

When Clouds Go Dry: an integrated model of deforestation, rainfall, and agriculture

Rafael Araujo*

November, 2024

Abstract

I develop a discrete choice model of deforestation with spatially heterogeneous endogenous climate. Using pixel-level data for the Amazon Rainforest and the Brazilian state of Mato Grosso, one of the most important agricultural hubs in the world, I estimate the model and consider a counterfactual where farmers are allowed to deforest protected areas. Deforestation reduces rainfall and increases temperatures for downwind farmers, creating cascading effects on a network of externalities. Ignoring climate feedback loops significantly underestimates the costs of deforestation, highlighting the importance of considering climate endogeneity beyond the production-emissions-temperature framework.

JEL: *Q23, Q56, Q15, C35, L73*

Keywords: *Climate, Deforestation, Amazon, Land Use, Discrete Choice Model*

*São Paulo School of Economics FGV EESP (e-mail: rafael.araujo@fgv.br). I thank Francisco Costa, Marcelo Sant'Anna, Arthur Bragança, Teevrat Garg, Juliano Assunção, Sophie Mathes, Robert Heilmayr, José A. Scheinkman, and Bruno Barsanetti for their invaluable contributions. I also thank seminar participants at FGV EPGE, FGV EESP, Climate Policy Initiative, Ridge, AERE OSWEET, SED Winter, and the committee of the Wallace E. Oates Outstanding Doctoral Dissertation Award. All errors are my own. This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior Brasil (CAPES) Finance Code 001, National Council for Scientific and Technological Development (CNPq), Norway's International Climate and Forest Initiative (NICFI) (18/0028), and the Gordon and Betty Moore Foundation.

1 Introduction

Climate is endogenous to people's actions. The integration of climate and human action models lead to a large literature on the endogeneity of productivity to temperature through emissions resulting from economic production (Nordhaus, 1991, 2019). However, outside this literature connecting production, emissions, and temperature - predominantly through macro-environmental models - the endogeneity of climate has been ignored. This includes the literature on deforestation (Souza-Rodrigues, 2019; Balboni et al., 2023) despite longstanding evidence from the natural sciences showing deforestation's significant impacts on climate (Salati et al., 1979; Nobre et al., 1991; Sampaio et al., 2007; Flores et al., 2024). In this paper, I build a micro-environmental model where land use and climate interact. Specifically, climate affects agricultural productivity, which affects deforestation, which in turn affects climate.

Salati et al. (1979) is the seminal work demonstrating that tropical forests can affect climate. Salati et al. (1979) showed that most of the Amazon's rainfall is produced by the forest itself. This occurs because tropical forests play a crucial role in the water cycle. Solar energy evaporates ocean water, which the atmosphere transports, creating rain over continental land. In the presence of tropical forests, rainwater is continuously used by trees in photosynthesis, with transpiration from the trees recycling the rainwater and recharging atmospheric humidity as it moves inland. These are the *flying rivers* of tropical forests (Marengo et al., 2004). Nobre et al. (1991) later showed that the Amazon influences not only its own rainfall but also South America's rainfall. The implication of this mechanism is that deforestation reduces rainfall.

As deforestation impacts rainfall and other climate outcomes, it subsequently affects key inputs for the agricultural sector, which is the main driver of tropical deforestation (Souza-Rodrigues, 2019; Araujo et al., 2020; Domínguez-Iino, 2021; Hsiao, 2021; Pendrill et al., 2022; Farrokhi et al., 2023). Since the flying rivers mechanism operates through atmospheric transport, the deforestation decisions of farmers upwind influence climate conditions for farmers downwind. This creates a network of externalities closely linked to the configuration of forest cover and atmospheric circulation. As a result, this network connects farmers separated by hundreds of kilometers in an economy where each farmer's productivity is endogenously affected by the deforestation decisions of others.

To model this externality network, I connect a discrete choice model of land use with a climate model of atmospheric circulation. In the discrete choice model, each farmer decides whether to deforest its plot and what to produce on it. Each potential agricultural production has a return that depends on climate, which is a function of the upwind forest

cover configuration, referred to here as *upwind forest*. This upwind forest is derived from the climate model, which tracks the paths of atmospheric trajectories from the ocean to a farm, storing information on land use along the way. A weighted sum of upwind forest pixels represents my measure of upwind forest.

To integrate the discrete choice model with the atmospheric circulation model, I use the fact that the discrete choice model implies a probability of the farmer choosing to maintain forest cover. Assigning this probability as the forest cover of a farmer means that each farmer's probability is a function of the probabilities of other farmers. Formally, this creates a map of forest probabilities into forest probabilities. Thus, an equilibrium in this economy, which is a Nash equilibrium, is a fixed point: a vector of land use decisions consistent with the climate it generates. I first demonstrate that an equilibrium always exists in this economy. I then derive conditions for this probability map to be a contraction, which implies a unique equilibrium. This condition essentially places a bound on the externality effect that one farmer can have on others.

