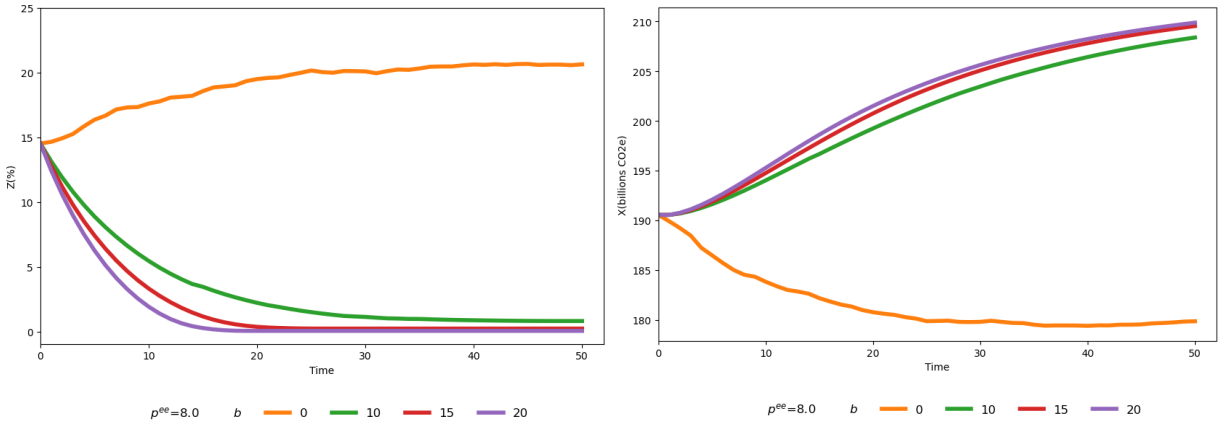


Figure 8: Agricultural Area and Carbon Stock Evolution (50 years) - deterministic case 1043 sites

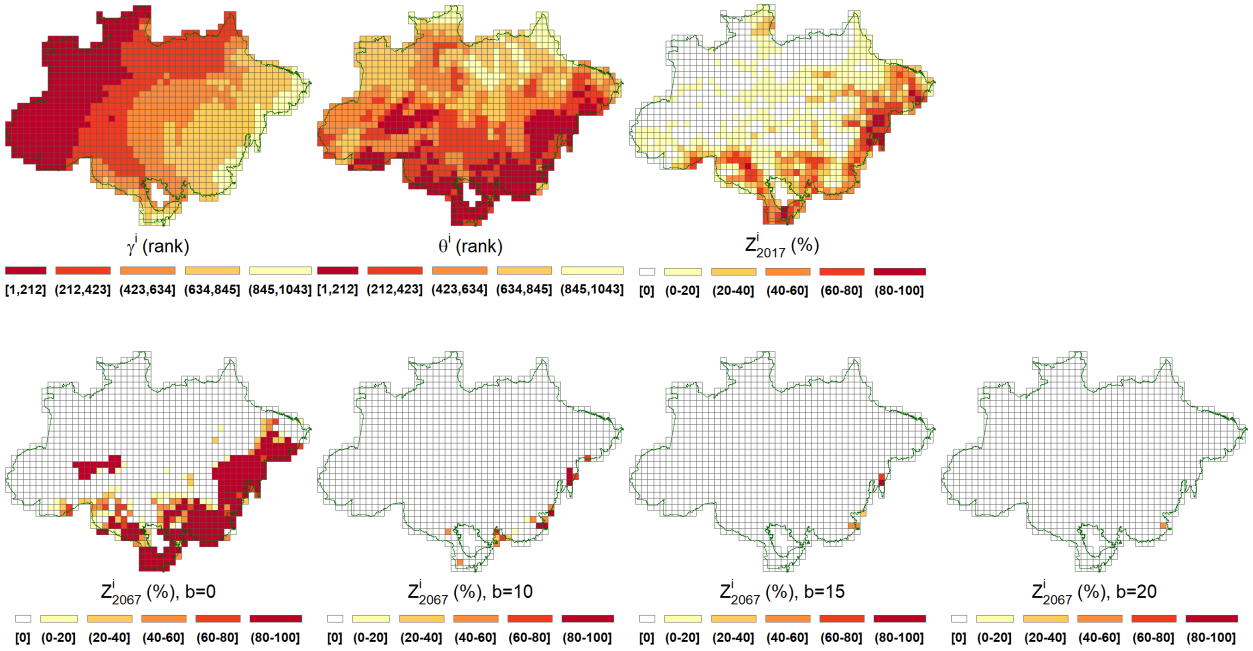


Figure 9: Agricultura Area Evolution by Agricultural Price - deterministic case 1043 sites

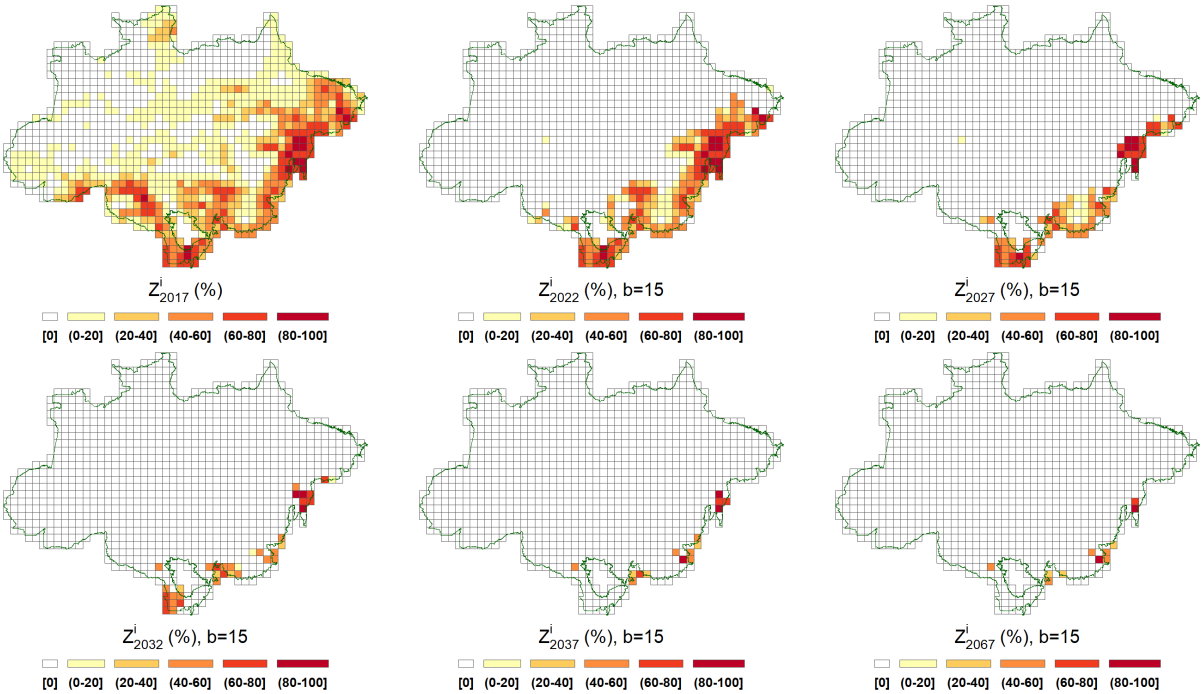


Figure 10: Agricultural Area Evolution by Time - deterministic case 1043 sites

Table 3: Planner Value Decomposition (200 years) - deterministic case 1043 sites

P^a (\$)	P^e (\$)	b (\$)	Agricultural Output (\$ 10^{11})	Net Transfers (\$ 10^{11})	Forest Services (\$ 10^{11})	Adjustment Costs (\$ 10^{11})	Planner Value (\$ 10^{11})
42.03	8.0	0	3.53	0.00	-1.25	0.07	2.20
42.03	18.0	10	0.55	1.15	0.92	0.11	2.50
42.03	23.0	15	0.32	1.93	1.03	0.17	3.11
42.03	28.0	20	0.23	2.68	1.07	0.22	3.77
42.03	33.0	25	0.18	3.43	1.09	0.26	4.45
44.76	8.0	0	4.05	0.00	-1.51	0.08	2.45
44.76	18.0	10	0.69	1.08	0.86	0.10	2.54
44.76	23.0	15	0.38	1.90	1.01	0.16	3.13
44.76	28.0	20	0.26	2.67	1.07	0.21	3.79
44.76	33.0	25	0.20	3.42	1.09	0.25	4.46

Notes: For P^a, 42.03 is the stationary price and 44.76 is the high price (75th percentile of the series). Forest services are calculated using baseline shadow price ($b = 0$).

