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)

February 22, 2024

Abstract

Reforestation in the Brazilian Amazon absorbs carbon, and deforestation for cattle-ranching produces carbon emissions. The social productivities vary across locations for the alternative activities. We analyze a spatial/dynamic model of efficient land allocation to establish a benchmark for policies. We treat cattle prices as stochastic and location-specific productivities as uncertain when assessing the consequences of imposing alternative prices of carbon emissions. Modest price increases would incentivize Brazil to choose policies that capture a significant amount of greenhouse gases in the next 30 years. Our analysis pinpoints tropical forest management as an important contributor to climate change mitigation.

*We thank Pengyu Chen, Bin Cheng, Patricio Hernandez, João Pedro Vieira, Daniel (Samuel) Zhao for their expert research assistance, to Joanna Harris and Diana Petrova for their helpful comments and to Carmen Quinn for editorial assistance. Assunção thanks the Climate Policy Initiative-Brazil for research support. Hansen thanks EPIC/Argonne National Laboratory and the Griffin Applied Economics Incubator Project for research support. Scheinkman's research was partially supported by Columbia Climate School and by Princeton University.

1 Introduction

This paper investigates the potential social gains of 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.)

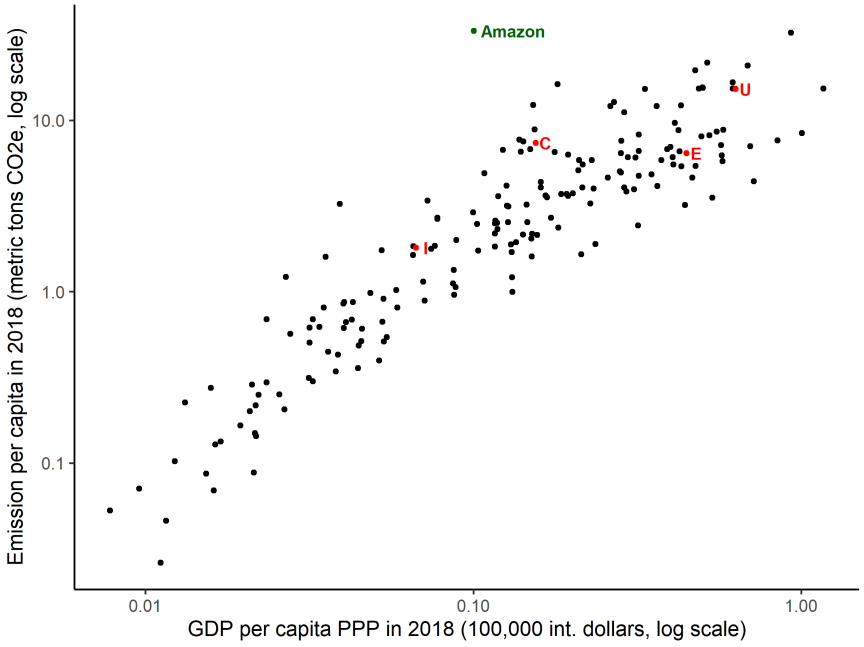


Figure 1: Each dot represents a country in 2018, except for the European Union and the Brazilian Amazon. Highlighted letters stand for (C)hina, (I)ndia, (E)uropean Union, and (U)nited States. Sources: World Bank Data, downloaded on March 2021; Fatos da Amazônia 2021 (www.amazonia2030.org).

We use our model of the Brazilian Amazon 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

[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.

in the Brazilian Amazon. This heterogeneity highlights the importance of incorporating a spatial dimension in the model. To account for these locational differences, we divide the Amazon region into various sub-regions or sites, each of which has its own capacity to support agriculture and forestry. 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 the associated 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 a numerical method, in this case a Markov chain Monte Carlo method based on Hamiltonian dynamics. Such a method is particularly valuable for high-dimensional problems and has substantial advantages over the familiar Metropolis-Hastings approach. We use the Hamiltonian Markov chain approach in a novel way to confront what is sometimes referred to as "deep uncertainty" about productivity parameters. Alternatively, we may interpret this treatment of subjective ambiguity as a robust Bayesian formulation of the control problem of interest.

To set the stage for our analysis, 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.[3] We use the inferred shadow price for 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.[4]

We study the impact of adding outside payments for *net* capture of CO_2 in the Amazon. We produce results in three steps. In step one, 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. Instead, we impose 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

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

[3]Their funding was a pay-for-performance scheme based on an emissions price of $5 per ton of CO2e. It generated USD 1.2 billion in payments in 2008-2017 (Angelsen (2017), Correa et al. (2019)) and provides an example of how deforestation may be influenced by outside payments.

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

and by how much Brazil would gain if the country signed an agreement for a set of hypothetical dollar transfers per net ton of CO_2 captured. Our model output shows the overall social gains to reallocating production in the cross-section, including the option to preserve or enhance the Brazilian rain forest.

In the second 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.

Central to our analysis, we allow the productivities for carbon absorption and agriculture to be site-specific. Estimates of these crucial productivity parameters are subject to non-trivial uncertainty. Moreover, while we have cross-sectional data that are informative, these data do not give us direct measurements for the site-specific productivities. Instead, we use regression methods to provide the inputs needed for our analysis. Uncertainty in the regression coefficients induces uncertainty in the implied site-specific productivity parameters. The typical approach to uncertainty quantification explores such parameter uncertainty from the perspective of an external analyst. Instead, we incorporate this uncertainty explicitly into the decision problem of our hypothetical social planner. While it is appealing to address the parameter uncertainty probabilistically, there is ambiguity as to what probabilities to impose.

In the third step, we explicitly consider parameter ambiguity from the perspective of the social planner. We again use the less refined 78-site partition of the Brazilian Amazon. We start with a baseline posterior distribution over site-specific productivity parameters for carbon sequestration and agricultural productivity implied by the regression estimation. While convenient, this baseline construction is *ad hoc* and uncertain. This leads us to engage in a sensitivity analysis that explores distributional sensitivity subject to penalization. This estimation is done prior to the decision-making of the social planner and so acts as "prior" when exploring alternative courses of action. Rather than fully embracing this distribution, the planner uses it as a baseline in a sensitivity analysis subject to penalization to ascertain which departures should be of most concern. The magnitude of a 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 ambiguity, we use the Markov chain Monte Carlo method based on Hamiltonian dynamics that we alluded to earlier.

While we assume that the planner can directly control deforestation, we 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.[5]

[5]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 in 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. (2023a) and Assunção et al. (2023b)).

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 shows how we confront parameter uncertainty. Section 5 summarizes how we use a large collection of relevant data sets to calibrate the model. Section 6 discusses the numerical methods used to compute solutions to the social planner's maximization problem. Our results are presented in Section 7, 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) focus on the potential of the Brazilian Amazon. As noted by Balboni et al. (2023), 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 in the design of policies, without explicitly modeling the carbon cycles associated with the forest. Araujo et al. (2022), on the other hand, present 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) are 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 forests 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 of 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 of 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.[6] 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 a 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.

[6]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 $t \in [0, T]$ is 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 lowercase 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.[7]

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.[8] 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 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 on the right side of (2) connects deforestation to a loss in captured carbon. The site-specific parameter $\gamma^i > 0$ denotes

[7] 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).

[8] For calibration of this and the other parameters see Section 5

the density of CO2e that is present in a primary forest in site i.[9] 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 $100[1-\exp(-\alpha 100)]\%$ of the maximal captured CO2e in 100 years as in Heinrich et al. (2021). In our case, we choose α such that $100[1-\exp(-\alpha 100)] = 99\%$. 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, one of the controls is always zero, which introduces additional binding constraints into the analysis. We write this constraint as: $U_t^i V_t^i = 0$.

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 parameter κ captures the emissions that result because of cattle farming.[10] We include a cost of adjustment to changes in the use of land with contributions from each site. It is measured by

$$\frac{\zeta}{2}\left[\sum_{i=1}^{I}\left(U_t^i + V_t^i\right)\right]^2.$$

Importantly, the adjustment cost depends on the aggregate change in land use.

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 non-negative 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

[9] 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.

[10] About 75 percent of emissions from agricultural activity in the Amazon is the result of the natural digestive process of cattle. Another approximately 21 percent is from soil management. Thus, for simplicity we assume that cattle herd per hectare does not vary and that productivity variations come mostly from transportation costs and carcass weights.

that describes the future evolution of the price P_t^a for cattle and maximizes

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

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 exogenously specified emissions price, P^e, is an input into the analysis that allows us to explore how costly it will be to make important changes in deforestation outcomes in Brazil. Its magnitude reflects the sum of the marginal value attributed by the planner to emission and any monetary transfers obtained from others, such as sales in carbon emission markets. In the computations that follow, we take this price to be fixed over time, but in future work we intend P^e to fluctuate over time.

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 a 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. A more complex model would introduce interactions across sites via non-linearities in the valuation of agricultural output and/or emissions, an extension worthy of exploration in future work.*

Remark 3.3. *To show 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 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 emission trajectories from the optimal solution, one could use geo-science inputs to inform the mapping of emissions from the Brazilian Amazon and elsewhere into carbon in the atmosphere to compute the impact on the evolution of carbon stocks. This would require a much more comprehensive model that is beyond the scope of this particular exercise.*

4 Parameter uncertainty

We treat parameter uncertainty probabilistically by starting with a baseline subjective prior over the parameters. We form this baseline "prior" conditioned on available data. Rather than assuming a full commitment to this baseline distribution, we allow for some skepticism by exploring sensitivity to distributional changes. We limit the scope of the sensitivity analysis by penalizing deviations from the baseline prior using a relative entropy or Kullback-Leibler measure of divergence. This divergence is well known to have convenient mathematical and conceptual implications. We implement this sensitivity analysis by converting our one-person maximization problem into a two-player game where the sensitivity analysis is conducted via minimization. This delivers a form of ambiguity aversion consistent with two alternative representations of ambiguity aversion: smooth ambiguity and variational preferences.[11]

Our model is dynamic and Markovian. As such, it could be formulated from the vantage point of an initial period or recursively. Even in the absence of parameter uncertainty, we find it convenient computationally to solve it from the perspective of an initial date. This makes it a "static problem." We adopt this same static perspective to explore the consequences of uncertainty. In contrast to the single-agent decision theory, this has conceptual implications beyond just computational considerations as the minimization is performed as well at the initial date. In effect, we treat the two-player formulation as a static max-min game where, as we noted, the minimizing player is used as a formal device to explore the sensitivity to changes in the distribution over parameters used in optimization.

We implement a static formulation of robustness to parameter uncertainty as follows: For each site, we consider the parameter pair (γ^i, θ^i) for $i = 1, 2, ..., I$. Let φ denote full parameter vector including all sites and hence of dimension $2 \times I$. We use a regression approach to construct baseline estimates of the site-specific productivities given site attributes. Each site may intersect multiple municipalities, and we write M^i for the set of municipalities that overlap site i. We construct site-specific productivities using:

$$\begin{bmatrix} \gamma^i \\ \theta^i \end{bmatrix} = \begin{bmatrix} \sum_{m \in M^i} w_m^i \exp\left(\beta_\gamma \cdot R_\gamma^m\right) \\ \frac{1}{P_{2017}^A} \sum_{m \in M^i} w_m^i \exp\left(\beta_\theta \cdot R_\theta^m\right) \end{bmatrix}. \tag{6}$$

In this regression, R_γ^m is a vector of geographical variables used to construct baseline estimates for the carbon-absorption productivity for municipality, m. Similarly, R_θ^m is a vector of such variables used to construct baseline estimates for the value of agricultural output per hectare for the same municipality. w_m^i is an area-based weight of the importance of municipality m in site i. Finally, P_{2017}^A is the price of cattle in 2017.

[11] See Klibanoff et al. (2005) for an initial axiomatic foundation for smooth ambiguity, Maccheroni et al. (2006) for an axiomatic foundation for variational preferences. Maccheroni et al. (2006) used Hansen and Sargent (2001)'s application of relative entropy divergence as motivation for their work. See Hansen and Sargent (2023) for a recent analysis relating both preference formulations to statistical decision theory and to robust control theory by embracing a different axiomatic perspective.