As an application of my model, I study the importance of protecting Indigenous Territories in the Amazon on the agricultural sector of the Brazilian state of Mato Grosso, one of the most important agricultural hubs in the world. Even though the model has this complicated network structure it can be estimated with standard micro-applied techniques.

I start by testing the mechanisms of the climate model, establishing that upwind forest affects both rainfall and temperature. Using three-dimensional data on wind velocity and direction (Copernicus, 2017), I construct atmospheric circulation patterns from 1985 to 2017 across South America. I then combine these trajectories with land use data from Mapbiomas, along with rainfall and temperature data from Copernicus (2017), to build a panel dataset on upwind forest across the Amazon rainforest and climate outcomes in Mato Grosso. Since this paper centers on climate endogeneity, this endogeneity introduces challenges for identification, including potential reverse causality. To address this, I develop an instrumental variable that leverages variation exclusively from changes in atmospheric circulation to identify the parameters of the climate model. I find a strong and significant effect of upwind forest on rainfall and temperature: a one standard deviation increase in upwind forest raises average rainfall by 10% and lowers average temperature by 0.12 °C.

I then estimate the parameters of the discrete choice model of land use. This model produces regressions that can be estimated with land use data and the panel dataset on upwind forest. I use detailed land use data from Simoes et al. (2020), which classifies land use at a 250-meter resolution in the Brazilian state of Mato Grosso. Importantly, this dataset incorporates information throughout the year to construct the land use classification algorithm, enabling it to distinguish between forest, pasture, crops, and double cropping,

the latter representing two harvests within a single year. I focus on the years 2008 and 2017, a period of regulatory stability (Burgess et al., 2019), and on data on private land. As with the climate model, climate endogeneity presents identification challenges. To address this, I again use variation in upwind forest caused by atmospheric circulation changes to instrument for the actual upwind forest. I find a significant effect of upwind forest on the returns of crops and double-cropping. However, due to the model's structure, understanding the true interpretation of the land use model parameters requires simulating a counterfactual and finding a new equilibrium using the contraction result.

In my counterfactual exercise, I simulate a policy rollback of the protection status of the Indigenous Territories in the Xingu Basin.¹ The Xingu, located within the Amazon Basin in Brazil, is one of the earliest initiatives in indigenous land protection, covering 141,000 square kilometers—twice the size of Portugal—and situated between rapidly advancing agricultural frontiers in Brazil. This type of policy where protection is rolled back, known as Protected Area Downgrading, Downsizing, and Degazettement (PADDD), has gained traction globally (Pack et al., 2016; Golden Kroner et al., 2019).

I apply the parameters of the land use model, estimated using data on private land, to the Xingu region, effectively rolling back its protection status. Deforestation in the Xingu would lead to cascading effects on climate and land use decisions, both within and outside the region. In the new equilibrium, 47% of the Xingu is deforested. As a consequence, some regions in Mato Grosso experience a loss of over one-fifth of their average rainfall and an increase in temperature of 0.3 °C. The increase in forest choice probability outside the Xingu, which I interpret as farmers abandoning land, totals 60,000 square kilometers. By summing the benefits of deforesting the Xingu—increased land availability for agriculture—with the costs—deterioration of climate conditions for agricultural production—I show that accounting for climate feedback offsets 40% of the benefit of expanding agricultural land. That is, ignoring the climate's endogeneity leads to a severe underestimation of the costs of deforestation. However, these aggregate effects mask significant spatial heterogeneity, with the winners and losers of such a policy being determined by atmospheric circulation.

This paper connects with various strands of literature both within and outside economics. Most broadly, it aligns with the literature on the endogeneity of climate. Most works in this area employ macro-environmental models, particularly Integrated Assessment Models, as pioneered by Nordhaus (1991, 2019), to quantify policy trade-offs in emissions reduction. More recently, this literature has begun to incorporate spatial dimensions using Dynamic Spatial Integrated Assessment Models (Desmet and Rossi-Hansberg, 2024; Cruz and Rossi-Hansberg,

¹Pronounced as [shing-goo]

2024). My contribution to this literature is to broaden the scope of factors considered as endogenously affecting climate, moving beyond the production-emissions-temperature nexus.

The literature on the economics of deforestation has largely overlooked the endogeneity of climate. Studies such as Souza-Rodrigues (2019), Araujo et al. (2020), Domínguez-Iino (2021), and Hsiao (2021) model drivers of deforestation but completely ignore the feedback effects of deforestation on climate, leading to an underestimation of deforestation costs.² Similarly, papers evaluating the effectiveness of policies to curb deforestation underestimate the benefits of these policies for the same reason (Assunção et al., 2013, 2019; Burgess et al., 2019). A notable exception is Grosset et al. (2023), which uses micro-applied tools to show that a large tree-planting program in the US significantly impacted local climate and farmers' decisions. My paper contributes to this literature by developing a model to assess counterfactual deforestation scenarios in a framework where climate is endogenous.