Table 4: Transfer Cost (30 years) - deterministic case 1043 sites

P^a (\$)	P^e (\$)	b (\$)	Net Captured Emissions (billion tons of CO2e)	Net Transfers (\$ 10^{11})	Effective cost (\$ per ton of CO2e)
42.03	8.0	0	-15.28	0.00	NaN
42.03	18.0	10	11.77	1.17	4.35
42.03	23.0	15	13.78	2.06	7.11
42.03	28.0	20	14.51	2.90	9.74
42.03	33.0	25	14.89	3.72	12.33

Notes: For P^a, 42.03 is the stationary price.

6.3 Results for the case of agricultural price uncertainty

In this section we summarize results for the case when cattle prices, P^a are generated by the calibrated Markov chain. We again focus on the grid of 78 sites. Under uncertainty, trajectories depend on the realization of agricultural prices so that our results involving a single trajectory should be taken as indicative. Given the historical prince experience during 1995-2008, the calibrated shadow price is now P^{ee} = USD 7, lower than the shadow price obtained in the deterministic case with the stationary price. Notice that the picture is much like in the previous case. As the trajectories in Figure 11 indicate, in the absence of transfers, the optimal choice again involves an increase in the agricultural area that may result on tipping of the eastern, southern and central Amazone, whereas transfers of at least \$10 per net ton, would lead to almost complete reforestation. When transfers change from \$0 to \$25 per ton, the present value over the next 200 years of cattle output is reduced from \$312 billion to \$16 billion, a loss that is more than compensated by (i) the increase in the

present value of net transfers from \$0 to \$346 billion, and (ii) the increase in forest services as valued by Brazilians (at the shadow price of \$7 per ton of net emissions captured) from the change in policy induced by the \$25 per ton transfers, which totals \$199 billion. The effective cost for emissions are quite close to the effective cost in the case of no uncertainty and cattle prices given by the stationary price associated with our estimated Markov chain. Figure 12 shows the pattern of concentration is similar to the case of no uncertainty. Figure 13 illustrates the spatial dynamics by showing that with the level of adjustment costs we imposed, much of the geographical change in land use occurs within the first ten years, and most of it in the first 20 years. Finally, Figure 14 illustrates the effect of the planner internalizing the potential for forest growth of the different sites. After 30 years the planner would completely reforested the site with the highest θ (in red) but would keep maintain the level of deforestation (47%) of the site with the second highest θ, because this site (in red) is less productive in carbon capture.

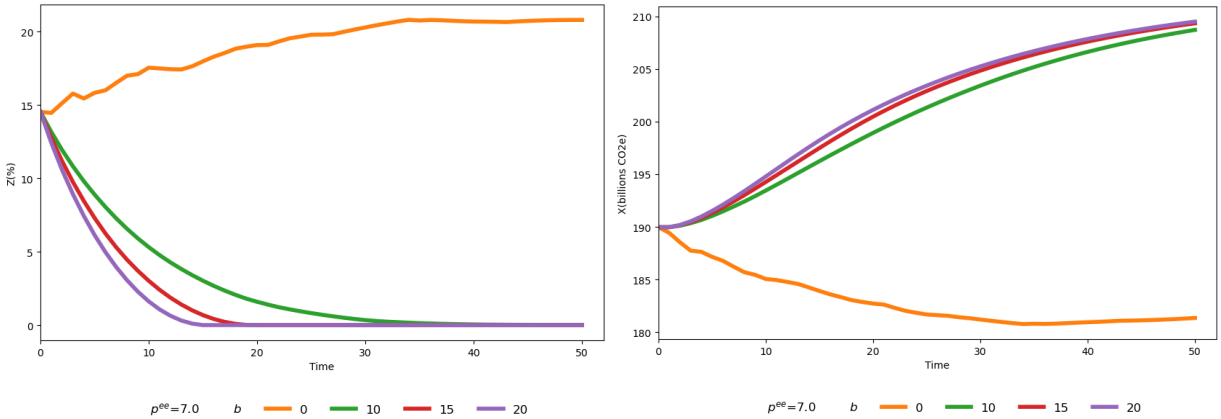


Figure 11: Agricultural Area and Carbon Stock Evolution (50 years) - MPC case 78 sites