Uncertainty in the composite regression parameter vector, $\beta' \stackrel{\text{def}}{=} (\beta_\gamma', \beta_\theta')$, induces uncertainty in the site-specific productivities, (γ^i, θ^i) for $i = 1, 2, ..., I$. The underlying dimension of the uncertainty is given by the number of the unknown regression parameters, which is substantially less than $2 \times I$. This reduction turns out to be important for implementation. We use a regression method because we do not have direct evidence for each of the site-specific productivities. Our productivity data are available at different resolutions than the sites within our model. For these reasons, we use attributes as right-hand side variables in the regression and feed in site-specific attributes. Moreover, we use the regression approach to fill in missing observations. Since the dependent variables in the regressions are expressed in terms of logarithms, we exponentiate the predictions implied by regressions.

We use historical evidence on land-use productivity to estimate the two regression equations as functions of attributes. We impose a familiar and convenient conjugate prior to produce a Bayesian posterior π, for the β's. See, for instance, Raiffa et al. (1961). The unknown regression coefficients are presumed to be normally distributed conditioned on the regression-error variances and the regression-error variance are posited to have an inverse gamma distribution. We impose uninformative limits on the priors and use quasi-analytical formulas to deduce posterior distributions for the regression predictions that we use as baseline probability distributions for the site-specific productivities.[12] While the baseline distributions π constructed in this matter are "posterior" from the perspective of our regression analysis, they play the role of a "prior" for the social planner acting conditioned on data used in the regression analysis.[13]

Our measure of divergence used to explore sensitivity restricts the alternative probabilities to be absolutely continuous with respect to the baseline distribution. With this restriction, it suffices to focus only on alternative distributions to the baseline specification for the regression coefficients. This simplifies substantially the numerical computations.

Remark 4.1. *An interesting extension could include a probabilistic specification of heterogeneity in the productivity parameters not captured by the regression formulation. This would add to the computational burden as the sensitivity analysis would be over a substantially larger space and would require a plausible probability specification of the unobserved heterogeneity.*

Given the vector of regression parameters, β, in a set, $\mathcal{B}$, we write $f(d, \beta)$ for the discounted utility obtained from a sequence of decisions d with implied productivity parameters given by formula (6). We abstract from uncertainty in this calculation, but this could also be incorporated into the analysis. We use a familiar relative entropy (or Kullback-Leibler) divergence measure to capture ambiguity about the parameter distribution given by:

$$\int_{\mathcal{B}} \log g(\beta) g(\beta) d\pi(\beta) \geqslant 0,$$

[12] See Appendix C for details

[13] Our static formulation of the two-player decision problem does not allow for the possibility of dynamic learning going forward. We abstract from this for computational tractability.

where g satisfies $\int g(\beta)d\pi(\beta) = 1$. Notice that the non-negative (relative) density, g, implies an alternative probability distribution given by $g(\beta)d\pi(\beta)$. Preferences over alternative decision sequences are given by:

$$\min_{g\geqslant 0, \int g d\pi = 1} \int_{\mathcal{B}} \left[f(d,\beta) + \xi \log g(\beta) \right] g(\beta) d\pi(\beta),$$

for a penalty parameter $\xi > 0$. For ξ arbitrarily large, these preferences are well approximated by expected utility preferences using the baseline distribution π. See Maccheroni et al. (2006).[14]

These preferences are recognizable as a special case of what are called variational preferences for a static decision problem. The minimization problem has a well-known quasi-analytical solution:

$$g^*(\beta) = \frac{\exp\left[-\frac{1}{\xi} f(d,\beta)\right]}{\int_{\mathcal{B}} \exp\left[-\frac{1}{\xi} f(d,\beta)\right] d\pi(\beta)} \tag{7}$$

with a minimized objective:

$$-\xi \log \int_{\mathcal{B}} \exp\left[-\frac{1}{\xi} f(d,\beta)\right] d\pi(\beta). \tag{8}$$

The minimizing g given in (7) induces an "exponential tilt" of the probabilities towards lower discounted utilities. The magnitude of ξ determines the strength of this tilt. We shall refer to limiting $\xi = \infty$ case as *ambiguity neutrality*. For this limit, the decision problem uses the familiar expected utility objective:

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

Remark 4.2. *The minimized objective given by* (8) *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* (8) *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.*

Remark 4.3. *The conjugate prior formulation we implemented leads naturally to a conditional posterior for the regression coefficients and a marginal posterior for the regression-error variances. As posed, the minimization is over the marginal posterior for the regression coefficients. With relative entropy divergence, we may equivalently solve the minimization problem using the relative entropy divergence over the conditional posterior distribution for the regression coefficients and the marginal posterior for the regression-error variances. We use this observation in our actual calculations to simplify our algorithmic implementation.*

[14] Maccheroni et al. (2006) used Hansen and Sargent (2001)'s application of relative entropy divergence as motivation for their work.

Given the parameter ambiguity adjustment, the implied decision problem is a two-player zero-sum game. Let the decision process d be constrained to be in a convex set $\mathcal{D}$.

Problem 4.4.

$$\max_{d\in\mathcal{D}} \min_{g\geqslant 0,\int g d\pi=1} \int_{\mathcal{B}} f(d,\beta)g(\beta)d\pi(\beta) + \xi \int_{\mathcal{B}} \log g(\beta)g(\beta)d\pi(\beta).$$

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

Problem 4.5.

$$\min_{g\geqslant 0,\int g d\pi=1} \max_{d\in\mathcal{D}} \int_{\mathcal{B}} f(d,\beta)g(\beta)d\pi(\beta) + \xi \int \log g(\beta)g(\beta)d\pi(\beta).$$

Under sufficient conditions that support the Max-min Theorem, the resulting objective for Problems 4.4 and 4.5 will be same and the minimizing g evaluated at the maximized d for the Problem 4.4 will agree with the minimizing g from Problem 4.5. Similarly, the optimized decision processes will agree.

Consider the inner maximization for Problem 4.5:

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

where we are free to drop the relative entropy penalty as it does not depend on the decision process d. Provided that this inner 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.[15]

5 Productivity measurement

As we discussed in section 4, we construct site specific productivity estimates from the output of regression equations. See formula (6). Appendix A describes in detail all of the data used for these regressions. What follows is a summary of the evidence that we draw on.

For input into the agricultural productivity regression, we use the year of 2017 as a reference for many variables, since this is the year of the latest Agricultural Census in Brazil. For the regressand, this census provides information on the value of cattle sold for slaughter per hectare of pasture land at the level of a municipality.[16] As regressors, we use geographical variables as stipulated in

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

[16] Considering the total revenue of the cattle ranching activity, we may be overestimating its importance. It accounts not only for the value created in the local economy, in the form of profits and wages but also for some inputs specifically consumed in production. Unfortunately, our data does not allow us to construct a reliable measure for the value added by grazing in the Brazilian Amazon.

Appendix A. The census provides observations on the value of cattle sold for slaughter per hectare of pasture land for 466 municipalities out of the 540 municipalities that intersect the biome. Since we have data on the regressors, and we use regression predictions to attribute values to the remaining municipalities and to obtain agricultural productivities, the θ^i's.

For measuring the productivity of carbon sequestration, the γ^i's, we first use data from MapBiomas [17] to select pixels of $100m \times 100m$ that can be considered primary forests. Given this partitioning, we used 2017 data from ESA Biomass[18] to obtain carbon per hectare. We then calculate average productivities for each municipality. Analogous to the procedure we used for the θ_i's, we run a regression using the MapBiomas data as regressands and geographical variables as regressors to obtain estimates at the spatial resolutions used in our model.

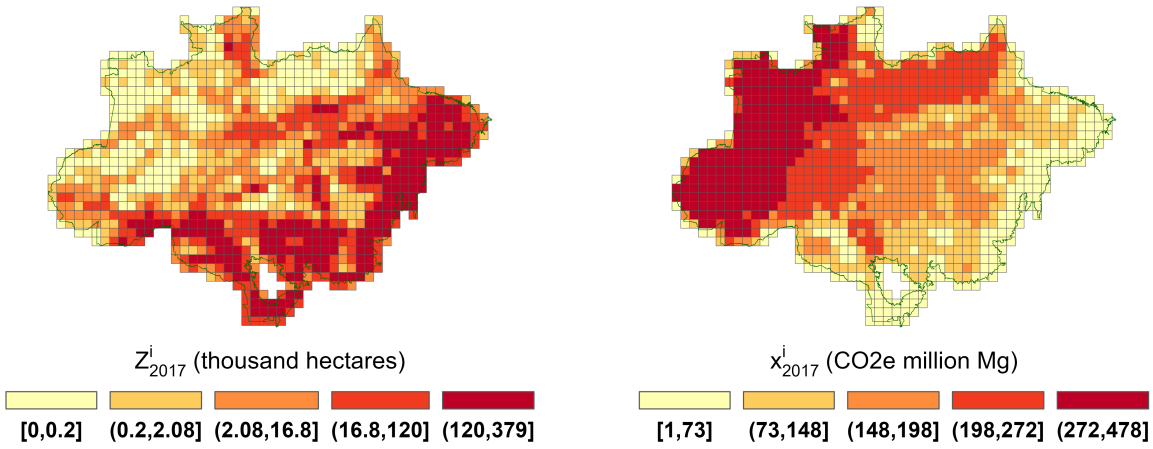
We use two different spatial resolutions for the results that we report. In both cases, we project municipal estimates into two grids of the Amazon biome. At the most detailed level we consider a regular grid of the Amazon region with pixels of 30m $\times$ 30m resolution from MapBiomas (Souza Jr et al., 2020). We then aggregate pixels to form 1887 sites that are 67.5 km $\times$ 67.5km. Many of these sites do not overlap the Amazon biome. We discard these and twenty others with less than 3% of their area in the Amazon biome. This reduced our number to 1043 sites.

For reasons of tractability, when we consider either a stochastic evolution of agricultural prices or uncertainty in the productivity parameters, we use a less refined grid of 130 sites that are approximately 270km $\times$ 270km. We obtain 78 sites after dropping sites that do not overlap the Amazon biome at all and four additional sites with less than 3% in the Amazon biome.

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 $-.35$ for the finer resolution and $-.44$ for the coarser resolution. Thus, while agricultural productivity and carbon absorption capacity are negatively correlated, this relationship is imperfect.

[17] Web address: www.mapbiomas.org (Collection 5).

[18] See Santoro and Cartus (2021).



z_0^i x_0^i

Figure 2: Initial values for agricultural area (z_0^i) and carbon stock (x_0^i)

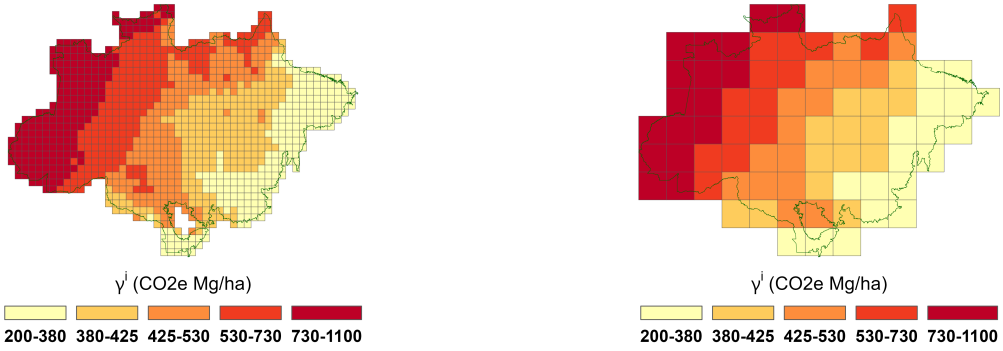


Figure 3: Carbon sequestration parameter heterogeneity (1043 sites left, 78 sites right)

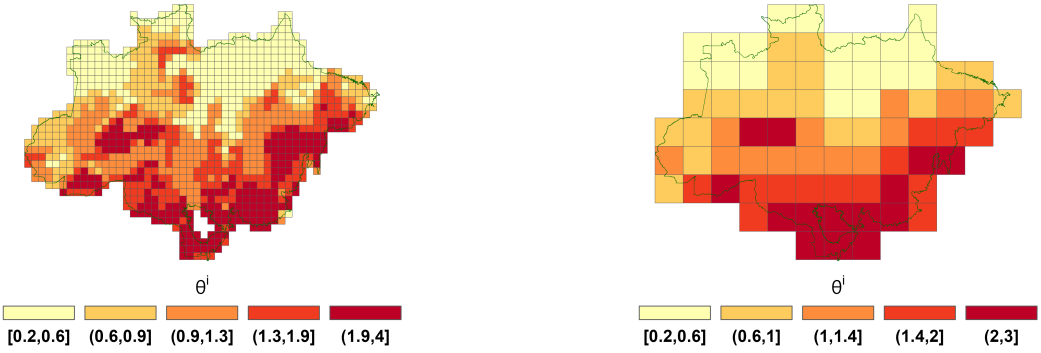


Figure 4: Agricultural productivity heterogeneity (1043 sites left, 78 sites right)

6 Solving the maximization problem

To achieve the needed degree of economic and spatial richness, we use numerical methods to obtain model solutions. Given the number of locations that we use (1043 or 78), we necessarily have a large number of 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 non-negative. 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.