There is a broader literature on the impacts of climate on agriculture which is an important source that informs many of the parameters in macro-environmental models. Focused on short-term weather variations, this literature aims at a deeper discussion on identification challenges in the relationship between agricultural productivity and rainfall or temperature (Schlenker and Roberts, 2009; Dell et al., 2014; Yang et al., 2024). My contribution to this literature is to introduce a modeling approach that addresses the inevitable reverse causality of land use and climate/weather inherent in such empirical analyses.

Finally, this paper builds on a large literature in natural sciences on the impacts of deforestation on climate. This literature is highly diverse. For example, there are papers testing physical implications of such systems, with Salati et al. (1979) being the seminal work on the field, followed by more recent papers such as Gatti et al. (2021). Much of the literature focus on simulated computable models (Nobre et al., 1991; Wunderling et al., 2022; Flores et al., 2024), but there is also correlational studies with Spracklen et al. (2012) being a pioneering work using observational data to establish the link of deforestation and rainfall. Leite-Filho et al. (2021) do consider local impacts of deforestation on climate and agriculture, but abstract from a notion of equilibrium using a simulated model calibrated from correlations. My contribution to this literature is twofold. First, I introduce econometric tools that allow for a more thorough discussion of identification challenges and statistical inference. Second, by modeling the equilibrium between climate and land use decisions, I enable a deeper understanding of the importance of adaptation and optimal behavior in response to a changing climate.

²Methodologically, this paper also connects with other papers using discrete choice models of land use (Scott, 2014; Sant'Anna, 2021), in particular Rafey (2020) uses a discrete choice model to study the impact of water markets on the Australian agricultural sector.

The remainder of this paper is organized as follows: Section 2 presents the model, connecting the climate and land use models in equilibrium. Section 3 discusses my counterfactual scenario of allowing deforestation in protected areas, showing empirical results on the impact of deforestation on climate and comparing equilibria across different land use policy scenarios. Section 4 addresses important caveats of the model. Section 5 concludes.

2 Model

This section has three parts. First, I build a model where climate is a function of land use. Second, I build a discrete choice model of land use that allows me to identify model parameters when climate is endogenous. Third, I show how these two models — climate and land use — can be combined into a single consistent framework.

Climate depends on land use

The seminal work by Salati et al. (1979) initiated a series of studies showing the role of tropical forests in regulating climate systems. One key mechanism by which tropical forests impact climate is through their role in the water cycle. Solar energy evaporates ocean water, creating humid air parcels that move toward land. As these parcels travel, they lose humidity through precipitation, while ground-level transpiration from trees replenishes air moisture, sustaining downwind rainfall. Through this process, forests influence rainfall in areas hundreds or even thousands of kilometers away — a phenomenon known as *flying rivers* (Marengo et al., 2004; Nobre, 2014). Consequently, deforestation in tropical forests reduces rainfall, impacting climate on a continental scale (Costa et al., 2012; Zemp et al., 2017; Staal et al., 2018).

The model of how land use affects climate is based on Spracklen et al. (2012). To model the causal effect of tropical forests on the climate of a location, three steps are needed: **First**, I build a back trajectory of atmospheric transport for that location. This trajectory traces the path of an air parcel from the ocean to the location of interest, identifying the terrain it passed over after reaching land. Figure 1 illustrates two examples of back trajectories that reach the target location (the white cell), shown as dashed black lines. Similar to tracing a river upstream to find its source, following wind patterns reveals the path taken by an air parcel from the ocean. By building such trajectories I can determine what type of terrain an air parcel has flown over since its arrival in continental land. **Second**, I calculate a weighted sum of forest pixels encountered along the back trajectory. For each step s along this trajectory – at each new pixel that the trajectory reaches – I discount the contribution of