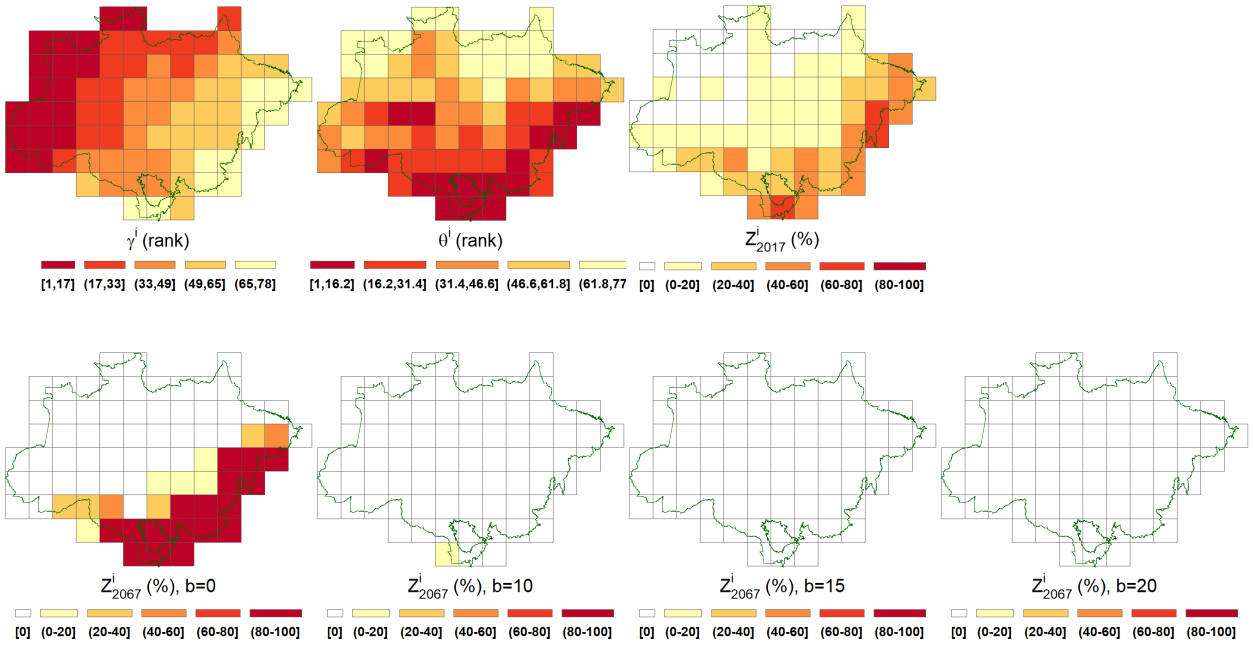


Figure 12: Agricultura Area Evolution by Agricultural Price - MPC case 78 sites

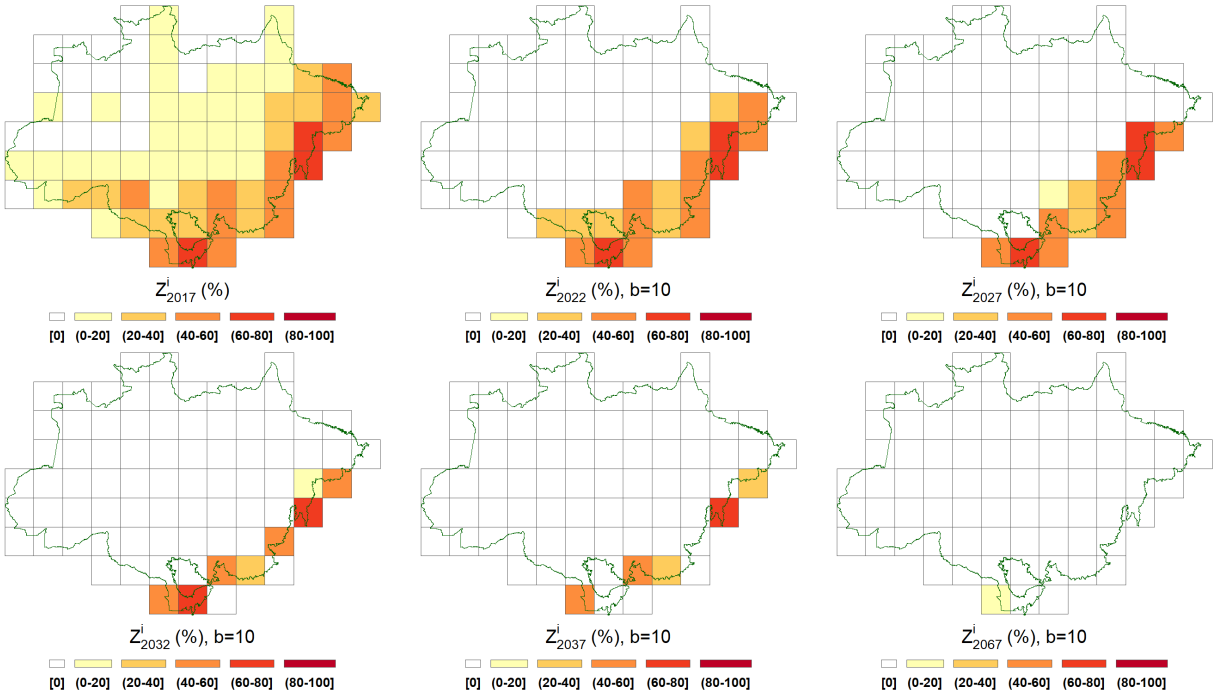


Figure 13: Agricultura Area Evolution by Time - MPC case 78 sites

Table 5: Planner Value Decomposition (200 years) - MPC case 78 sites

P^e (\$)	b (\$)	Agricultural Output (\$ 10^{11})	Net Transfers (\$ 10^{11})	Forest Services (\$ 10^{11})	Adjustment Costs (\$ 10^{11})	Planner Value (\$ 10^{11})
7.0	0	3.12	0.00	-1.02	0.05	2.05
17.0	10	0.36	1.22	0.86	0.12	2.33
22.0	15	0.24	1.98	0.92	0.18	2.97
27.0	20	0.19	2.72	0.95	0.23	3.64
32.0	25	0.16	3.46	0.97	0.26	4.32

Notes: Forest services calculated using baseline shadow price ($b = 0$).

Table 6: Transfer Cost (30 years) - MPC case 78 sites

P^e (\$)	b (\$)	Net Captured Emissions (billion tons of CO2e)	Net Transfers (\$ 10^{11})	Effective cost (\$ per ton of CO2e)
7.0	0	-13.22	0.00	NaN
17.0	10	12.39	1.23	4.83
22.0	15	14.19	2.12	7.76
27.0	20	14.74	2.94	10.54
32.0	25	15.03	3.75	13.30

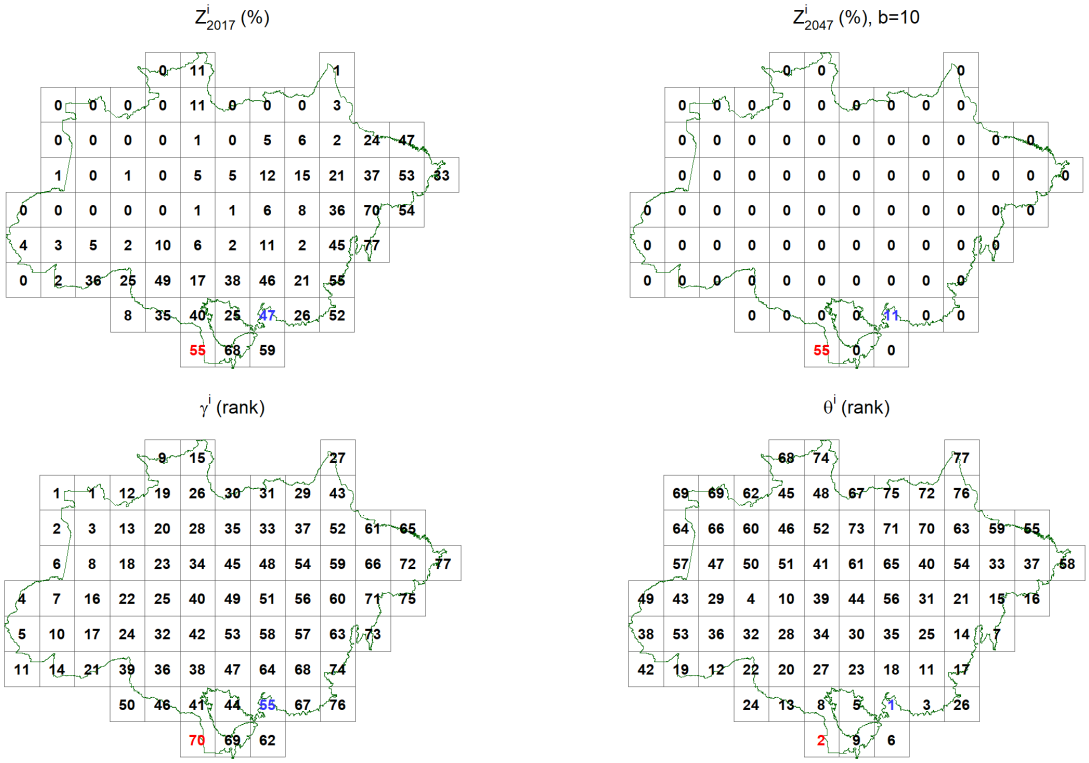


Figure 14: Share of agricultural area by site and ranking of parameters by site, MPC case 78 sites

6.4 Results with robustness to parameter uncertainty

In this section we present results when the planner is uncertain about cattle productivity and CO_2. capture capacity of the 78 sites, for a planner that assumes that P^a is the stationary price associated with our estimated Markov chain and has $\xi = 2$. Ther calculated "shadow price" in this case is \$6.5/ton.

Figure 15 shows adjusted and unadjusted distributions for parameters γ and θ when $b = 25$, for a few sites. Notice that the (negative) adjustment on γ are proportionally much larger then the θ adjustments. This result is related to the fact that, with $b = 25$, at the baseline distribution (no ambiguity) most of the forest would be restored.