6.1 Solution with parameter ambiguity neutrality

In the absence of price and ambiguity neutrality, we solve a static optimization problem over all possible trajectories for the next 200 years for 1043 sites. This problem is deterministic evaluated at the *ex-ante* average parameters using baseline distributions. Given obvious uniform bounds on possible utility flows and discount rates of at least 2%, the resulting trajectories give reasonable approximations to the infinite horizon optima.

6.2 Solution with parameter ambiguity aversion

For solving the robustly optimal problem numerically, we take an iterative approach, supported by the Min-max Theorem. Specifically, we proceed as follows:

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

ii) For a given d, we solve the minimization problem with the relative entropy penalization to obtain a new candidate for g^*.

iii) We repeat the steps until we achieve convergence.

Our computations in step ii) will exploit the quasi-analytical formulas given in (7), which we repeat here as:

$$g^*(\beta) = \frac{\exp\left[-\frac{1}{\xi} f(d,\beta)\right]}{\int_{\mathcal{B}} \exp\left[-\frac{1}{\xi} f(d,\beta)\right] d\pi(\beta)},$$

where the distribution of interest is $g^* d\pi$. We take given decision process, d, and evaluate the discounted objective to obtain the numerator for formula for g^*. The denominator, however, must be computed numerically. Here we use a Monte Carlo method that was originally developed for computing Bayesian posteriors. This method is based on Hamiltonian dynamics and is often more computationally efficient for high-dimensional problems than the familiar Metropolis-Hastings method. See, for instance, Neal et al. (2011) and Carpenter et al. (2017), with software support given by Stan Development Team (2023).[19] See Appendix D for more details.

[19] From a mathematical standpoint, this calculation is equivalent to the computing a Bayesian posterior where

6.3 Solution with price stochasticity under ambiguity neutrality

To account for stochastic prices in a tractable way, we construct an approximate Markov chain with two states $P^a_\ell < P^a_h$ and transitions that match the empirical transitions from monthly data.[20] When considering a stochastic evolution of prices, 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.[21] 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 $\tau + 1, \dots T$ are set equal to the value that prevails at τ. Prior to date $\tau + 1$, we confront randomness in this problem by imposing appropriate "measurability" restrictions on the controls as functions of potentially realized states. We then apply the interior point method to find the optimal trajectory at zero given P^a_0. We keep the optimal date $t = 1$ states computed at time $t = 0$ and repeat. That is, we consider the problem stating at $t = 1$ with the new state vector and divide the future into two segments: an uncertainty segment of length τ and a remaining period of $T - \tau - 1$. This step will determine an optimal state at period 2. We continue to apply this procedure to produce the optimal state at periods 3, 4, ..., T supported by the corresponding optimal controls.

In practice, the dimensionality of the stochastic problem increases geometrically as a function of the uncertainty horizon, τ. Consequently, 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 $\tau = 0, 1, \dots, \tau^*$. In our coarse grid with 78 sites and price randomness, we chose $\tau^* = 5$.

Many of the results we will show entail projections into the future. We report results based on two hundred simulated sequences of cattle prices, P^a_t, $t = 1, \dots T$, using the observed P^a_0 and the calibrated Markov chain.

7 Results

In this section, we report our quantitative findings. We start by constructing a benchmark "business-as-usual" set of results. We accomplish this by deducing an implied social price of carbon that supports current aggregate implications. Then, in succession, we study i) deterministic solutions for higher social prices; ii) stochastic solutions with price randomness; and iii) solutions with parameter uncertainty.

$-\frac{1}{\xi} f(d, \beta)$ plays the role of a log-likelihood function and π plays the role of a prior. As we commented in Remark 4.3, for programming simplicity, we worked with an augmented parameter space when computing the distribution of interest.

[20] See Appendix A for details.

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

7.1 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 an interval $[\underline{t}, \bar{t}]$ and then select a time invariant price for emissions, denoted by P^{ee}, to match the aggregate deforestation predicted by the model at a final observation period $\bar{t}$. We let $(X^o_{\underline{t}}, Z^o_{\underline{t}})$ denote the initial observed state vector. We also input the realized history of agricultural prices $\{P^a_t : \underline{t} \leqslant t \leqslant \bar{t}\}$. We then compute the optimal trajectory for the state variables implied by our model for alternative choices of P^{ee} and find the choice of P^{ee} that matches the observed value of the aggregate deforestation at $\bar{t}$:[22]

$$\sum_{i=1}^{I} Z^i_{\bar{t}}.$$

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 money contributed mostly by Norway.

The business-as-usual price, P^{ee}, depends on the model specifications, as we document in Table 1. For instance, since an increase in the agricultural price makes cattle farming more productive, the business-as-usual carbon price must increase to achieve the same deforestation outcome. Also, with more sites at a finer resolution, some of the smaller sites will have agricultural productivities that are relatively higher, increasing the incentive to deforest. As Table 1 shows, the business-as-usual price is larger when we have a finer partition of sites. Finally, uncertainty in the productivity parameters leads to a smaller business-as-usual price for reasons that will become clear in our subsequent discussion. In what follows, we will consider solutions to the optimization problem starting in 2017 and a discount rate of 2 percent.

We will explore implications when the planner uses $P^e = P^{ee} + b$ for $b = 0, 10, 15, 20,$ and 25 where b represents transfers per ton of *net* captured emissions to the planner. Specifically, when net emissions total E tons of CO_2, the planner receives a transfer of bE. Note that even $b = 25$ corresponds to a social price that is low when compared to the prevailing price of emissions in some regions of the globe. The forces that lead to changes in shadow prices have a direct and partially offsetting effect on deforestation. Consequently, the implied optimal (or robustly optimal) trajectories for each choice of b will be less sensitive to the particular model specification. Essentially, the same argument applies to changes in the subjective discount rate.[23]

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

[23] Since emissions are a low duration asset relative to cattle raising, a larger discount rate would imply less deforesting for cattle thus lowering P^{ee}, approximating future trajectories for a given b. In fact, our simulations show that future trajectories do not change much for each value of b when we move from 2 to 3 percent.

Table 1: Business-as-usual prices

number of sites	agricultural price	ξ	carbon price (P^{ee})
1043	$p^a = 41.1$	∞	7.6
78	$p^a = 41.1$	∞	7.1
78	$p^a = 41.1$	1	5.3
78	stochastic	∞	6.9

Notes: The agricultural price $P^a = 41.1$ is the mean under the stationary distribution and the price of $P^a = 44.3$ is the upper value used in our two-state Markov chain. The productivity parameters for the $\xi = \infty$ specifications equal the mean computed from the baseline probability distributions for the productivity parameters.

7.2 Results for the case without stochasticity or ambiguity aversion

In this section, we discuss results for a model with a constant price for cattle that equals the average price in the stationary distribution for the estimated 2-state Markov chain ($41.10). We first discuss results for 1043 sites, and then include results for 78 sites for comparison.

As Figure 5 shows, with "business-as-usual" ($P^e = P^{ee} = \$7.6$), the optimal choice involves an increase in the agricultural area from 15% to more than 25% of the biome. This increase may actually cause sufficient deforestation for the hydrological cycle of the Amazon to degrade to the point of being unable to support rain forest ecosystems (Lovejoy and Nobre (2018)). The predicted trajectories are much different with an additional per ton payment to the planner of $10, $15, $20 or $25. Figure 6 reports the trajectories over time of the transfer payments for $b = \$15$ and $b = \$25$. The peak payments occur after about 12 years for both values of b. As expected the transfer payments for $b = \$25$ are much larger that the corresponding payments for $b = \$10$.

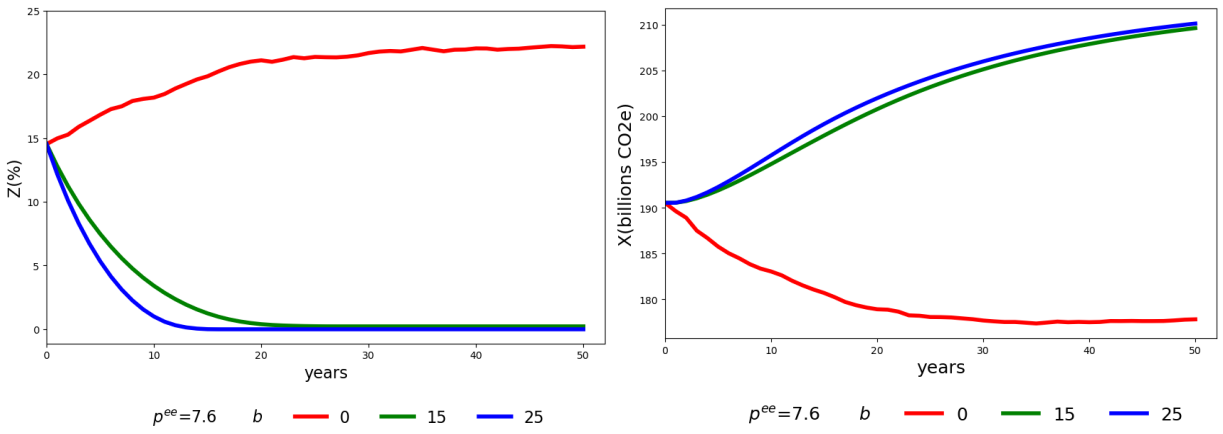


Figure 5: Agricultural area and carbon stock evolution.

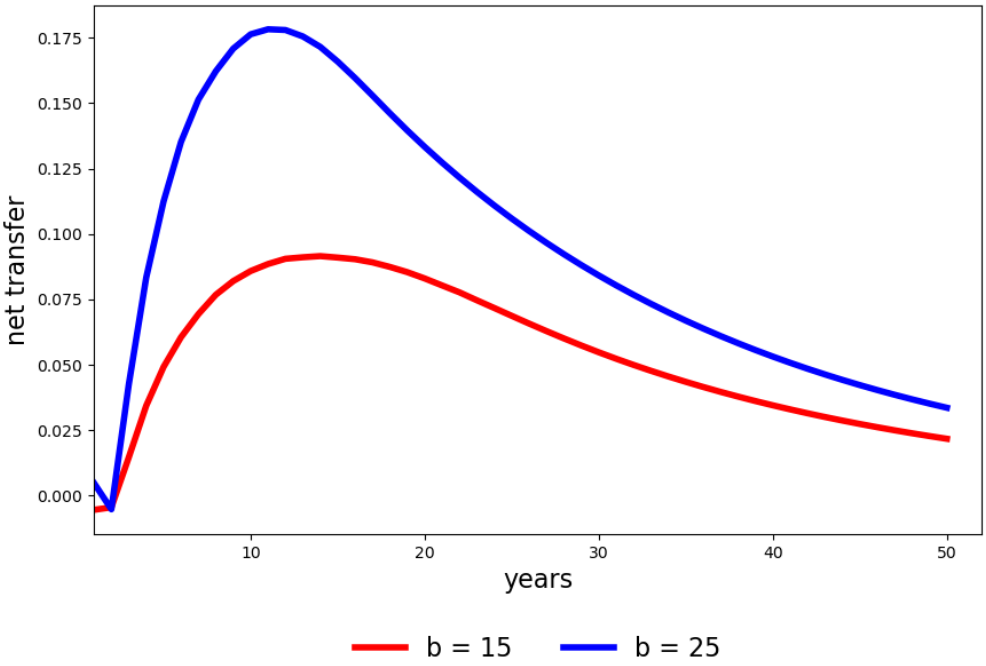


Figure 6: Evolution of transfer payment for two choices of b.