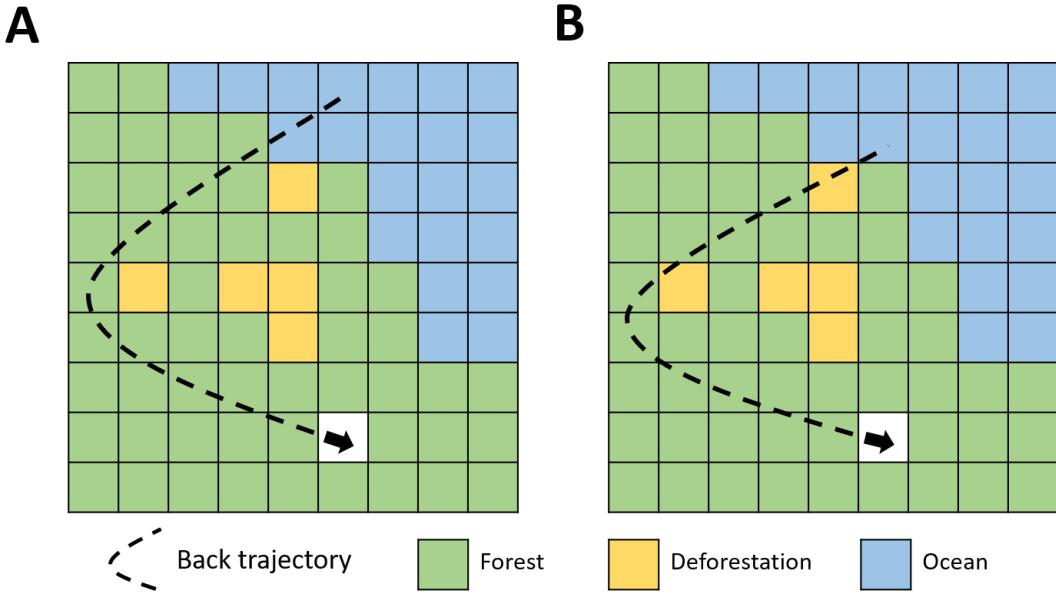


Figure 1: Elements of the Climate Model. This Figure illustrates the steps needed to build the climate model. The dashed black lines represent back trajectories of atmospheric transport, starting in the white cell and walking back until the Ocean is reached. Along each trajectory, different types of terrain are covered. In A, the back trajectory is exposed to 14 pixels of forest, while in B the trajectory is exposed to 10 pixels of forest. That way, it is likely that scenario A will deliver more rainfall for the white cell than scenario B.

the additional pixel by θ^s given a parameter $0 < \theta \leq 1$. This parameter θ accommodates the idea that more distant locations may have a reduced impact. When $\theta = 1$, for example, then there is no weighting and I simply count the number of forest pixels along the trajectory. In this case, the back trajectory shown in Figure 1A would yield a total exposure to the forest of 14 forest pixels, while that in Figure 1B would yield 10 forest pixels. Thus, it is likely that the scenario 1A will deliver more rainfall for the target location than the scenario 1B; **Third**, I connect rainfall and other climate outcomes at the target location to the weighted sum of forest pixels, termed *upwind forest*, using a linear model.

Formally, these three steps are described as follows.

First, given a location o in a time period q , I aim to determine the most likely trajectory of the air parcel over location o during the previous Δ periods of time. Let $W_{jq} = \{x_{jq}, y_{jq}, z_{jq}\}$ denote latitude, longitude and pressure level (altitude) of an air parcel j in time q and let V_{jq} denote the gradient of W_{jt} , which is the air velocity vector. Then this air parcel can be traced one step back as:

$$W_{j,q-1} = W_{j,q} - V_{j,q}$$

Variations in wind speed and direction across different spatial points determine how a particle will move within this wind surface. I then compound Expression 2 to build a back trajectory for an arbitrary time window. Given a time window of Δ , the compounded back trajectory is given by:

$$W_{j,q-\Delta} = W_{j,q} - \sum_{s=1}^{\Delta} V_{j,q-s} \quad (1)$$

Given a initial state $W_{j,0}$ and a time window Δ , I can compute a back trajectory. A location o already defines a latitude and a longitude (x_{jt}, y_{jt}) . By setting a pressure level (z_{jt}) and a time window (Δ), I can compute the back trajectory of an air parcel that reaches location o . For now, take the initial pressure level and the time window as given.

Second, once the back trajectory is built, I need to connect it with the underlying land cover. Let B_o denote the back trajectory from a location o , which is the set of all locations that the air parcel, which precipitates over o , has passed through during the previous Δ time periods. Let h_o denote the measure of forest cover at location o and let $s(o, j)$ denote the number of steps separating location o from location j . Then the upwind forest of a back trajectory is defined as:

$$H_o = \sum_{j \in B_o} \theta^{s(o,j)} h_j \quad (2)$$

Referring again to Figure 1, along each back trajectory, the upwind forest (H_o) is the weighted sum of forest pixels.

Third, in the last step I model climate outcomes, primarily rainfall, over location o as a function of the upwind forest. The model for rainfall is parametrized by:

$$R_o = \gamma_R H_o + \epsilon_o \quad (3)$$

Where R_o denotes rainfall at location o . The marginal effect of deforestation on rainfall is given by γ_R , if the deforestation occurs along the back trajectory path; otherwise the effect is zero. For now, I only assume an additive functional form for the impact of land use on rainfall, as shown above, without making further assumptions about the relationship between H_o and ϵ_o . The impact of deforestation on rainfall may have additional consequences for the climate. Temperature, in particular, is affected by precipitation. Thus, analogous to Expression 3, I also model temperature as a function of upwind forest:

$$T_o = \gamma_T H_o + \varepsilon_o \quad (4)$$

Where T_o denotes temperature at location o . Estimating the empirical effects of deforestation on climate involves several challenges, but I defer a discussion on data and identification until after presenting the complete theoretical model of land use and climate. This section showed how land use affects climate. The next section builds a model of how climate affects land use.