Figure 16 shows that, when $b = 25$, ambiguity has a very minor effect on the aggregate trajectory. However as Figure 17 ambiguity does affect the order in which sites are reforested. Ambiguity postpones the start of reforestation of the sites colored red but reduces the start of reforestation of the sites colored blue. Comparing the value to the planner in the HMC case, Table 7, to the case with no uncertainty with the same number of sites and P^a, Table 1, we find that before accounting for the uncertainty costs given by the relative entropy entropy term, there is a minor drop ($<$ 10%) for all level of transfers. Although relative to planner's value uncertainty costs are small for most level of transfers, they are more substantial for the no transfer case, increasing the benefit of transfers to the planner. The slight increase in emissions in the baseline case ($b = 0$) results in

a lower effective cost per ton but the slower reforestation in the first years when ($b = 25$) implies lower net captured emissions, compared to the case of no uncertainty (Table 2)

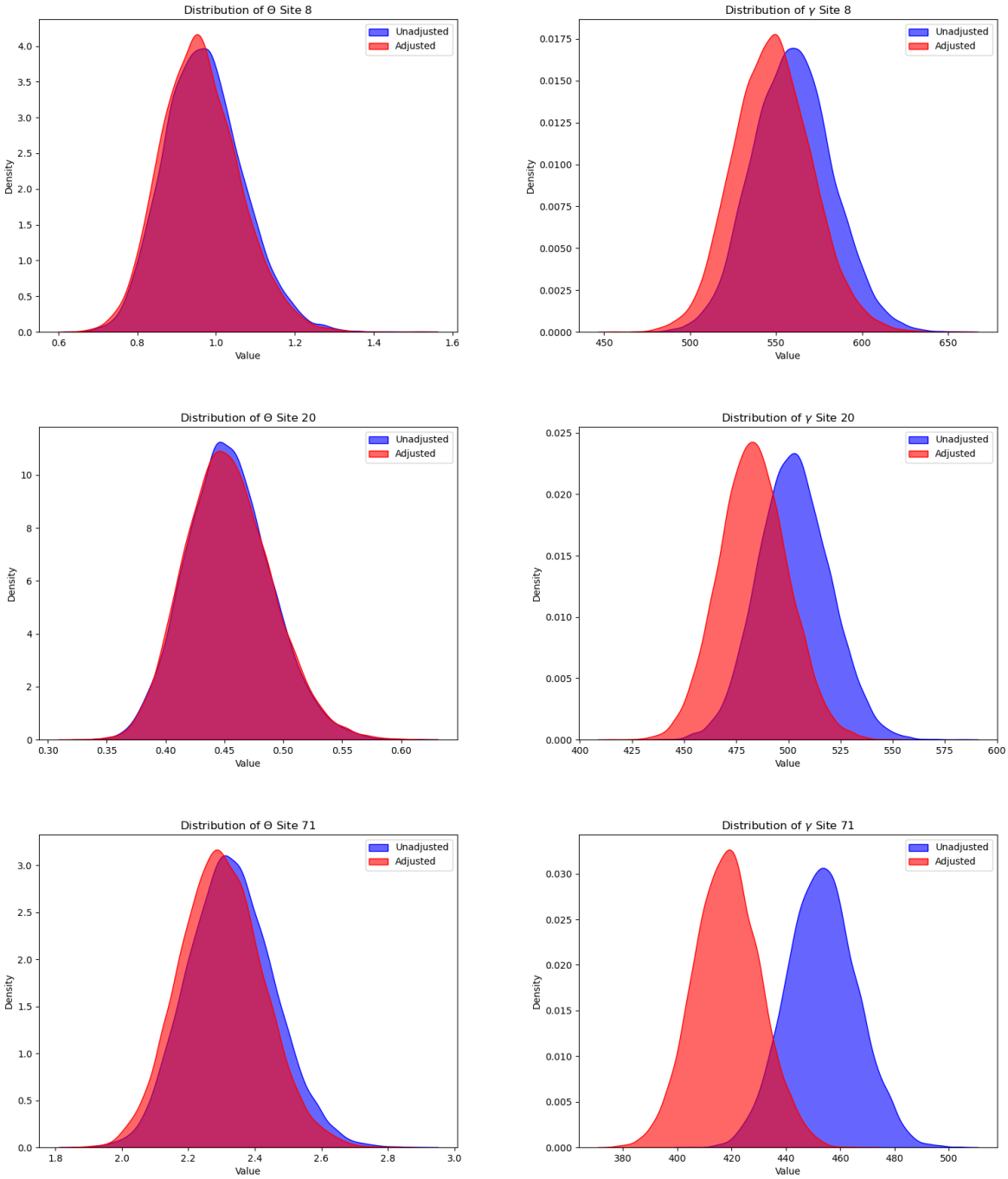


Figure 15: Ambiguity adjustment, $b = 25$

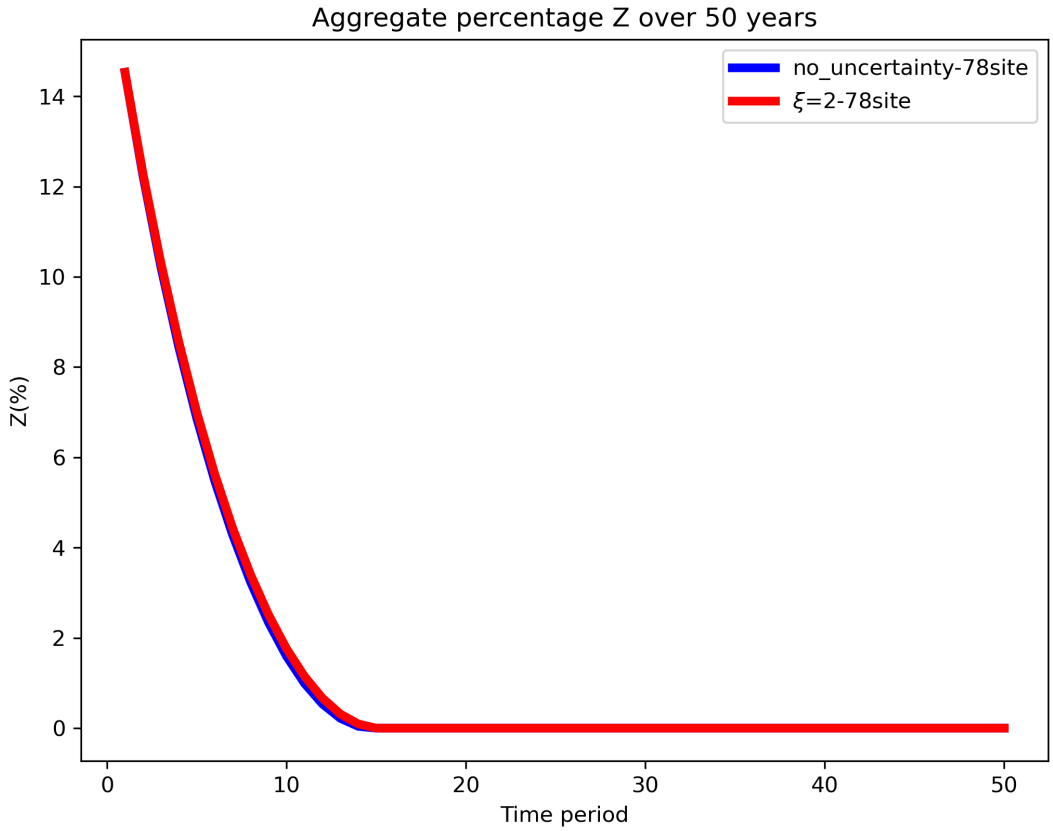
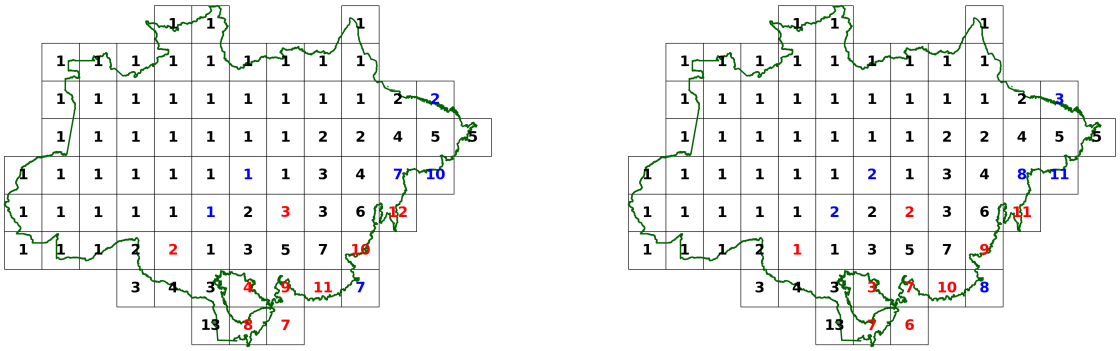