As we now show, these transfer payments result in a substantial decrease in agricultural area and a corresponding increase in forested area. The first five rows of Table 2 give the discounted value to the planner of a commitment to receive b of net transfers for each ton captured of CO_2, when P^a is the stationary price. It also gives a decomposition of this present value to interpretable components. Among these components, "forest services" are measured at the implied Brazilian shadow price for business-as-usual. The net transfers to Brazil are reported separately. Even transfers of \$10 per ton are enough to compensate the losses of agricultural output, but the largest contributor to the gains is the increase in forest services. The larger transfer of \$25 per ton of net captured CO_2 almost doubles the value for the planner - a net gain of \$226 billion. This net gain is composed of a loss of \$354 billion in the value of cattle output,[24] which is more than compensated by \$352 billion in transfers and \$246 billion in forest services. Adjustment costs are a small part of the story.

[24]Recall, however, that we use a measure of full output as value added. Thus, we have exaggerated the loss of agricultural output.

Table 2: Present-value decomposition - 1043 sites

P^e ($)	b ($)	agricultural output value ($ 10^{11})	net transfers ($ 10^{11})	forest services ($ 10^{11})	adjustment Costs ($ 10^{11})	planner value ($ 10^{11})
7.6	0	3.73	0.00	-1.39	0.08	2.26
17.6	10	0.58	1.17	0.89	0.12	2.52
22.6	15	0.33	1.98	1.00	0.18	3.14
27.6	20	0.24	2.76	1.05	0.23	3.82
32.6	25	0.19	3.53	1.07	0.27	4.52

Notes: For P^a, 41.1 is the mean of the agricultural price in the stationary distribution and 44.3 is the upper value in the two-state Markov chain. Forest services are calculated using baseline shadow price ($b = 0$). The present values are computed for two hundred years.

In Table 2, we also report results for an optimistic P^a, the highest price in our 2-state Markov chain. The price P^{ee} that matches observed deforestation now increases to offset the stronger output price incentive for allocating land to agriculture. The present value measure of agricultural output value increases and the corresponding measure of forest services diminishes.

Table 3 displays the total effect of transfers per ton of net CO_2 captured in years 15 and 30. For the business-as-usual carbon price, the planner chooses deforestation that induces carbon emissions of about 12 and 18 billion tons per year in 15 and 30 years, respectively. This table uses this baseline in featuring the "effective cost." We calculated this as the ratio of discounted net transfers to the difference between the net carbon captured and the corresponding baseline value when $b = 0$. With transfers of, say, \$15/ton, optimal management induces capture of about 6.6 billion tons by year 15 and an additional 7.2 billion tons by year 30. The effective costs are about \$4.5 and \$4.9, considerably less than the per ton subsidies captured by the b's. With transfers of \$25/ton, there are modest increases in the captured carbon with effective prices that are almost double, but still about one third the transfer payments per ton. Thus, the results in Table 3 illustrate the gains from trade in instituting a contract that pays Brazil per net ton of CO_2 captured.

Table 3: Transfer costs - 1043 sites

		15 years			30 years		
P^e ($)	b ($)	net captured emissions (billion tons of CO2e)	discounted net transfers ($ 10^{11})	discounted effective cost ($ per ton of CO2e)	net captured emissions (billion tons of CO2e)	discounted net transfers ($ 10^{11})	discounted effective cost ($ per ton of CO2e)
7.6	0	-12.04	0.00	NaN	-17.66	0.00	NaN
17.6	10	5.20	0.44	2.53	11.67	0.86	2.93
22.6	15	6.66	0.84	4.50	13.85	1.55	4.92
27.6	20	7.58	1.28	6.53	14.62	2.21	6.85
32.6	25	8.16	1.73	8.56	15.00	2.86	8.75

Notes: Agricultural price $P^a = 41.11$, which is the mean of the agricultural price in the stationary distribution.

Figure 7 exhibits the initial distribution of land allocation over 30 years for transfers per ton = \$0, \$10, and \$25. Specifically, Figure 7 shows that for the case of transfers that exceed \$10 per ton of net emissions, the area of the biome that is occupied by cattle farming after 30 years would be substantially reduced in comparison to the 2017 allocation. This is in sharp contrast to what transpires in the $b = 0$ business-as-usual specification in which agricultural production becomes quite intense in the lower right sites.

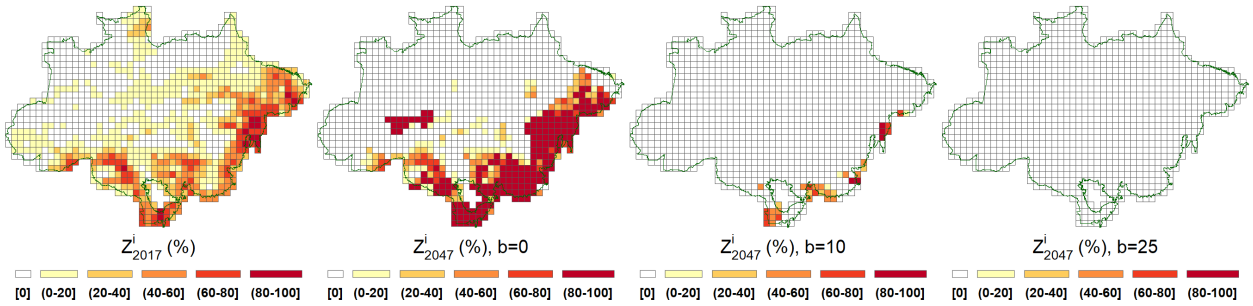


Figure 7: Agricultural area changes after 30 years.

Figure 8 provides a more complete spatial dynamic characterization for transfers of \$15/ton. In the optimal solution, much of the change in land occupation occurs within the first 15 years. Appendix A shows how much this process slows down for higher choices of the adjustment cost.

Table 4: Present-value decomposition - case 78 sites

P^a (\$)	P^e (\$)	b (\$)	agricultural output value (\$ 10^{11})	net transfers (\$ 10^{11})	forest services (\$ 10^{11})	adjustment costs (\$ 10^{11})	planner value (\$ 10^{11})
41.1	7.1	0	3.31	0.00	-1.10	0.06	2.14
41.1	17.1	10	0.43	1.23	0.87	0.12	2.42
41.1	22.1	15	0.26	2.02	0.95	0.17	3.06
41.1	27.1	20	0.21	2.79	0.99	0.23	3.75
41.1	32.1	25	0.17	3.54	1.00	0.26	4.45

Notes: For P^a, 41.1 is the mean of the agricultural price in the stationary distribution. Forest services are calculated using baseline shadow price ($b = 0$).

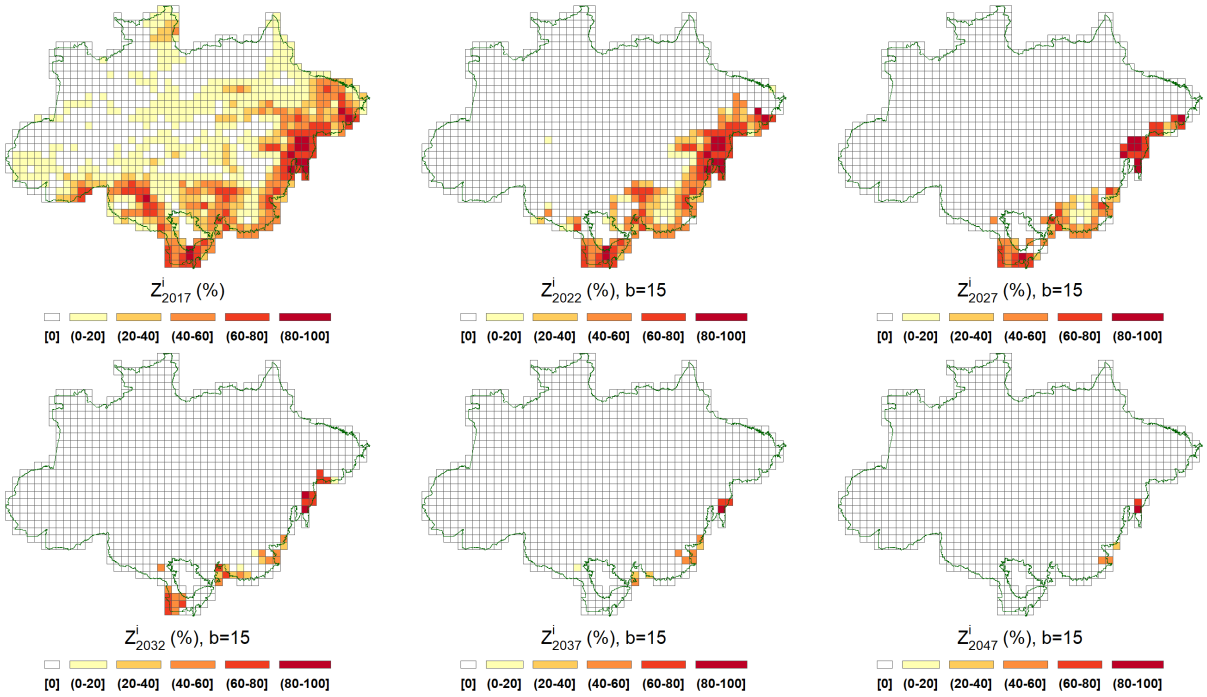


Figure 8: Agricultural area evolution over time

Our computations with parameter uncertainty and agricultural price stochasticity use a coarser grid with 78 sites. Before exploring those results, Table 4 shows what happens to the present-value decomposition if we only reduce the resolution. Overall, the findings are quite similar for the two resolutions. Since some of the highly productive refined sites for either agriculture or carbon absorption have been aggregated with other less productive sites, the planner's options are more limited. Consequently, there are small reductions in agricultural output values and in the magnitudes of the forest services. This can be seen by comparing the respective columns of Table 4 with those in Table 2.

7.3 Results with robustness to parameter uncertainty

In this section, we present results when the planner is uncertain about cattle productivity and CO_2 capture potential. While it is revealing to perform robustness calculations for several values of ξ, here we report results only for $\xi = \infty$ and $\xi = 1$. We refer to the former as 'ambiguity neutral' and the latter as 'ambiguity averse.' We report results for other values of ξ in the appendix A.11. The implied ambiguity adjustments to the probabilities help us gauge the plausibility of different values of ξ. The calculated shadow price, as reported in Table 1, is \$7.1/ton under ambiguity neutrality and a considerably lower value of \$5.3/ton under ambiguity aversion. The shadow price reduction under ambiguity aversion compensates for the slower destruction of the forests when there is ambiguity in the agricultural productivities for the various sites.

Figure 9 shows the baseline and ambiguity-adjusted distributions for parameters γ and θ when $b = 0$ and $b = 15$. We display results for the two sites with largest divergence between the baseline

probability distributions and the ambiguity-adjusted counterparts for sake of illustration.[25] When looking across all sites, the adjustments are very heterogeneous; and for some sites there is very little difference between the two distributions.

Under the business-as-usual benchmark, the adjustments are more pronounced for the θ distributions than for the γ distributions. In other words, it is the uncertainty about agricultural productivity that the social planner finds to be most concerning since the planner is not incentivized to preserve the rain forest without transfer payments.[26] The uncertainty adjustment to the probabilities are substantially different when $b = 15$. Now the adjustments are more pronounced for the γ probability distributions because reforestation becomes a more prominent ambition for the planner.