2.1 Land use depends on climate

A farmer owns a plot of land i at location o and chooses a land use from a discrete choice set. The farmer can either keep the plot as forest or clear it for agricultural activities. Formally, let F denote the choice of keeping the land as forest while $c \in C$ denotes using the land for agriculture, where c denotes a choice within the possibility set C of available agricultural activities.

If the farmer chooses an agricultural activity c , the farmer receives as return:

$$\tilde{\pi}_i^c = Climate_o \tilde{\alpha}_c + X_o \beta_c + \tilde{\xi}_o^c + \tilde{\varepsilon}_i^c$$

Where, $Climate_o$ denotes a vector of climate outcomes such as rainfall and temperature at the location level, and $\tilde{\alpha}_c$ is a vector of parameters that determines how climate affects land use returns; X_o denotes a set of variables, such as, transportation cost and soil characteristics; β_c is another vector of parameters; $\tilde{\xi}_o^c$ is an aggregate shock at the location level; $\tilde{\varepsilon}_i^c$ is an idiosyncratic shock at the farm level. Given the climate model discussed in the previous section, this expression introduces some identification problems since the climate model uses the same variation - changes in upwind forest - for different climate outcomes. If, for example, $Climate_o$ includes rainfall and temperature, $Climate_o \alpha_c$ will end up being a linear combination of upwind forest. This will be true for any additional climate outcomes, such as days without rain and growing degree days as in Schlenker and Roberts (2009). Therefore, my approach is to anticipate this constraint and incorporate it by directly rewriting the farmer's return as:

$$\tilde{\pi}_i^c = \alpha_c H_o + X_o \beta_c + \xi_o^c + \varepsilon_i^c \quad (5)$$

As is in the previous section, H_o denotes upwind forest at location o . In Expression 5, the climate model is incorporated directly into the land use model, such that the shocks ξ_m^c and ε_i^c also include climate shocks net of the upwind forest mechanism.

The return a farmer receives by keeping his plot of land as forest is described by a constant:

$$\tilde{\pi}_i^F = k_F + \xi_o^F + \varepsilon_i^F \quad (6)$$

Where k_F is the constant, ξ_o^F is an aggregate shock at the location level, and ε_i^F is an idiosyncratic shock. I model the farmer's discrete choice problem as a logit. In Assumption 1 I formalize it.

Assumption 1 *The vector of land unobservables $\{\varepsilon_i^F, \{\varepsilon_i^c\}_{c \in C}\}$ has the joint cumulative distribution given by*

$$\exp \left[-e^{-\varepsilon_i^F} - \sum_{c \in C} e^{-\varepsilon_i^c} \right]$$

It will be useful to denote by π_o^c and π_o^F Expressions 5 and 6 free of the idiosyncratic shock. Then, with Assumption 1, the probability of choosing land use $j \in \{C, F\}$ at location o is given by:

$$P_o(j) = \frac{e^{\pi_o^j(H_o,.)}}{\sum_{n \in \{C, F\}} e^{\pi_o^n(H_o,.)}} \quad (7)$$

Lastly, it will be useful to rewrite the returns net of the forest return, so that the choice probability for forest can be written as:

$$P_o(F) = \frac{1}{1 + \sum_{c \in C} e^{\pi_o^c(H_o,.) - k_f - \xi_o^F}} \quad (8)$$

Given my specifications, Expression 8, which is the probability that a farmer will not use their land for agricultural activities, depends on land use choices of other farmers through the H_o term. That is, a farmer's decision depends on other farmers' decisions. I now turn to a characterization of the equilibrium in this economy with endogenous climate.

2.2 Land use and climate in equilibrium

Land use depends on climate, and climate depends on land use. Integrating the land use model and the climate model amounts to solving a fixed point problem, that is, finding a land use that is consistent with the climate generated by it.

To translate the land use decision of the discrete choice model to the deforestation scenario of the climate model, let $P(F)$ denote the vector of probabilities for all plots of land where each entry is given by $P_o(F)$, that is, the probability of location o having forest (Expression 8). I interpret h_o as $P_o(F)$, that is the measure of forest cover in the climate model is the

probability of forest cover in the land use model. This allows me to map forest probabilities into forest probabilities, that is, $P(F)$ depends on $P(F)$. Formally, I can define a map $M : P(F) \rightarrow P(F)$ as:

$$\begin{aligned} P(F) &= [P_1(F), \dots, P_N(F)] \\ &= \left[\frac{1}{1 + \sum_c e^{\pi_1^c - k_f - \xi_1^F}}, \dots, \frac{1}{1 + \sum_c e^{\pi_N^c - k_f - \xi_N^F}} \right] \\ &= MP(F) \end{aligned}$$

A fixed point of the map M can be interpreted as a Nash equilibrium of this economy, where actions are deforest and not-deforest and mixed strategies are allowed. In equilibrium a farmer chooses their best possible land use given the land use choices of all the other farmers. The existence of such an equilibrium is a straightforward application of Kakutani's fixed point theorem. That is because the correspondence M maps the set $[0, 1]^N$ to itself and the map is smooth. This is the advantage of using a discrete choice model in this application, the probabilities guarantee the compactness of the domain and the image and the functional form for the choice probabilities guarantees the necessary smoothness of the map. I formalize this in the next lemma.