Figure 16: Evolution of agricultural area, $b = 25$



(a) ambiguity

(b) no ambiguity

Figure 17: Year in which reforestation in site i begins, $b = 25$

Table 7: Planner Value Decomposition (200 years) - HMC case 78 sites

ξ	P^e (\$)	b (\$)	Agricultural Output (\$ 10^{11})	Net Transfers (\$ 10^{11})	Forest Services (\$ 10^{11})	Adjustment Costs (\$ 10^{11})	Planner Value (\$ 10^{11})	Uncertainty Costs (\$ 10^{11})
2	6.5	0	2.83	0.00	-0.95	0.05	1.82	0.45
2	16.5	10	0.50	1.11	0.72	0.10	2.23	0.06
2	21.5	15	0.28	1.87	0.81	0.16	2.81	0.11
2	26.5	20	0.21	2.57	0.83	0.21	3.41	0.16
2	31.5	25	0.18	3.25	0.84	0.25	4.02	0.23

Notes: $P^a = 42.03$, the stationary price. Forest services are calculated using baseline shadow price ($b = 0$).

Table 8: Transfer Cost (30 years) - HMC case 78 sites

ξ	P^e (\$)	b (\$)	Net Captured Emissions (billion tons of CO2e)	Net Transfers (\$ 10^{11})	Effective cost (\$ per ton of CO2e)
2	6.5	0	-13.74	0.00	NaN
2	16.5	10	10.95	1.09	4.43
2	21.5	15	13.29	1.99	7.37
2	26.5	20	13.88	2.77	10.05
2	31.5	25	14.09	3.52	12.65

Notes: $P^a = 42.03$, the stationary price.

7 Conclusions

We used a rich data set to study the impact of carbon prices on optimal forest preservation over time and space in the Brazilian Amazon. We produced results for three exercises. First we calculated the shadow prices for emissions that would justify historical deforestation. We then calculated the impact of payments for net emissions captured. The optimal strategy under the calculated shadow price would eventually produce enough deforestation to threaten the survival of the Amazon as a tropical forest. However transfer prices above \$15 per ton of $C0_2$e, increase Brazil's welfare and lead to substantial net reforestation and carbon capture in the optimal trajectory. Furthermore, carbon prices above \$25 would generate enough monetary transfers to fully compensate Brazil for the loss of output in the re3forested areas. As a third application, we considered optimal robust controls when the capacity of each site to capture carbon and and the corresponding agricultural productivity is uncertain to the planner.

Our results suggest international carbon payments of \$25 USD/ton can reduce emissions by 27 to 30 GtCO2e, in 30 years. This amount represents not only the total GtCO2e of carbon captured by natural regeneration, for which Brazil will receive the payments, but the avoided emissions from the deforestation that could happen in the 'business as usual' scenario. This combination implies a leveraged carbon mechanism for which the effective carbon cost around \$13 USD/ton.

According to Griscom et al. (2017), nature-based solutions such as forest restoration, avoided land conversion, forest management and other practices have the potential of capturing about 11.3 Gt of CO2 per year globally with costs no greater than \$100 USD/ton. Our baseline simulation in

2 suggests that optimal management of the Brazilian Amazon can deliver at least 9% of this total at an effective cost that would not exceed $12.5 USD/ton.

Our calculations in this paper ignore some important costs of deforestation. We do not include, for instance, the effect of deforestation on agricultural productivity in the Amazon (Leite-Filho et al. (2021)) or in other regions in Brazil, a country that is currently the fourth largest agricultural producer and exporter in the world. We also do not take into account the loss of biodiversity or the possible role of Amazon deforestation as a global tipping point (Steffen et al. (2018)). Finally we do not account for the direct effect of deforestation in one site on forests in other sites.[19]

These are important considerations for future research.

[19]SeeAraujo et al. (2023) for an estimate

A Data construction

Total available area: $\bar{z}^i$

To compute $\bar{z}^i$, the amount of available area for the planner's choice (forest or cattle farming) in each site i, we first calculate the fraction of 30m-pixels in site i classified as agriculture (crops + pastures) or forests in MapBiomas 2017 (Souza Jr et al., 2020). We then multiply this fraction by the area (within the biome) of the site, to obtain a measure in hectares. Notice $\bar{z}^i$ comprises the total site area, excluding areas such as rivers, roads, cities and etc.

Parameters related to carbon dynamics: γ^i, α, and κ

We first extract a random sample of 1.2M 30m-pixels and select 893,753 pixels with no deforestation during 1985-2017, which we treat as primary forests as of 2017. We add *above ground* biomass density data for the year 2017 from ESA Biomass[20]. The biomass data also comes in a grid format with ∼100m resolution, so we spatially match it to our sample. The original data is measured in biomass density (Mg per ha) but we convert it to carbon per hectare, by dividing by 2 (carbon is approximately 50% of the biomass), and then obtain CO2 equivalent by multiplying by 44 and dividing by 12 (based on atomic mass). We calculate average of CO2 density (MG/ha) for each municipality and run the following regression

$$\begin{aligned}\log(\text{co2e\_ha}) = \beta_0^\gamma + \beta_1^\gamma(\log(\text{historical\_precip})) + \\ \beta_2^\gamma(\log(\text{historical\_temp}))+ \\ \beta_3^\gamma(\log(\text{lat})) + \beta_4^\gamma(\log(\text{lon})) \ + \varepsilon_\gamma\end{aligned}$$

Table 9 below summarizes results from this regression

Table 9: Regression for γ

	log(co2e_ha)
	(1)
Intercept	−7.63
	(5.03)
log(precip)	0.43***
	(0.10)
log(temp)	−4.56***
	(0.65)
log(lat)	3.20***
	(0.41)
log(lon)	−1.64***
	(0.10)
R^2	0.51
Adj. R^2	0.51
Num. obs.	538

Notes: ***$p < 0.001$; **$p < 0.01$; *$p < 0.05$, inside parentheses are standard errors.

[20] (Santoro and Cartus, 2021)

For each municipality m, we consider the fitted value $log(\widehat{\text{co2e\_ha}})_m$ and take the exponential to get $\widehat{\text{co2e\_ha}}_m$. Finally, γ^i for each site i is $\gamma^i = \sum_{m \in i} w_{i,m} \widehat{\text{co2e\_ha}}_m$, with the weights given by $w_{i,m} := \frac{g_{i,m}}{\sum_{m \in i} g_{i,m}}$, where $g_{i,m}$ is the overlapped area between municipality m and site i. The parameter α is a carbon depreciation parameter, assumed to be constant across sites. It is set so that the 99% convergence time of the carbon accumulation process is 100 years (see Heinrich et al. (2021)), that is $\alpha = 1 - (1 - 0.99)^{1/100}$.

Finally, the parameter κ is calibrated using the agricultural net annual emission data at the state level available from the system SEEG.[21] We compute κ as the average of agricultural net emission divided by the agricultural area from MapBiomas for all states within the Amazon biome, weighting by the area of each state overlap with the biome, from 1990 to 2019.