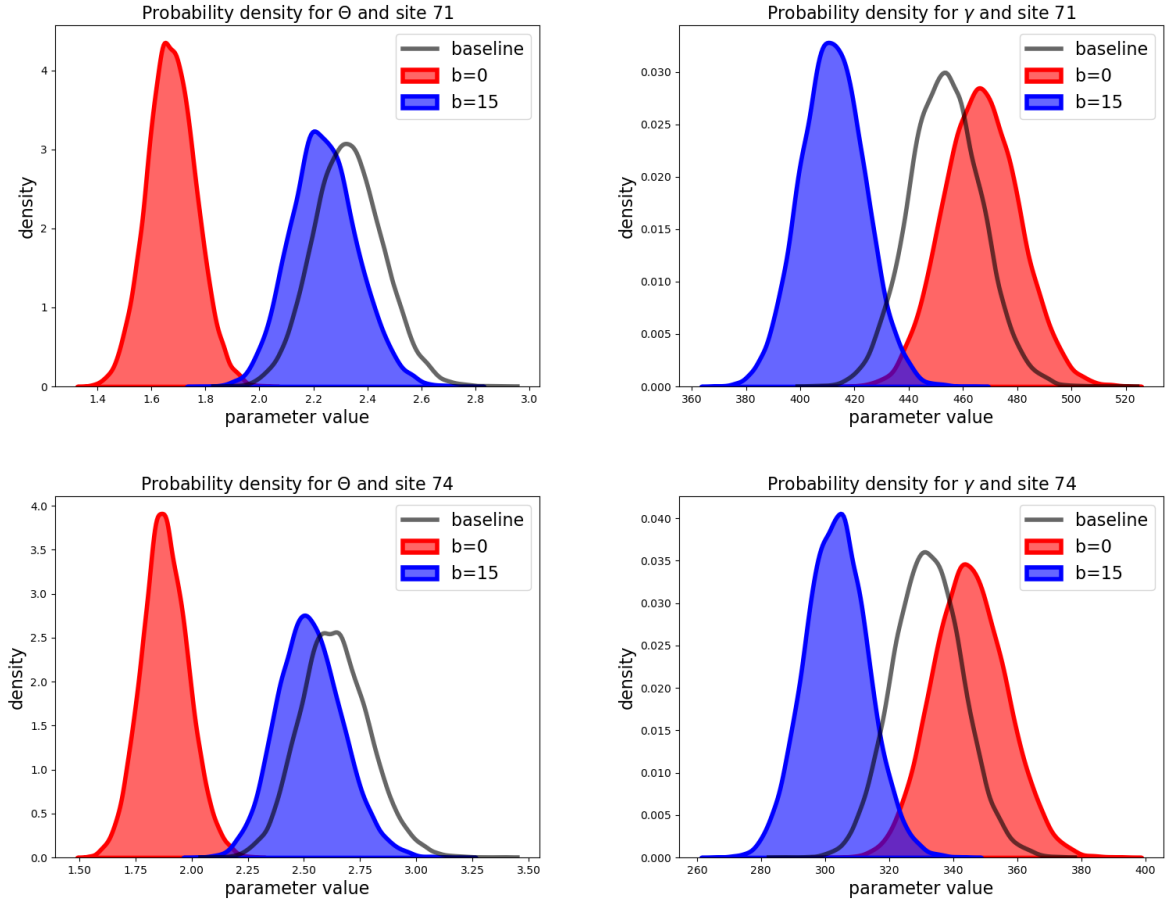


Figure 9: Baseline and ambiguity adjusted densities for $b = 0$ and $b = 15$ for two sites.

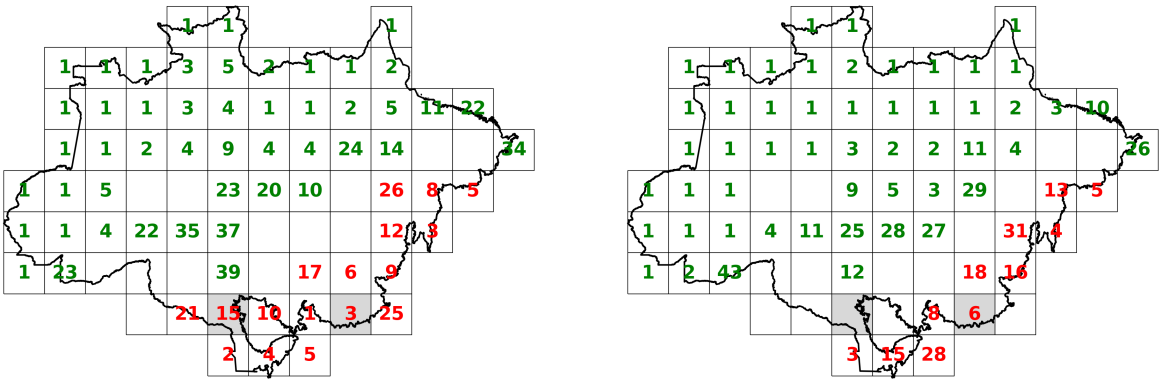
Some sites are deforested, and some sites are reforested by the planner in the absence of external transfer payments. Figures 10, 12, and 13 show the years in which one of the two controls is maximal

[25] The sites were selected at the imposed business-as-usual price under ambiguity averse. See Figure 16 in Appendix A.10 for a map of the site locations.

[26] Notice that the uncertainty adjusted γ distributions actually shift to the right when $b = \$0$.

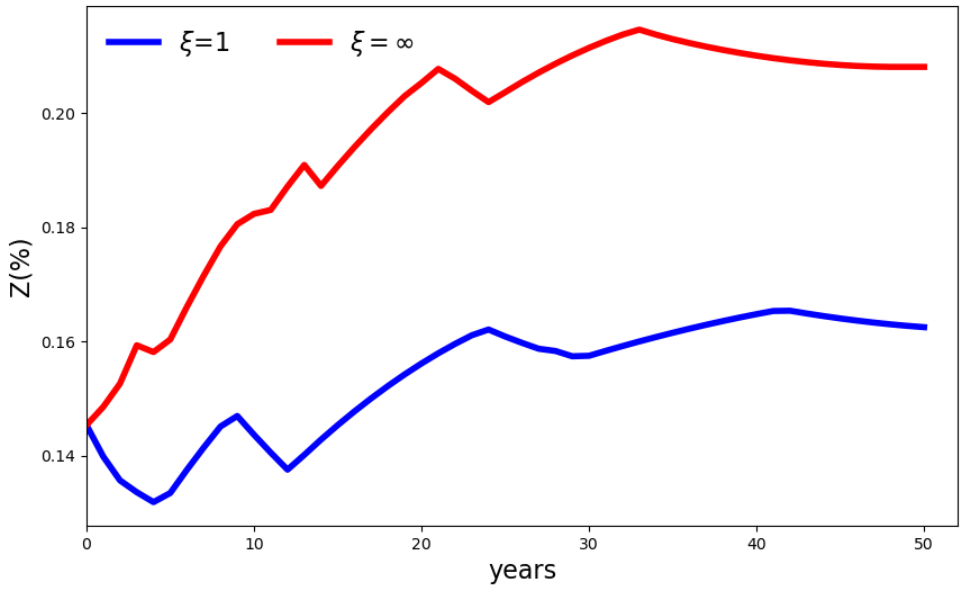
for the individual sites. Recall that $U_t^i > 0$ is when site i is being deforested, and $V_t^i > 0$ when the site i is being reforested. Only one of these can be strictly positive at any date t.

Figure 10 compares what happens when $b = 0$ under ambiguity aversion, using the same business-as-usual carbon emission price as we used under ambiguity neutrality. As examples, the two sites featured in Figure 9 are shaded in. For instance, for site 71, the change in land allocation moves from year fifteen to one in which there is no change in land allocation over the time span of 50 years. For site 74 the change is from year three under ambiguity neutrality to year six under ambiguity aversion. The delay is because the planner is particularly skeptical of the baseline distribution for the site-specific agricultural productivity as displayed in Figure 9. More generally, Figure 12 shows how the delays in deforestation and accelerations of reforestation for all of the 78 sites under ambiguity aversion, in the absence of transfer payments. Overall, the ambiguity aversion decreases substantially the land allocated to agriculture as is evident from Figure 11.



(a) ambiguity neutral (b) ambiguity averse

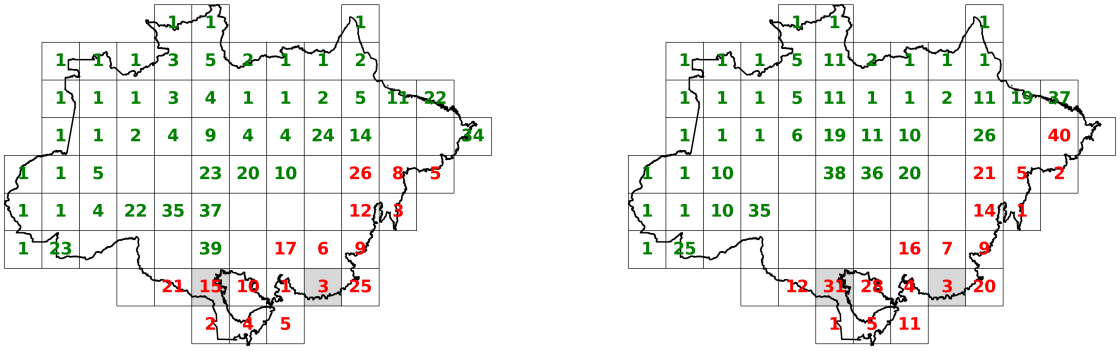
Figure 10: Spatial allocations for $b = 0$ imposing the business-as-usual carbon price $P^e = P^{ee} =$ \$7.1. The green number in a site is the year in which the rate of reforestation is maximal. The red number in each site is the year in which deforestation is maximal. Sites 71 and 74 are shaded.



(a) b=0

Figure 11: Evolution of agricultural area under ambiguity neutrality and ambiguity aversion at a common business-as-usual carbon price, $P^{ee} = 7.1$.

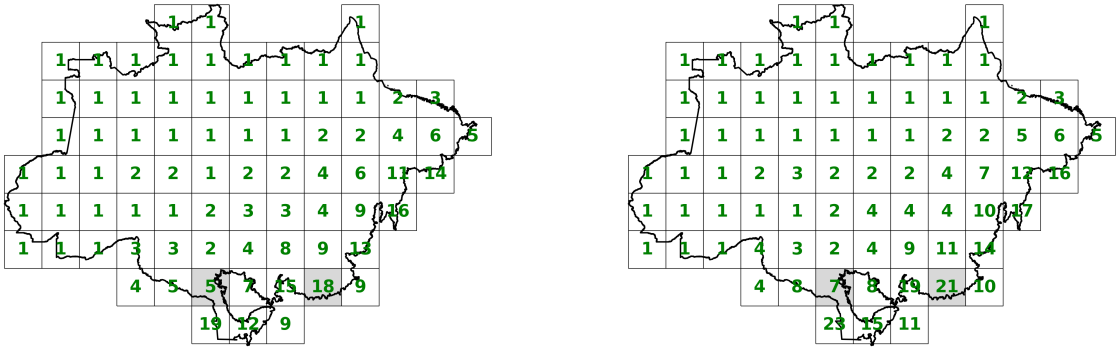
The previous results hold the business-as-usual price fixed as we introduce ambiguity aversion. As we know from Table 1, this price decreases endogenously when we impose ambiguity aversion from 7.1 to 5.3. Not surprisingly, this has a big impact on both the spatial-dynamic land allocation as we show in Figure 12. While ambiguity aversion still shifts the dates of the maximal responses, it now does so in a much more muted way than in Figure 10.



(a) ambiguity neutral (b) ambiguity averse

Figure 12: Spatial allocations for $b = 0$, $P^e = 7.1$ for ambiguity neutral and $P^e = 5.3$ for ambiguity averse, the corresponding shadow prices. The green number in a site is the year in which the rate of reforestation is maximal. The red number in each site is the year in which deforestation is maximal. Sites 71 and 74 are shaded.

As shown in Figure 13, when $b = 15$, there is only reforestation. The dates of maximal reforestation turn out show only limited sensitivity to the ambiguity aversion in this case. For instance, consider site 71, which we featured in the top panel of Figure 9 and is shaded in Figure 13. Under ambiguity neutrality, the planner's peak reforesting takes place in year five; and under ambiguity aversion, this action is delayed until year seven. For site 74, the maximal response is delayed from year eighteen to year twenty-one. The source of these and other modest delays are due to uncertainty about carbon sequestration productivities.

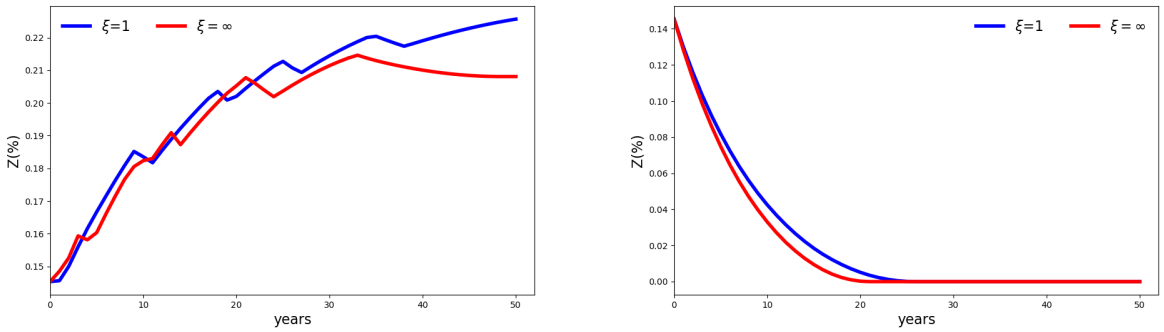


(a) ambiguity neutral (b) ambiguity averse

Figure 13: Spatial allocation for $b = 15$, using the corresponding shadow prices. The number in each site is the year in which reforestation in that site is maximal.