Lemma 1 *Existence of Equilibrium*

Let $P = [0, 1]^N$ be the closed N -dimensional unit cube. Suppose $M : P \rightarrow P$ is a correspondence defined by $M(P) = [P_1(P), P_2(P), \dots, P_N(P)]$, where each $P_o : P \rightarrow [0, 1]$ is a continuously differentiable function mapping P into the closed interval $[0, 1]$. Then M has a fixed point.

PROOF. By Kakutani's fixed point theorem, a correspondence $M : P \rightarrow P$ has a fixed point if:

1. P is a non-empty, compact, and convex subset of a Euclidean space,
2. M is upper semi-continuous,
3. $M(P)$ is non-empty, compact, and convex for each $P \in [0, 1]^N$.

In this case, $P = [0, 1]^N$ is a compact and convex subset of \mathbb{R}^N , satisfying the first condition. Each function P_o is continuously differentiable, ensuring continuity and hence the upper semi-continuity of the correspondence M . Finally, since $M(P)$ consists of compact

convex sets within $[0, 1]^N$ (by continuity of the P_o functions), the images of M are non-empty, compact, and convex. Thus, by Kakutani's fixed point theorem, there exists $P^* \in P$ such that $P^* \in M(P^*)$. ■

The first lemma proves the existence of equilibrium, but in principle there could be many equilibria. Uniqueness of the equilibrium is more complicated and it will depend on the effect that each location has on other locations, that is, it will depend on the model's parameters. I formalize this condition in the next lemma.

Lemma 2 Sufficient condition for the uniqueness of the equilibrium

Let $\bar{\alpha} = \max_c\{\alpha_c\}$. If $\frac{1}{1-\theta}\frac{\bar{\alpha}}{4} < 1$, then the correspondence $M : P(F) \rightarrow P(F)$ is a contraction map, has a unique fixed point (equilibrium), and this fixed point can be found by iterating M starting from any initial point.

PROOF. I start with the partial derivative of $P_o(F)$ with respect to $P_j(F)$. This derivative is either zero, if j is not on the back trajectory of i ($j \notin B_i$) or is given by:

$$\frac{\partial P_o(F)}{\partial P_j(F)} = -\theta^{s(o,j)} P_o(F) \frac{\sum_c \alpha_c e^{\pi_o^c}}{1 + \sum_c e^{\pi_o^c}}$$

Let $\bar{\alpha} = \max_c\{\alpha_c\}$, then I can bound the absolute value of the derivative as:

$$\left| \frac{\partial P_o(F)}{\partial P_j(F)} \right| \leq \left| -\theta^{s(o,j)} P_i(F) \frac{\sum_c \alpha_c e^{\pi_o^c}}{1 + \sum_c e^{\pi_o^c}} \right| \leq \bar{\alpha} |\theta^{s(o,j)} P_o(F)(1 - P_o(F))| \leq \frac{\theta^{s(o,j)} \bar{\alpha}}{4}$$

By the mean value theorem,

$$|P_o(F|P_j(F)) - P_o(F|P'_j(F))| \leq \frac{\theta^{s(o,j)} \bar{\alpha}}{4} |P_j(F) - P'_j(F)|$$

A change in $P_j(F)$ will have an effect on $P_o(F)$ but also on all other locations that are on the path of the trajectory that crosses location j . For locations that are not on this trajectory the effect is zero and the inequality above will always be true. As each location is, by construction, one step farther than the other and $\sum_{s=1}^{\infty} \theta^s = 1/(1-\theta)$ I can write:

$$\sum_o |P_o(F|P_j(F)) - P_o(F|P'_j(F))| \leq \frac{1}{1-\theta} \frac{\bar{\alpha}}{4} |P_j(F) - P'_j(F)|$$

This is the bound on the effect of location j on all locations of the economy. With this inequality I can bound the total change in map M using the triangular inequality as:

$$\begin{aligned}
|MP - MP'| &\leq |M[P_1, P_2, \dots, P_N] - M[P'_1, P_2, \dots, P_N]| + \\
&\quad |M[P'_1, P_2, \dots, P_N] - M[P'_1, P'_2, \dots, P_N]| + \\
&\quad \dots \\
&|M[P'_1, P'_2, \dots, P_N] - M[P'_1, P'_2, \dots, P'_N]| \leq \\
&\frac{1}{1-\theta} \frac{\bar{\alpha}}{4} \sum_j |P_j(F) - P'_j(F)|
\end{aligned}$$

Thus, if $\frac{1}{1-\theta} \frac{\bar{\alpha}}{4} < 1$, I have

$$|MP - MP'| < \Theta |P - P'|, \quad \text{with } \Theta < 1$$

That is, if $\frac{1}{1-\theta} \frac{\bar{\alpha}}{4} < 1$, the map M is a contraction, there exists an unique fixed point (equilibrium) and I can find it by iterating the map M starting from any initial point. ■

Lemma 2 provides a sufficient condition for the uniqueness of equilibrium. This condition is intuitive: the externality effect of an individual's decision cannot be too strong, a condition that can be tested after estimating the model's parameters. Furthermore, this condition relates to more general results concerning Nash equilibrium in networks and the necessity of weakness of externality (Bayer et al., 2023).