Cattle farming productivity: θ^i

Since almost 90% of the historically deforested land in the Amazon Biome that was used for agricultural activities in 2017 was used for pasture, we focus on the productivity of cattle farming for each site. Since we do not have measurements concerning the cost of attracting or redeploying variable inputs to the cattle farming sector, we focus on revenue per hectare. This choice leads to an overvaluation of the contribution of cattle farming in the Amazon to the Brazilian economy.[22] We consider the value of cattle sold for slaughter per hectare of pastureland at the municipal level, from the 2017 Agricultural Census (IBGE, 2017). In order to build a smoother representation of technology, and fill missing values, we used the predicted values of the following specification;

$$\begin{aligned} \log(\text{Slaughter\_value}) = \beta_\theta^0 \quad &+ \quad \beta_\theta^1(\text{historical\_precip}) \ + \\ \beta_\theta^2(\text{historical\_temp}) \quad &+ \quad \beta_\theta^3(\text{historical\_temp}^2) \ + \\ \beta_\theta^4(\text{lat}) + \beta_\theta^5(\text{lat}^2) \quad &+ \\ \beta_\theta^6 \log(\text{cattleSlaughter\_farmGatePrice}) \quad &+ \quad \beta_\theta^7(\text{distance}) \ + \varepsilon_\theta, \end{aligned} \tag{8}$$

where slaughter value is the value of cattle sold per hectare of pasture area in 2017 (USD/ha), precipitation and temperature are the average annual precipitation (mm) and temperature (degrees Celsius), respectively, for the period of 1970-2000 (Fick and Hijmans, 2017), latitude is the geographical coordinates of the municipality centroids, farm gate price is the price of cattle slaughter (SEAB-PR, 2021), and distance is measured the distance from the municipality to the state capital. Since the area dedicated to agriculture varies substantially across municipalities, we opted for weighted least squares estimation with a diagonal matrix W_θ where W_θ^{ii} is equal to the 2017 pasture area in each municipality i; that is, we assume that the variance of ε_θ is $\sigma_\theta^2 W_\theta^{-1}$, where σ_θ^2 is a scalar. Table 10 below summarizes results from this regression:

The inclusion of farm gate prices on the right hand side of this regression is reasonable, because

[21] Sistema de Estimativas de Emissões e Remoções de Gases de Efeito Estufa. Available in http://seeg.eco.br/.

[22] In contrast to other areas in Brazil, average value of slaughter per hectare of pasture in the Amazon, decreased between 2006 and 2017, making it doubtful that future productivity will increase

Table 10: Regression for θ

	log(Cattle_Slaughter)
	(1)
Intercept	−111.55***
	(23.35)
precip	−4.32e−3**
	(1.33e−3)
temp	3.94*
	(1.85)
temp^2	−0.08*
	(0.04)
lat	1.30e−5***
	(1.76e−6)
lat^2	−7.16e−13***
	(9.63e−14)
distance	−3.65e−4**
	(1.27e−4)
log(FarmGatePrice)	2.59***
	(0.17)
R^2	0.53
Adj. R^2	0.52
Num. obs.	466

Notes: ***$p < 0.001$; **$p < 0.01$; *$p < 0.05$, inside parentheses are standard errors.

variations in farm gate prices across municipalities mostly reflect unobserved costs to bring cattle to stockyards and meat to markets such as proximity to roads or rivers, which are not fully controlled by our geographical variables.

To obtain site slaughter value for sites we use

$$\widehat{\text{Slaughter\_value}}^i = \sum_{m \in i} w_{i,m} \exp(\log\widehat{\text{Slaughter\_value}}_m)$$

with the weights $w_{i,m} := \frac{g_{im}}{\sum_{m \in i} g_{i,m}}$, where $g_{i,m}$ is the area overlap of municipality m in site i. Finally, we obtain θ^i by dividing $\widehat{\text{Slaughter\_value}}^i$ by by the observed ($\bar{P}^a_{2017}$) (expressed in 01/2017 USD)

Discount rate (δ) and adjustment costs cost (ζ)

We use the discount rate $\delta = 0.02$ and calibrate ζ using the difference in price between forested land and cleaned land and the amount of annual deforestation that occurred from 2008 to 2017 based on Araujo et al. (2022). Notice that the difference in price should reflect both the cost of deforestation and any value of wood obtained in the process. Unfortunately we did not have data that would allow us to compute a separate adjustment cost for decreasing (as opposed to increasing) deforestation, so we opted for symmetry.

Initial values: z^i_0, x^i_0

The approach for computing the initial value for the agricultural area, z^i_0, is similar to that used for the total available area $\bar{z}^i$. The only difference is that we focus only on the fraction of pixels

classified as agriculture (crops + pastures) in 2017 before multiplying by the site's area in order to obtain a measure in hectares.

The initial value for the carbon stored in the forests x_0^i is calibrated as $x_0^i = \gamma^i(\bar{z}^i - z_0^i)$, i.e., the carbon stock per hectare of forest times the forest area. Notice that x_0^i is measured in CO2e (Mg). Notice that we assume that all forest at the initial point is primary, which is compatible with equation (2).

Agricultural prices

We use a data series on monthly deflated cattle prices (reference date 01/2017).[23] from 1995, the year in which the Real Plan stabilized the Brazilian currency, until 2017. In this paper, we fit a two-state Markov process. Let $\tilde{P}_t^a$ denote observed prices. We calculated a high (low) price $P^{a,h}$ ($P^{a,\ell}$) as the 75th (25th) percentile of the series $\tilde{P}_t^a$. We discretized the prices to $P^{a,h}$ and $P^{a,\ell}$ based on the following criteria. If the initial price is above (below) the median, we start the series at $P^{a,h}$ ($P^{a,\ell}$). If at time t, $P_t^a = P^{a,h}$ then $P_{t+1}^a = P^{a,h}$ unless $\tilde{P}_{t+1}^a \leqslant (P^{a,\ell}$. Symmetrically, if at time t, $P_t^a = P^{a,\ell}$ then $P_{t+1}^a = P^{a,\ell}$ unless $\tilde{P}_{t+1}^a \geqslant P^{a,h}$. We calculate the discrete transition probabilities using observed frequencies. The probability from high to high, for example, is defined as the number of consecutive periods with high prices divided by the number of periods with high prices.

B Baseline distribution

To characterize the baseline distribution $\pi(\boldsymbol{\rho})$ we consider $\pi(\beta_\theta, \sigma_\theta^2)$ and $\pi(\beta_\gamma, \sigma_\gamma^2)$ separately; below we present the derivation of $\pi(\beta_\theta, \sigma_\theta^2)$, with the derivation of $\pi(\beta_\gamma, \sigma_\gamma^2)$ following analogously, except that we assume $W_\gamma = I$.

Rewriting the linear regression model at the municipality level, given by 8 and the matrix W_θ, and assuming a Gaussian error term,

$$y_\theta \;=\; X_\theta \beta_\theta + \varepsilon_\theta, \quad \varepsilon_\theta \sim \mathcal{N}(0,\, \sigma_\theta^2 W_\theta^{-1}) \tag{9}$$

We assume that the priors for parameters $(\beta_\theta, \sigma_\theta^2)$ are:

$$\beta_\theta \mid \sigma_\theta^2 \;\sim\; \mathcal{N}(m_0,\, \sigma_\theta^2 Q_0^{-1}) \tag{10}$$

$$\sigma_\theta^2 \;\sim\; \text{Inv-Gamma}(a_0,\, b_0) \tag{11}$$

Because this prior is conjugate to the likelihood implied by equation 9, the posterior for $(\beta_\theta, \sigma_\theta^2)$ has the closed-form expression:

[23]Commodity prices from SEAB-PR. Secretaria da Agricultura e do Abastecimento do Estado do Paraná (SEAB-PR). 2021. "Preço Médio - Recebido pelo Agricultor: boi gordo, arroz (em casca), cana-de-açúcar, milho, mandioca, 1990-2021." Secretaria da Agricultura e do Abastecimento do Estado do Paraná, Departamento de Economia Rural [publisher], Instituto de Pesquisa Econômica Aplicada, Ministério da Economia [distributor]. http://www.ipeadata.gov.br (accessed February 22, 2021)

$$\beta_\theta \mid \sigma_\theta^2, Q, m \ \sim \ \mathcal{N}(m,\, \sigma_\theta^2 Q^{-1}) \tag{12}$$

$$\sigma_\theta^2 \mid a, b \ \sim \ \text{Inv-Gamma}(a,\, b) \tag{13}$$

with (Q, m, a, b) given by,

$$Q = X_\theta' W_\theta X_\theta + Q_0, \tag{14}$$

$$m = Q^{-1}(X_\theta' W_\theta y_\theta + Q_0 m_0), \tag{15}$$

$$a = a_0 + \frac{n}{2}, \tag{16}$$

$$b = b_0 + \frac{1}{2}(y_\theta' W_\theta y_\theta + m_0' Q_0 m_0 - m' Q m). \tag{17}$$

We set $Q_0 = 0_{k\times k}$, where k is the number of regressors, and $a_0 = b_0 = 0$, which corresponds to the limit of an (improper) uniform prior over $(\beta_\theta, \sigma_\theta^2)$. Together with equations (12 - 17), this defines our baseline density:

$$d\pi(\beta_\theta, \sigma_\theta^2 | X_\theta, y_\theta) = d\pi(\beta_\theta | \sigma_\theta^2, X_\theta, y_\theta) d\pi(\sigma_\theta^2 | X_\theta, y_\theta). \tag{18}$$

Equation 18 implies that the marginal distribution $\pi_{\boldsymbol{\beta}}$ is a t-distribution.

C Hamiltonian Monte Carlo

As is standard in MCMC we can ignore normalizations so we are interested in only on the numerator of 7. Although we could sample from the marginal distribution for β, Stan, the software we employ, actually recommends that, if users intend sampling from a t-distribution, to avoid fat tails, users should re-parameterize their model to sample from a joint normal-inverse gamma distribution instead and discard the extra parameters (see Stan Development Team, 2023). This means that we should sample from

$$\exp\left[-\frac{1}{\xi} f(\mathbf{d}, \boldsymbol{\beta})\right] d\pi(\beta_\theta, \sigma_\theta^2 | X_\theta, y_\theta) d\pi(\beta_\gamma, \sigma_\gamma^2 | X_\gamma, y_\gamma) \tag{19}$$

By taking logs, multiplying by -1, and substituting 18, we get the potential energy term $\mathcal{U}$:

$$\mathcal{U}(\boldsymbol{\rho}) = \frac{1}{\xi} f(\mathbf{d}, \boldsymbol{\beta}) - \log d\pi(\beta_\theta, \sigma_\theta^2 | X_\theta, y_\theta) - \log d\pi(\beta_\gamma, \sigma_\gamma^2 | X_\gamma, y_\gamma) \tag{20}$$

$$\begin{aligned} &= \frac{1}{\xi} f(\mathbf{d}, \boldsymbol{\beta}) - \log d\pi(\beta_\theta | \sigma_\theta^2, X_\theta, y_\theta) - \log d\pi(\beta_\gamma | \sigma_\gamma^2, X_\gamma, y_\gamma) \\ &\quad - \log d\pi(\sigma_\theta^2 | X_\theta, y_\theta) - \log d\pi(\sigma_\gamma^2 | X_\theta, y_\gamma) \end{aligned} \tag{21}$$

HMC relies on an auxiliary momentum vector ω of the same dimension as $\boldsymbol{\rho}$, where $\omega \sim \mathcal{N}(0, M)$ and M is a symmetric, positive-definite mass matrix. The Hamiltonian is then defined as:

$$\mathcal{H}(\boldsymbol{\rho}, \omega) := \mathcal{U}(\boldsymbol{\rho}) + \frac{1}{2} \omega' M^{-1} \omega \tag{22}$$

The HMC algorithm then consists of:

1. Initialize $\boldsymbol{\rho}_{(0)}$.
2. Sample momentum $\omega_{(0)} \sim N(0, M)$.
3. Generate a state proposal $(\tilde{\boldsymbol{\rho}}_{(0)}, \tilde{\omega}_{(0)})$ by evolving its position according to Hamilton's equations, using the leapfrog integrator with step size ϵ and a number of steps L:

$$\frac{d\boldsymbol{\rho}}{dt} = \frac{\partial \mathcal{H}}{\partial \omega} \tag{23}$$

$$\frac{d\omega}{dt} = -\frac{\partial \mathcal{H}}{\partial \boldsymbol{\rho}} \tag{24}$$

4. Perform a Metropolis test to accept or reject the state update $(\boldsymbol{\rho}_{(1)}, \omega_{(1)}) \leftarrow (\tilde{\boldsymbol{\rho}}_{(0)}, \tilde{\omega}_{(0)})$, with the acceptance probability given by:

$$\min \left\{1,\ \exp\left(\mathcal{H}(\boldsymbol{\rho}_{(0)}, \omega_{(0)}) - \mathcal{H}(\tilde{\boldsymbol{\rho}}_{(0)}, \tilde{\omega}_{(0)})\right)\right\}$$

5. Repeat steps 2-4 until the desired number of samples is reached.

We then iterate between solving the planner's problem for $\boldsymbol{d}$ and sampling $\boldsymbol{\rho}$ as follows:

1. Initialize $\boldsymbol{\varphi}_{(0)}$ as the transformed mean of the baseline distribution π.
2. Solve the planner's problem for decision vector $\boldsymbol{d}_{(0)}$ using the updated parameters.
3. Sample $\{\boldsymbol{\rho}_{(s)}\}_{s=1}^{4000}$ from 19 by running HMC simultaneously across 4 independent Markov chains, taking 1000 samples and 500 burn-in samples per chain.
4. Transform samples $\{\boldsymbol{\beta}_{(s)}\}_{s=1}^{4000}$ back into the $\boldsymbol{\varphi}$ space, compute $\bar{\boldsymbol{\varphi}}$ as the mean across samples, and update $\boldsymbol{\varphi}$ using $\boldsymbol{\varphi}_{(t+1)} := w\bar{\boldsymbol{\varphi}} + (1-w)\boldsymbol{\varphi}_{(t)}$, with $w = 0.25$.

5. Repeat steps 2 – 4 until $||\boldsymbol{\varphi}_{(t+1)} - \boldsymbol{\varphi}_{(t)}||_\infty < 0.001$.

Computational implementation details

To sample from 19, we rely on the Stan software (Carpenter et al., 2017, Stan Development Team, 2023) for high-performance statistical computation. The Stan implementation for HMC makes a few adaptations to the algorithm described above to improve computation speed and sampling efficiency. We summarize these below:

- To ensure convergence onto the stationary target distribution, Stan discards the pre-specified number of burn-in samples at the start of the sampling process.
- Stan utilizes the No U-turn sampling (NUTS) variant of HMC, which adaptively determines the number of leapfrog steps L at each iteration to avoid U-turns in the state trajectory (Hoffman and Gelman, 2014, Betancourt, 2016).
- Stan determines the leapfrog step size ϵ using the dual averaging Nesterov algorithm (Nesterov, 2009).
- By default, Stan utilizes a diagonal matrix for M which is estimated using the burn-in samples collected at the start of the algorithm.
- Stan uses reverse-mode automatic differentiation to compute the Hamiltonian gradient.

References

Angelsen, Arild. 2017. REDD+ as result-based aid: General lessons and bilateral agreements of Norway. *Review of Development Economics* 21 (2):237–264.

Araujo, Rafael, Francisco Costa, and Marcelo Sant'Anna. 2022. Efficient forestation in the Brazilian Amazon: Evidence from a dynamic model.

Araujo, Rafael, Juliano Assunção, Marina Hirota, and José A Scheinkman. 2023. Estimating the spatial amplification of damage caused by degradation in the Amazon. *Proceedings of the National Academy of Sciences* 120 (46):e2312451120.