The left panel of Figure 14 shows a notable contrast with Figure 11. Resetting the business-as-usual price under ambiguity aversion has big impact on the aggregate land allocation dynamics

when $b = \$0$. Consistent with the spatial dynamics reported in Figures 12 and 13, Figure 14 shows that once we adjust the business-as-usual price change, there is only a very modest impact in the aggregate land allocation with and without ambiguity aversion. This holds true for $b = \$15$ as well.



(a) b=0 (b) b=15

Figure 14: Evolution of agricultural area under ambiguity neutrality and ambiguity aversion, using the corresponding shadow prices.

Finally, we consider the present values under ambiguity aversion ($\xi = 1$) in comparison to ambiguity neutrality ($\xi = \infty$) in Table 5. As should be expected, the ambiguity aversion induces smaller present values since they are computed with uncertainty adjusted probabilities. The specification of b determines either a net tax or subsidy depending on its magnitude. Over the range of transfers implied by b's in excess of \$10 per ton, the planner's discounted objective increases and is greater than that for $b = \$0$ under both ambiguity neutrality and ambiguity aversion. In this range, $\$b$ functions as a net subsidy, implying that the planner would strictly prefer outcomes under these scenarios to the business-as-usual outcome.

Table 5 also reports the discounted present value of agricultural output. For $b = 0$, the drop in the present value contribution of agriculture is over twenty percent. This drop is consistent with the reduction in land allocated to agriculture under ambiguity aversion as reported in Figure 14. In addition, the ambiguity-adjusted probability densities provide a more conservative assessment of agricultural productivity than the baseline densities. For the other choices of b in Table 5, the contributions are very small but increase under ambiguity aversion. The reason for the increase is that, under ambiguity aversion, the planner makes a more cautious assessment of the ability of the Brazilian rain forest to absorb carbon, leading to a very small increase in the land allocated to agriculture.

Table 5: Present-value decomposition - parameter ambiguity

	agricultural output value ($ 10^{11})			planner value ($ 10^{11})		
b ($)	ambiguity neutral	ambiguity aversion	percent change	ambiguity neutral	ambiguity aversion	percent change
0	3.31	2.57	-22.4	2.14	1.64	-23.4
10	0.41	0.55	33.6	2.41	2.08	-13.9
15	0.26	0.30	14.2	3.06	2.62	-14.4
20	0.20	0.23	12.9	3.75	3.19	-15.0
25	0.17	0.19	11.9	4.45	3.74	-15.8

Notes: P^a = \$41.11, the average price under the stationary distribution. Forest services are calculated using baseline shadow price (b = \$0).

Remark 7.1. *In contrast to land allocation process, Z, with parameter ambiguity, the state vector process, X, of captured carbon is disguised to the planner, because the initial condition and the dynamics of carbon captured depend on the value of γ. This has ramification for policy since we presume transfer payments are based on carbon reduction. Under ambiguity aversion, our planner uses the ambiguity-adjusted probabilities to compute these payments. In ad hoc policy-making settings distinct from our fictitious planner formulation, one could imagine differences in perspective among providers and recipients of transfers opening the door to explicit consideration of differences in their aversion to uncertainty. Also, parameter uncertainty opens the door to explicit learning that we abstract from here. Rather than being purely passive, this learning could offer the potential for experimentation as a way to gain a better understanding of site-specific productivities.*

7.4 Results with stochastic variation in agricultural prices

For our final set of results, we explore implications allowing for an explicit randomness in the agricultural price process. We generate these results using the MPC method described previously. To keep things tractable, we have the social planner assume a two state Markov process for the price process. We obtained the inputs into this specification by estimating a hidden-state Markov process with Gaussian noise as we describe in Appendix A.8. Of course, this is a rather substantial reduction in the price stochastic structure of agricultural prices, but it allows to engage in an initial explore of price randomness in a tractable way. Under the two-state Markov chain, there are two price realizations: $p^a = \$35.7, \$44.3$. The implied annual transition probabilities for staying put in each state are: .71 for the low state and .83 for the high state.[27] From Table 1 we see only a small, and not necessarily very meaningful, drop in the business-as-usual price (\$6.9 instead of \$7.1) once we introduce fluctuations in agricultural prices.[28]

[27] Appendix A.8 gives results for a second estimation of the hidden-state Markov process in which Gaussian shock variances are constrained to be the same. In this case, both realized states are lower and most of the time is spent in the higher of the two states.

[28] To construct the business-as-usual price for emissions when the agricultural prices are stochastic, we used the smoothed probabilities reported in left panel of Figure 15 in Appendix A.8 to assign the discrete states in our computations. While we used a probability .5 threshold for this assignment, many of the probabilities are actually close to zero or one.

Table 6 presents our results for present values decompositions. In addition to reporting three quantiles, for comparison we include results imposing constant agricultural prices equal to the mean under the implied stationary distribution. The notable outcomes are as follows. First, the variation while evident for the business-as-usual case ($b = \$0$) is much less consequential when $b = \$15$ and $b = \$25$ for the overall planner value. It is proportionately similar for the agricultural contribution to the value for all three choices of b. Second, imposing the mean as a constant price gives quite an accurate approximation to the median response. In this sense, the stochastic specification for agricultural prices only has modest impact on our results. Finally, Table 12 in Appendix A.8 exhibits the transfer costs and, again, these show only a modest impact of the presence uncertainty.

Table 6: Present-value decomposition with stochastic agricultural prices

	agricultural output value ($\$ 10^{11}$)	net transfers ($\$ 10^{11}$)	forest services ($\$ 10^{11}$)	adjustment costs ($\$ 10^{11}$)	planner value ($\$ 10^{11}$)
p^a = stochastic					
b = \$0					
10%	3.19	0.00	-1.07	0.06	2.10
50%	3.30	0.00	-1.05	0.06	2.19
90%	3.39	0.00	-1.03	0.06	2.27
b = \$15					
10%	0.24	1.99	0.92	0.19	2.97
50%	0.26	2.00	0.92	0.19	2.99
90%	0.27	2.00	0.92	0.19	3.00
b = \$25					
10%	0.16	3.48	0.96	0.27	4.33
50%	0.18	3.48	0.96	0.28	4.34
90%	0.19	3.48	0.96	0.28	4.35
p^a = \$41.1					
b = 0	3.31	0.00	-1.10	0.06	2.14
b = 15	0.26	2.02	0.95	0.17	3.06
b = 25	0.17	3.54	1.00	0.26	4.45

Notes: Forest services were calculated using baseline shadow price ($b = \$0$). The quantiles were computed based on two hundred simulations. The agricultural price $P^a = 41.1$ is the mean under the stationary distribution. Shadow prices are $P^{ee} = 7.1$ for $P^a = 41.1$ and $P^{ee} = 6.9$ for P^a = stochastic.

Remark 7.2. *As we have seen, the impact of making the agricultural prices stochastic is very modest. One possible way to amplify this impact is to introduce specification uncertainty into the analysis. Anderson et al. (2003) propose a recursive way to make this adjustment for Markov specifications that include continuous-time jump components. This approach also uses a relative entropy divergence, relative to a baseline specification, to form penalties that limit the potential misspecifications that are explored by the planner.*[29] *Thus, the MPC approach could potentially be*

[29] Anderson et al. (2003) show the connection between this form of misspecification concern and recursive utility-based risk adjustments, extending insights from robust control theory.

extended to incorporate this extension, opening the door to greater adjustments for uncertainty.

8 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 project out the implications of "business-as-usual" for land allocation in the Brazilian amazon. Formally, we deduce the shadow price for emissions that would justify historical deforestation. When we impose this shadow price looking forward, the implied land allocation we find that the resulting deforestation would eventually threaten the survival of the Amazon as a tropical forest. Second, we explore implications of augmenting this shadow price with transfer payments. We find that transfer prices as low as \$15 per ton of $C0_2e$ lead to a substantial net reforestation and carbon capture in an efficient dynamic spatial land allocation. Furthermore, a time invariant carbon price of \$25 would generate enough monetary transfers to fully compensate Brazil for the loss of output in the reforested areas. As a third exercise we consider the impact of robustly optimal controls on the spatial outcomes when the capacity of each site to capture carbon as well as agricultural productivity is uncertain to the planner.

Our results suggest international carbon payments of \$25 USD/ton can reduce emissions by about 20GtCO2e in 15 years and by about 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. 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 3 suggests that optimal management of the Brazilian Amazon can deliver about 10% of this total at a much lower effective cost.

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 resiliency including possibility that Amazon deforestation triggers a tipping point with broad based consequences (Steffen et al. (2018) and Flores et al. (2024)). Finally, we do not account for the direct effect of deforestation in one site on forests in other sites.[30] These are important considerations for future research.

[30] See Araujo et al. (2023) for an estimate

A Data construction

A.1 Total available area

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.

A.2 Carbon absorption

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[31]. 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). In Appendix C we exposit how we use the data to obtain a baseline distribution of the vector of site-specific carbon absorption productivities, $(\gamma^1, \ldots, \gamma^I)$.

A.3 Carbon depreciation

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 $\alpha = 1 - (1 - 0.99)^{1/100} = .045$.

A.4 Emissions contributed by agriculture

The parameter κ is calibrated using the agricultural net annual emission data at the state level available from the system SEEG.[32] We use $\kappa = 2.0942$, which is 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.

A.5 Cattle farming productivities

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

[31] (Santoro and Cartus, 2021)

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

an overvaluation of the contribution of cattle farming in the Amazon to the Brazilian economy.[33] 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 Appendix C we exposit how we derive a baseline distribution for the vector of site-specific cattle farming productivities, $(\theta^1, \ldots, \theta^I)$.

A.6 Discount rate (δ) and adjustment costs cost (ζ)

We use the discount rate $\delta = 0.02$ and calibrate $\zeta = 1.66e-4$ 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.

A.7 Initial values: Z_0^i, X_0^i

The approach for computing the initial value for the agricultural area, Z_0^i, 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 assumed to be given by $X_0^i = \gamma^i(\bar{Z}^i - Z_0^i)$, 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).

A.8 Agricultural prices

We use a data series on monthly deflated cattle prices (reference date 01/2017).[34] from 1995, the year in which the Real Plan stabilized the Brazilian currency, until 2017.

For the model inputs, we fit a two-state Markov process as a hidden state Markov chain with with Gaussian noise. We estimated two versions of this model using the **hmmlearn** package in python. This package provides a collection of software tools for analyzing Hidden State Markov Models. In estimation, the hidden states were initialized in the implied stationary distribution of the transition probabilities through an iterative process. The implied calibration we used for results reported in the main body of the paper allowed for the normally distributed variances to be different depending on the state. We also considered a specification in which the variances are

[33] 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

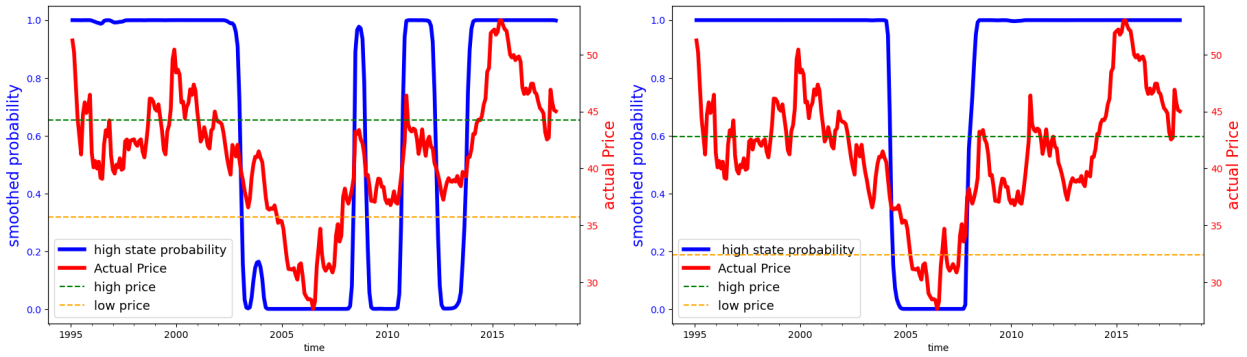
[34] Commodity prices from SEAB-PR. Secretaria da Agricultura e do Abastecimento do Estado do Paraná (SEAB-PR). 2021. "Preço Médio - Recebido pelo Agricultor: boi gordo, arroz (em casca), cana-de-açúcar, milho, mandioca, 1990-2021." Secretaria da Agricultura e do Abastecimento do Estado do Paraná, Departamento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

the same. The state realizations and transition probabilities for the two specifications are given in Table 7.

	distinct variances		a common variance	
	low price	high price	low price	high price
	35.71	44.26	32.44	42.78
s.d.	0.105	0.075	0.088	0.088
		transition probabilities		
	low	high	low	high
low	0.706	0.294	0.766	0.234
high	0.171	0.829	0.046	0.954

Table 7: Estimates for the hidden-state Markov models

The smoothed probabilities for both models are given in Figure 15. The more constrained estimation picks lower values for both states but assumes the process spends most of its time the higher of the two state.