My model is static, so that the economy jumps between equilibria. But, mathematically, the contraction map is also a result showing the convergence to an equilibrium through a learning process among agents. The contraction map shows that if each agent updates its decision based on what other agents are doing, an unique equilibrium will eventually be reached.³ This finding is noteworthy because, even tough there exists some general results on the convergence of best-responses to a Nash equilibrium in networks with reciprocal links (Bramoullé et al., 2014), the same is not true when interactions are not reciprocal (Bayer et al., 2023). In my model, interactions are not reciprocal: the network of externalities is determined by atmospheric circulation, which resembles a directed acyclic graph. As demonstrated in Bayer et al. (2023), such settings can create best-response cycles, rendering the Nash equilibrium a questionable equilibrium concept. However, this issue does not arise in my model.

³It is easy to think of settings where there is no best-response convergence to the equilibrium. The game of rock-paper-scissors is a simple example.

3 Application

For an application of the model, I focus on the importance of protecting Indigenous Territories in the Amazon for the agricultural sector of the Brazilian state of Mato Grosso. Figure 2A shows the location of the rainforest and of the state of Mato Grosso.

Mato Grosso is one of the most important agricultural hubs in the world, being responsible for 10% of the global production of soybeans. This achievement was made possible by the introduction of new agricultural techniques that expanded the application of double cropping (Bustos et al., 2016; Xu et al., 2021). Double cropping refers to a farming practice where two harvest seasons occur in the same year. In Mato Grosso, this typically involves planting soybeans followed by a mid-season corn crop. Another significant agricultural activity in the state is cattle ranching, which corresponds to almost 40% of land use.

I work with data from 2008 and 2017, a stable period in regulatory framework of land use in Brazil (see Burgess et al. (2019)). To estimate the model, I exclude observations inside protected areas, e.g., conservation units and Indigenous land. These protected areas are under a different regulatory framework than other regions and have significantly less deforestation. Figure 2B shows the locations of protected areas. With this selection of locations, the estimation of the land use model yields parameters that govern the decisions of agents on private land. Later, in my counterfactual exercise, I allow agents to act on Indigenous Territories as if they were in private land, effectively rolling back the protection status of these territories.

3.1 Data.

I work primarily with two types of data: high-resolution land use data and lower-resolution climate data. I interpret a pixel of land use data (with approximately 250 meters resolution) as a farm indexed by i in the model and a pixel of climate data (with approximately 25 kilometers of resolution) as a location indexed by o in the model. Below I describe each data set separately.

Land use

For this application, the set of available agricultural activities (C) includes growing cash crops and using land for pasture grazing. I define $C = \{\text{double}, \text{single}, \text{pasture}\}$ for the choices of double cropping (soybeans and corn), single cropping (soybeans), and pasture grazing.