Assunção, Juliano, Clarissa Gandour, Romero Rocha, and Rudi Rocha. 2020. The effect of rural credit on deforestation: evidence from the Brazilian Amazon. *The Economic Journal* 130 (626):290–330.

Assunção, Juliano, Clarissa Gandour, and Romero Rocha. 2022a. DETERring deforestation in the Brazilian Amazon: environmental monitoring and law enforcement. *American Economic Journal: Applied Economics (Forthcoming)* .

Assunção, Juliano, Robert McMillan, Joshua Murphy, and Eduardo Souza-Rodrigues. 2022b. Optimal environmental targeting in the amazon rainforest. *The Review of Economic Studies (Forthcoming)* .

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.

Balboni, Clare, Aaron Berman, Robin Burgess, and Benjamin A Olken. 2022. The Economics of Tropical Deforestation. *The Annual Review of Economics, forthcoming* .

Bemporad, Alberto, Manfred Morari, Vivek Dua, and Efstratios N. Pistikopoulos. 2002. The explicit linear quadratic regulator for constrained systems. *Automatica* 38:3–20.

Betancourt, Michael. 2016. Identifying the Optimal Integration Time in Hamiltonian Monte Carlo.

Cai, Yongyang and Kenneth L. Judd. 2023. A Simple But Powerful Simulated Certainty Equivalent Approximation Method for Dynamic Stochastic Problems. *Quantitative Economics* Forthcoming.

Cai, Yongyang, Kenneth Judd, and Jevgenijs Steinbuks. 2017. A nonlinear certainty equivalent approximation method for dynamic stochastic problems. *Quantitative Economics* 8:117–147.

Carpenter, Bob, Andrew Gelman, Matthew D Hoffman, Daniel Lee, Ben Goodrich, Michael Betancourt, Marcus A Brubaker, Jiqiang Guo, Peter Li, and Allen Riddell. 2017. Stan: A probabilistic programming language. *Journal of statistical software* 76.

Correa, Juliano, Richard van der Hoff, and Raoni Rajão. 2019. Amazon Fund 10 years later: lessons from the world's largest REDD+ program. *Forests* 10 (3):272.

Dominguez-Iino, Tomas. 2021. Efficiency and redistribution in environmental policy: An equilibrium analysis of agricultural supply chains. Tech. rep., Working Paper.

Fick, Stephen E and Robert J Hijmans. 2017. WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *International journal of climatology* 37 (12):4302–4315.

Friedlingstein, Pierre, Matthew W Jones, Michael O'Sullivan, Robbie M Andrew, Dorothee CE Bakker, Judith Hauck, Corinne Le Quéré, Glen P Peters, Wouter Peters, Julia Pongratz, et al. 2022. Global carbon budget 2021. *Earth System Science Data* 14 (4):1917–2005.

Gandour, Clarissa. 2018. *Forest Wars: A Trilogy on Combating Deforestation in the Brazilian Amazon*. Ph.D. thesis, Economics Department, Pontifícia Universidade Católica do Rio de Janeiro.

Griscom, Bronson W, Justin Adams, Peter W Ellis, Richard A Houghton, Guy Lomax, Daniela A Miteva, William H Schlesinger, David Shoch, Juha V Siikamäki, Pete Smith, et al. 2017. Natural climate solutions. *Proceedings of the National Academy of Sciences* 114 (44):11645–11650.

Hansen, Lars Peter and Thomas J Sargent. 2013. Risk, Uncertainty, and Value. *Princeton, New Jersey: Princeton* .

Heinrich, Viola HA, Ricardo Dalagnol, Henrique LG Cassol, Thais M Rosan, Catherine Torres de Almeida, Celso HL Silva Junior, Wesley A Campanharo, Joanna I House, Stephen Sitch, Tristram C Hales, et al. 2021. Large carbon sink potential of secondary forests in the Brazilian Amazon to mitigate climate change. *Nature communications* 12 (1):1–11.

Hoffman, Matthew D. and Andrew Gelman. 2014. The No-U-Turn Sampler: Adaptively Setting Path Lengths in Hamiltonian Monte Carlo. *Journal of Machine Learning Research* 15 (47):1593–1623.

IBGE. 2017. Censo Agropecuário: Tabelas 6882, 6911.

Klibanoff, Peter, Massimo Marinacci, and Sujoy Mukerji. 2005. A Smooth Model of Decision Making Under Uncertainty. *Econometrica* 73:1849–1892.

Leite-Filho, Argemiro Teixeira, Britaldo Silveira Soares-Filho, Juliana Leroy Davis, Gabriel Medeiros Abrahão, and Jan Börner. 2021. Deforestation reduces rainfall and agricultural revenues in the Brazilian Amazon. *Nature Communications* 12 (1):2591.

Lovejoy, Thomas E and Carlos Nobre. 2018. Amazon tipping point.

Malhi, Yadvinder, Daniel Wood, Timothy R. Baker, James Wright, Oliver L. Phillips, Thomas Cochrane, Patrick Meir, Jerome Chave, Samuel Almeida, Luzmilla Arroyo, Niro Higuchi, Timothy J. Killeen, Susan G. Laurance, William F. Laurance, Simon L. Lewis, Abel Monteagudo,

David A. Neill, Percy Nunez Vargas, Nigel C. A. Pitman, Carlos Alberto Quesada, Rafael Salomao, Jose Natalino M. Silva, Armando Torres Lezama, John Terborgh, Rodolfo Vasquez Martinez, and Barbara Vinceti. 2006. The regional variation of aboveground live biomass in old-growth Amazonian forests. *Global Change Biology* 12 (7):1107–1138.

Neal, Radford M et al. 2011. MCMC using Hamiltonian dynamics. *Handbook of markov chain monte carlo* 2 (11):2.

Nesterov, Yurii. 2009. Primal-dual subgradient methods for convex problems. *Mathematical programming* 120 (1):221–259.

Raiffa, Howard, Robert Schlaifer, et al. 1961. *Applied statistical decision theory*. Wiley.

Santoro, Maurizio and Oliver Cartus. 2021. ESA Biomass Climate Change Initiative (Biomass_cci): Global datasets of forest above-ground biomass for the years 2010, 2017 and 2018, v3.

Scokaert, Pierre O.M. and James B. Rawlings. 1998. Constrained linear quadratic regulation. *IEEE Transactions on Automatic Control* 43:1163–1169.

Scott, Paul. 2014. Dynamic discrete choice estimation of agricultural land use. Tech. rep., TSE Working Paper.

SEAB-PR. 2021. Preço Médio - Recebido Pelo Agricultor: Boi Gordo, Arroz (Em Casca), Cana-de-Açúcar, Milho, Mandioca, 1990-2021.

Souza Jr, Carlos M, Julia Z. Shimbo, Marcos R Rosa, Leandro L Parente, Ane A. Alencar, Bernardo FT Rudorff, Heinrich Hasenack, Marcelo Matsumoto, Laerte G. Ferreira, Pedro WM Souza-Filho, et al. 2020. Reconstructing three decades of land use and land cover changes in brazilian biomes with landsat archive and earth engine. *Remote Sensing* 12 (17):2735.

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

Stan Development Team. 2023. Stan Modeling Language Users Guide and Reference Manual. Version 2.33.

Steffen, Will, Johan Rockström, Katherine Richardson, Timothy M Lenton, Carl Folke, Diana Liverman, Colin P Summerhayes, Anthony D Barnosky, Sarah E Cornell, Michel Crucifix, et al. 2018. Trajectories of the Earth System in the Anthropocene. *Proceedings of the National Academy of Sciences* 115 (33):8252–8259.

Thangavel, S., S. Lucia, R. Paulen, and S. Engell. 2018. Dual robust nonlinear model predictive control: A multi-stage approach. *Journal of Process Control* 72:39–51.