(a) distinct variances (b) common variance

Figure 15: Smoothed probabilities for the two hidden state Markov chain models

Table 8 reports the likelihoods and AIC and BIC model selection diagnostics for both models. The AIC criterion picks the less constrained of the two models and the BIC criterion just the opposite.

	distinct variances	common variance
log likelihood	270.16	268.04
aic	-528.32	-526.08
bic	-502.97	-504.36

Table 8: Likelihood ratios and information criteria for the hidden state Markov chain estimation

In Table 9, we report the counterpart to Table 6 constructed using the implied calibration the same variances for each state. The differences between results are modest.

Table 9: Present-value decomposition with stochastic agricultural prices

	agricultural output value ($\$ 10^{11}$)	net transfers ($\$ 10^{11}$)	forest services ($\$ 10^{11}$)	adjustment costs ($\$ 10^{11}$)	planner value ($\$ 10^{11}$)
p^a = stochastic					
b = \$0					
10%	3.22	0.00	-1.01	0.05	2.19
50%	3.34	0.00	-1.00	0.05	2.29
90%	3.43	0.00	-0.97	0.05	2.37
b = \$15					
10%	0.24	2.00	0.84	0.19	2.89
50%	0.26	2.00	0.84	0.19	2.91
90%	0.26	2.00	0.84	0.19	2.91
b = \$25					
10%	0.16	3.48	0.96	0.27	4.33
50%	0.18	3.48	0.96	0.28	4.34
90%	0.19	3.48	0.96	0.28	4.35
p^a = \$41.1					
b = 0	3.31	0.00	-1.10	0.06	2.14
b = 15	0.26	2.02	0.95	0.17	3.06
b = 25	0.17	3.54	1.00	0.26	4.45

Notes: Hidden state Markov chain model with a common variance. Forest services were calculated using baseline shadow price (b = \$0). The quantiles were computed based on two hundred simulations. For this stochastic specification of the agricultural prices, P^{ee} = \$6.3.

A.9 Transfer costs

In this subsection we report in Table 10 and Table 11 transfer costs under ambiguity aversion and Table 12 transfer costs under stochastic variation in agricultural prices.

Table 10: Transfer costs under ambigutiy - 15 years

	ambiguity neutral			ambiguity aversion		
b ($)	net captured emissions (billion tons of CO2e)	discounted net transfers ($ 10^{11})	discounted effective cost ($ per ton of CO2e)	net captured emissions (billion tons of CO2e)	discounted net transfers ($ 10^{11})	discounted effective cost ($ per ton of CO2e)
0	-9.89	0.00	NaN	-9.59	0.00	NaN
10	5.17	0.43	2.87	4.24	0.35	2.55
15	6.69	0.84	5.08	5.63	0.70	4.65
20	7.65	1.29	7.36	6.47	1.08	6.77
25	8.21	1.74	9.60	6.97	1.47	8.88

Notes: Agricultural price P^a = \$41.1, which is the mean of the agricultural price in the stationary distribution.

Table 11: Transfer costs under ambiguity - 30 years

	ambiguity neutral			ambiguity aversion		
b ($)	net captured emissions (billion tons of CO2e)	discounted net transfers ($ 10^{11})	discounted effective cost ($ per ton of CO2e)	net captured emissions (billion tons of CO2e)	discounted net transfers ($ 10^{11})	discounted effective cost ($ per ton of CO2e)
0	-15.25	0.00	NaN	-15.08	0.00	NaN
10	11.91	0.87	3.21	10.18	0.74	2.93
15	14.10	1.57	5.37	12.57	1.39	5.03
20	14.75	2.23	7.43	13.15	1.97	6.97
25	15.08	2.87	9.47	13.29	2.51	8.86

Notes: Agricultural price P^a = \$41.1, which is the mean of the agricultural price in the stationary distribution.

Table 12: Transfer costs with stochastic agricultural prices

	15 years			30 years		
	net captured emissions (billion tons of CO2e)	discounted net transfers ($ 10^{11})	discounted effective cost ($ per ton of CO2e)	net captured emissions (billion tons of CO2e)	discounted net transfers ($ 10^{11})	discounted effective cost ($ per ton of CO2e)
p^a = stochastic						
b = $0						
10%	-9.28	0.00	NaN	-14.41	0.00	NaN
50%	-8.52	0.00	NaN	-13.79	0.00	NaN
90%	-7.82	0.00	NaN	-13.33	0.00	NaN
b=$15						
10%	6.22	0.77	4.91	13.89	1.53	5.41
50%	6.25	0.77	5.21	13.91	1.53	5.53
90%	6.29	0.78	5.43	13.94	1.54	5.62
b=$25						
10%	7.61	1.59	9.27	14.78	2.78	9.49
50%	7.63	1.60	9.78	14.79	2.78	9.70
90%	7.65	1.60	10.14	14.81	2.78	9.86
p^a = $41.1						
b=0	-9.89	0.00	NaN	-15.25	0.00	NaN
b=15	6.69	0.84	5.08	14.10	1.57	5.37
b=25	8.21	1.74	9.60	15.08	2.87	9.47

Notes: The quantiles were computed based on two hundred simulations. The agricultural price P^a = $41.1, which is the mean under the stationary distribution. Shadow prices are P^{ee} = 7.1 for P^a = 41.1 and P^{ee} = 6.9 when P^a is stochastic.

A.10 Catalog of the sites

Figure 16 lists the site numbers in the Amazon map as a reference.

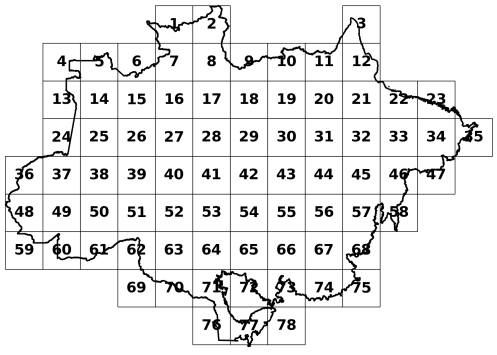


Figure 16: Locations of the different site numbers

A.11 Alternative values of ξ

In this section, we report results for parameter uncertainty with $\xi = 2$. The calculated business as usual price is \$6.1. Table 13 shows the present values under $\xi = 2$ in comparison to $\xi = \infty$. Figure 17 and Figure 18 shows the baseline and ambiguity-adjusted distributions.

Table 13: Present-value decomposition - parameter ambiguity $\xi = 2$

	agricultural output value (\$ 10^{11})			planner value (\$ 10^{11})		
b (\$)	$\xi = \infty$	$\xi = 2$	$\xi = 1$	$\xi = \infty$	$\xi = 2$	$\xi = 1$
0	3.31	2.87	2.57	2.14	1.86	1.64
10	0.41	0.50	0.55	2.41	2.22	2.08
15	0.26	0.28	0.30	3.06	2.81	2.62
20	0.20	0.22	0.23	3.75	3.43	3.19
25	0.17	0.18	0.19	4.45	4.06	3.74

Notes: P^a = \$41.11, the average price under the stationary distribution. Forest services are calculated using baseline shadow price (b = \$0).The business as usual price P^{ee} = 6.1 for $\xi = 2$.

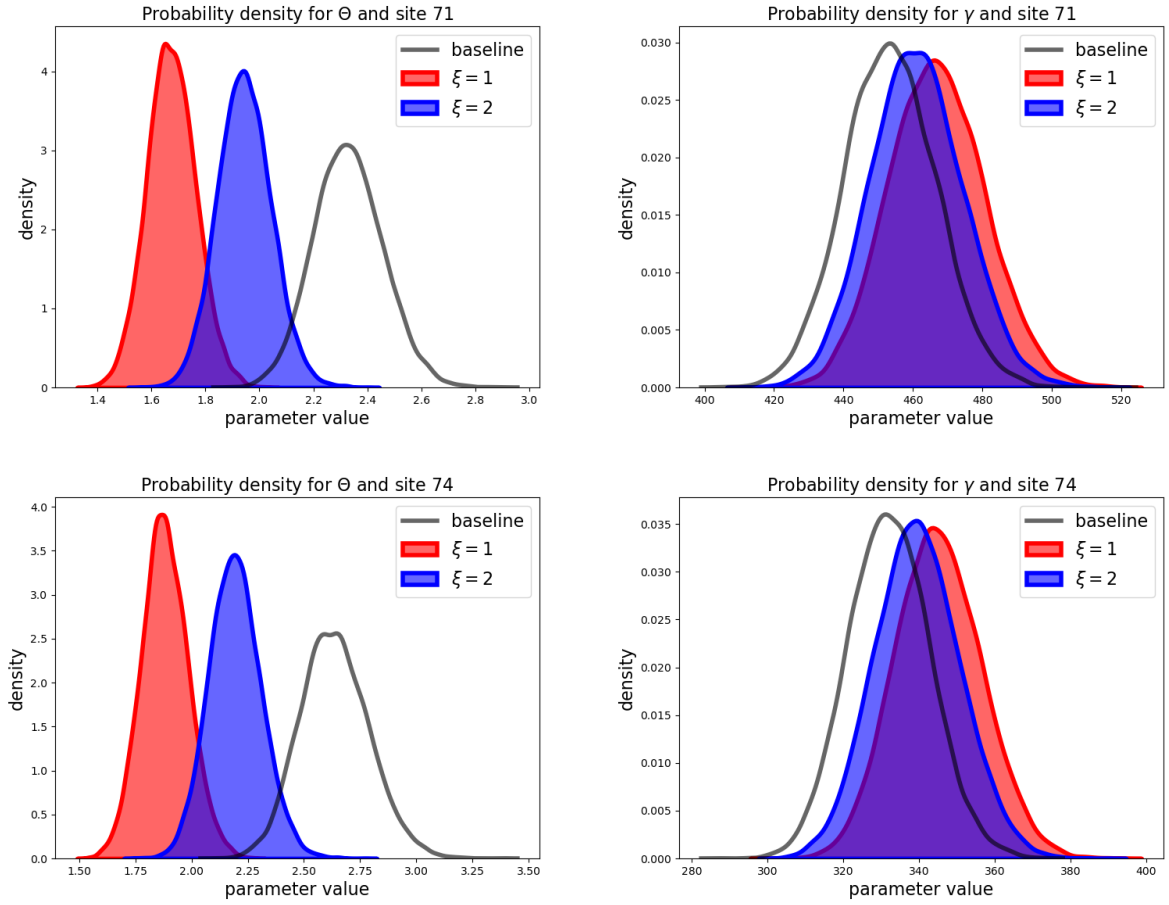


Figure 17: Baseline and ambiguity adjusted densities for $b = 0$.

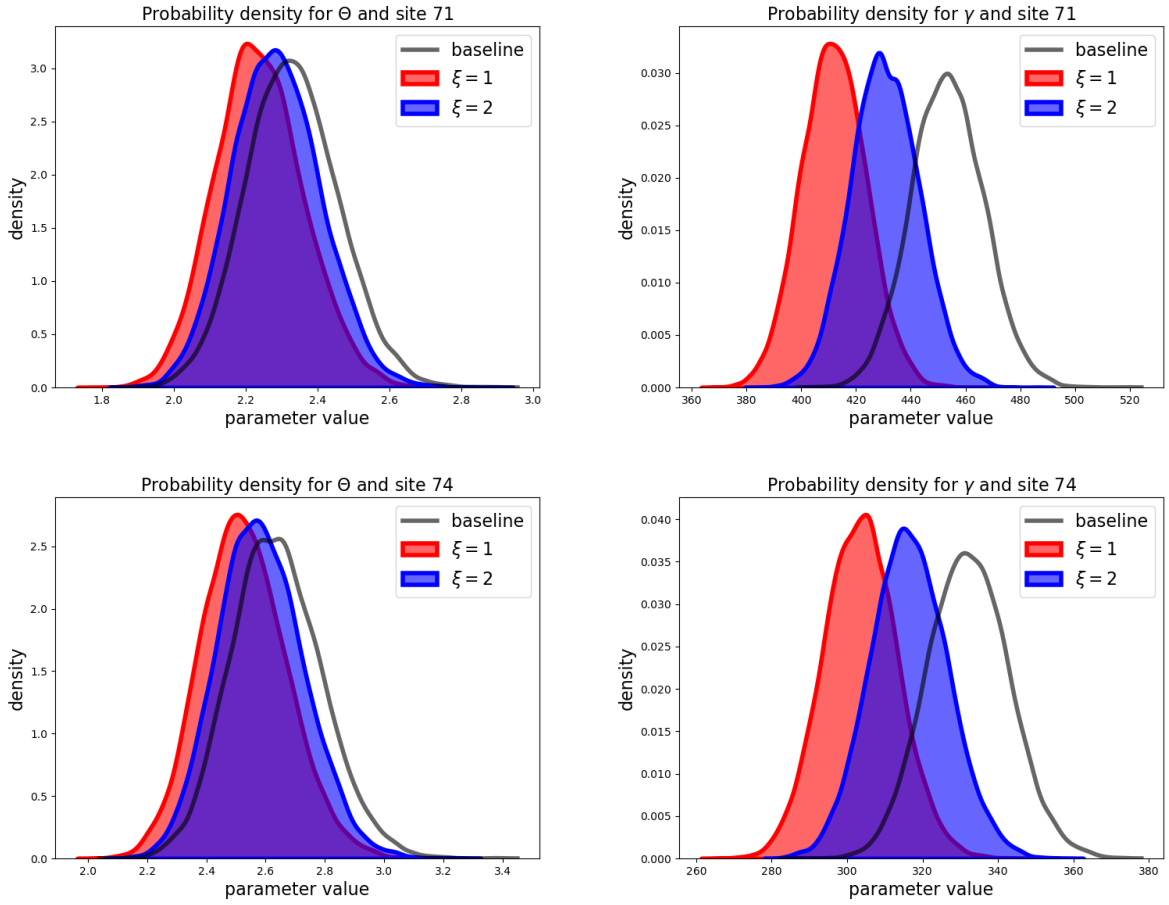


Figure 18: Baseline and ambiguity adjusted densities for $b = 15$.

B Model discretization

In order to obtain numerical solutions for the social planner problem, we solve the following discrete-time approximation, for a finite horizon of $T = 200$ years[35]:

$$\max_{\{U_t, V_t\}_{t=1}^{T}} \sum_{t=0}^{T} e^{-\delta t} \left[-P^e \sum_{i=1}^{I} \kappa Z_{t+1}^i - (X_{t+1}^i - X_t^i) + P^a \sum_{i=1}^{I} \theta^i Z_{t+1}^i - \frac{\zeta}{2} \left(\sum_{i=1}^{I} U_t^i + V_t^i \right)^2 \right], \qquad (9)$$

[35] Since period-payoff can be bounded by a constant, given the discount rates we use, the loss in precision for trajectories in the first 30 years, which is our period of interest, is small.

subject to the initial conditions in A.7 and the constraints:

$$X^i_{t+1} = X^i_t - \gamma^i U^i_t - \alpha\left[X^i_t - \gamma^i(\bar{z}^i - Z^i_t)\right] \qquad \forall i = 1,\dots I \text{ and } t = 0,\dots,T \tag{10}$$

$$Z^i_{t+1} = Z^i_t + U^i_t - V^i_t \qquad \forall i = 1,\dots I \text{ and } t = 0,\dots,T \tag{11}$$

$$U^i_t \geqslant 0, \quad V^i_t \geqslant 0 \qquad \forall i = 1,\dots I \text{ and } t = 0,\dots,T \tag{12}$$

C Benchmark distributions

Equation 6 gave the formula for constructing measurement of site-specific productivities from regression coefficients and measurements of municipality attributes. In what follows, we first outline the municipality regression models used for γ and θ, and then we describe the procedure that we used constructing baseline Bayesian posteriors for the regression coefficients.

C.1 $\theta's$

To construct a measurement of the θ's, we run the regression specification below with probabilistic output.[36]

$$\log(\text{Slaughter value}) = R_\theta \beta_\theta + \epsilon_\theta \tag{13}$$

where

$$\begin{aligned} R_\theta\beta_\theta \overset{\text{def}}{=} \quad & \beta^0_\theta + \beta^1_\theta(\text{historical\_precip}) + \beta^2_\theta(\text{historical\_temp}) + \beta^3_\theta(\text{historical\_temp}^2) \\ & + \beta^4_\theta(\text{lat}) + \beta^5_\theta(\text{lat}^2) + \beta^6_\theta \log(\text{cattleSlaughter\_farmGatePrice}) + \beta^7_\theta(\text{distance}) \end{aligned}$$

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 weighting observations by the 2017 pasture area in each municipality.

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

[36] We standardize the regressors prior to the posterior estimation.

C.2 $\gamma's$

We calculate average of CO2 density (MG/ha) for each municipality and run the following regression:[37]

$$\log(\text{co2e\_ha}) = R_\gamma \beta_\gamma + \varepsilon_\gamma$$

where

$$\begin{aligned} R_\gamma \beta_\gamma \stackrel{\text{def}}{=} \quad & \beta_0^\gamma + \beta_1^\gamma(\log(\text{historical\_precip})) + \beta_2^\gamma(\log(\text{historical\_temp})) \\ & \beta_3^\gamma(\log(\text{lat})) + \beta_4^\gamma(\log(\text{lon})). \end{aligned}$$

C.3 Posterior estimation

To estimate the benchmark posterior distribution π, 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.

Using a weighted regression for the model given by 13 with a Gaussian error term is equivalent to assuming

$$Y_\theta \;=\; R_\theta \beta_\theta + \varepsilon_\theta, \quad \varepsilon_\theta \sim \mathcal{N}(0,\, \sigma_\theta^2 W_\theta^{-1}), \tag{14}$$

where W_θ is the diagonal matrix of weights. For the γ case, $W_\gamma = I$.

We assume that the priors/posteriors for parameters $(\beta_\theta, \sigma_\theta^2)$ are the familiar conjugate form:

$$\beta_\theta \mid \sigma_\theta^2 \;\sim\; \mathcal{N}(m,\, \sigma_\theta^2 Q^{-1}) \tag{15}$$

$$\sigma_\theta^2 \;\sim\; \text{Inv-Gamma}(a,\, b) \tag{16}$$

where

$$\begin{aligned} Q &= R_\theta' W_\theta R_\theta + Q_0, \\ m &= Q^{-1}(R_\theta' W_\theta y_\theta + Q_0 m_0), \\ a &= a_0 + \frac{n}{2}, \\ b &= b_0 + \frac{1}{2}(y_\theta' W_\theta y_\theta + m_0' Q_0 m_0 - m' Q m), \end{aligned}$$

and Q_0, m_0, a_0, b_0 are prior inputs. We impose the familiar improper priors:

$$Q_0 = 0 \quad m_0 = 0 \quad a_0 = 0 \quad b_0 = 0,$$

which implies that the posteriors inputs are familiar regression statistics. In the following tables, we present quantiles for the posterior distributions described above:

[37] We again standardize the regressors.

Table 14: Quantiles for θ posterior estimation

	β_0^θ	β_1^θ	β_2^θ	β_3^θ	β_4^θ	β_5^θ	β_6^θ	β_7^θ
10%	3.899	-0.189	1.274	-5.250	5.466	-7.871	0.454	-0.132
50%	3.943	-0.135	3.201	-3.343	6.607	-6.732	0.495	-0.091
90%	3.986	-0.082	5.080	-1.404	7.729	-5.576	0.537	-0.051

Table 15: Quantiles for γ posterior estimation

	β_0^γ	β_1^γ	β_2^γ	β_3^γ	β_4^γ
10%	5.918	0.055	-0.172	0.148	-0.313
50%	5.936	0.079	-0.146	0.177	-0.291
90%	5.955	0.102	-0.119	0.206	-0.269

D Hamiltonian Monte Carlo

As is standard in applications of MCMC, we use numerical simulation to compute the probability distribution induced by the density of (7). While we have partially analytic expression for the numerator, the integration in the denominator is more problematic. This challenge is a typical starting point for MCMC computations.

Although we are interested in only the distribution for the regression coefficients, for numerical tractability we use the conditional normal inverse gamma specification in our computations, which includes the regression error variances. We then construct numerically the implied marginal distribution for the regression coefficients consistent with formula (7).

This means that we sample from

$$\exp\left[-\frac{1}{\xi}f(\mathbf{d},\boldsymbol{\beta})\right]d\pi(\beta_\theta,\sigma_\theta^2|R_\theta,y_\theta)d\pi(\beta_\gamma,\sigma_\gamma^2|R_\gamma,y_\gamma) \tag{17}$$

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

$$\mathcal{U}(\boldsymbol{\rho}) = \frac{1}{\xi}f(\mathbf{d},\boldsymbol{\beta}) - \log d\pi(\beta_\theta,\sigma_\theta^2|R_\theta,y_\theta) - \log d\pi(\beta_\gamma,\sigma_\gamma^2|R_\gamma,y_\gamma) \tag{18}$$

$$\begin{aligned} &= \frac{1}{\xi}f(\mathbf{d},\boldsymbol{\beta}) - \log d\pi(\beta_\theta|\sigma_\theta^2,R_\theta,y_\theta) - \log d\pi(\beta_\gamma|\sigma_\gamma^2,R_\gamma,y_\gamma) \\ &\quad - \log d\pi(\sigma_\theta^2|R_\theta,y_\theta) - \log d\pi(\sigma_\gamma^2|R_\theta,y_\gamma) \end{aligned} \tag{19}$$

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:

$$\mathcal{H}(\boldsymbol{\rho}, \omega) := \mathcal{U}(\boldsymbol{\rho}) + \frac{1}{2}\omega' M^{-1}\omega \tag{20}$$

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:

$$\frac{d\boldsymbol{\rho}}{dt} = \frac{\partial \mathcal{H}}{\partial \omega} \tag{21}$$

$$\frac{d\omega}{dt} = -\frac{\partial \mathcal{H}}{\partial \boldsymbol{\rho}} \tag{22}$$

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 17 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 $||\boldsymbol{\varphi}_{(t+1)} - \boldsymbol{\varphi}_{(t)}||_\infty < 0.001$.

D.1 Computational implementation details

To sample from 17, 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

Anderson, Evan W, Lars Peter Hansen, and Thomas J Sargent. 2003. A quartet of semigroups for model specification, robustness,prices of risk, and model detection. *Journal of the European Economic Association* 1:68–123.

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. 2023a. DETER-ing Deforestation in the Amazon: Environmental Monitoring and Law Enforcement. *American Economic Journal: Applied Economics* 15 (2):125–156.

Assunção, Juliano, Robert McMillan, Joshua Murphy, and Eduardo Souza-Rodrigues. 2023b. Optimal environmental targeting in the amazon rainforest. *The Review of Economic Studies* 90 (4):1608–1641.

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. 2023. The economics of tropical deforestation. *Annual Review of Economics* 15:723–754.

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* 14 (2):651–687.

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.

Flores, Bernardo M., Encarni Montoya, Boris Sakschewski, Nathália Nascimento, Arie Staal, Richard A. Betts, Carolina Levis, David M. Lapola, Adriane Esquível-Muelbert, Catarina Jakovac, Carlos A. Nobre, Rafael S. Oliveira, Laura S. Borma, Da Nian, Niklas Boers, Susanna B. Hecht, Hans ter Steege, Julia Arieira, Isabella L. Lucas, Erika Berenguer, JoséA. Marengo, Luciana V. Gatti, Caio R. C. Mattos, and Marina Hirota. 2024. Critical transitions in the Amazon forest system. *Nature* 626 (7999):555–564.

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. 2001. Robust Control and Model Uncertainty. *The American Economic Review* 91:60–66.

———. 2023. Risk, Ambiguity, and Misspecification: Decision Theory, Robust Control, and Statisitics. *Journal of Applied Econometrics* online:1–31.

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. *Science Advances* 4 (2):eaat2340.

Maccheroni, Fabio, Massimo Marinacci, and Aldo Rustichini. 2006. Ambiguity aversion, robustness, and the variational representation of preferences. *Econometrica* 74:1447–1498.

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.