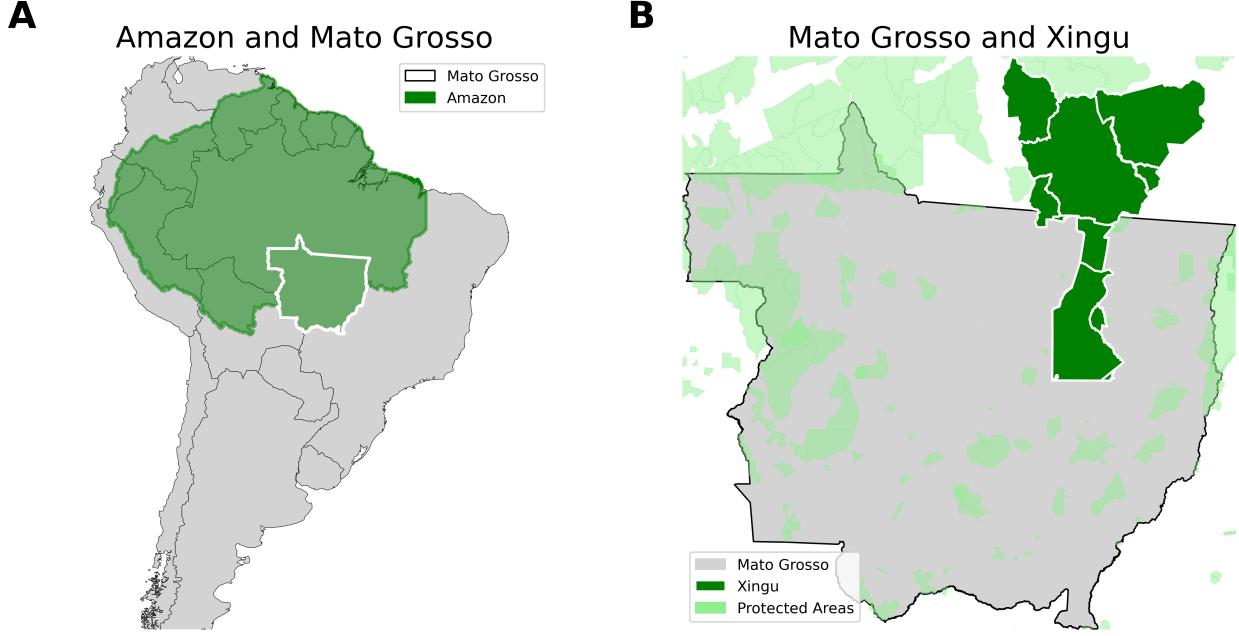


Figure 2: These maps show: (A) the location of the Amazon rainforest and of the Brazilian state of Mato Grosso; (B) the location of the Xingu Indigenous Territories, in dark green, and remaining protected areas, in light green, within and around Mato Grosso.

Land use data is from Simoes et al. (2020). The data classifies land use from 2001 to 2017 for pixels of 250-meter resolution for the Brazilian state of Mato Grosso, but I focus only on the years 2008 and 2017 to take a long difference during this stable period of regulatory framework. Importantly, the construction of this dataset uses information throughout the year to build the land use classification algorithm, allowing it to identify double cropping. That is, it identifies whether the same pixel had two harvests in the same year. For Mato Grosso, double cropping consists of soybeans followed by corn. Furthermore, the data identifies single cropping (soybeans), forest, and pasture areas. Figure A.1 in the Appendix shows the land use data.⁴ For every location o , I compute the share of land use pixels classified as single, double, pasture, or forest for the years of 2008 and 2017.

One empirical problem is that I end up with many pixels where the share of one type of land use is zero, which cannot be used in my estimation of the land use model due to the logarithmic transformation. There are various ways to address this issue, such as smoothing

⁴I collapse some of the land uses present in the data. Soybeans and millet is collapsed into single cropping, since millet is a intercrop option generally not sold but instead used to avoid leaving the soil bare. I also collapse cotton and sugarcane as single cropping since they represent a very small proportion of crops. Other double cropping systems in the data, such as soybeans-cotton and soybeans-sunflower, are also a very small proportion (< 1%) of the data. I just collapse them to double cropping together with soybeans-corn. I do not distinguish between wetland, rainforest, cerrado and secondary vegetation in the choice set, defining all these categories as the choice F . Finally I drop observations of urban areas and water.

the probabilities in space (Scott, 2014) or using a partial identification approach (Lima, 2024). I opt for the most straightforward and transparent solution: when estimating the land use model, I restrict my sample to pixels with positive shares for all choices and years. Figure A.2 in the Appendix shows the three types of pixels I use in this paper: pixels in private land with zero shares, pixels in private land with only positive shares (used in my estimation of the land use model), and pixels in protected areas.

Climate.

Climate variables are: rainfall, temperature, and the upwind forest.

Rainfall and temperature. Average annual rainfall and temperature data by location is from Copernicus (2017). I collect climate data from 1985 to 2018 at the resolution of 0.25° (approximately 25 kilometers).

Back Trajectories of Atmospheric Transport. To build the back trajectory of atmospheric transport, I use monthly wind data from 1985 to 2018 (Copernicus, 2017), also at the resolution of 0.25° . This dataset provides a velocity vector (u-component, v-component, and vertical velocity) across time, latitude, longitude and pressures levels (altitude). I depart from Spracklen et al. (2012) by using the prevailing winds of each month-year as my data for the velocity vector, instead of hourly data. This approach allows me to use a longer panel when estimating the impact of upwind forest on climate outcomes. The resolution of the data is kept at the native resolution of 0.25° instead of using the resolution of 1° as in Spracklen et al. (2012). Given a month, a year, and a starting point of interest, the back trajectory is computed using Expression 1. I set the time window of the back trajectory to 5 days and the initial pressure level to 800hPa as in Spracklen et al. (2012).

The back trajectories are computed for endpoints in the state of Mato Grosso, as this is the region where I want to model the climate effect. The back trajectories themselves can span across all South America. Figure 3 shows a sample of different back trajectories.

Upwind forest. Spracklen et al. (2012) uses a Leaf Area Index (LAI) as a forest cover measure, my h_o variable. To better align the data with my model, I use binary forest data of forest cover, which has the advantage of being consistent with the outputs of the land use model. I set my forest cover measure h_o for location o equal to the proportion of pixels in o classified as forest. Forest pixel data are from Mapbiomas⁵, which classifies land use for the entire Amazon from 1985 to 2018 at a resolution of 30 meters. I create a map where pixels classified as forest are assigned a value of 1, while non-forest pixels are assigned a value of

⁵ Available at www.amazonia.mapbiomas.org

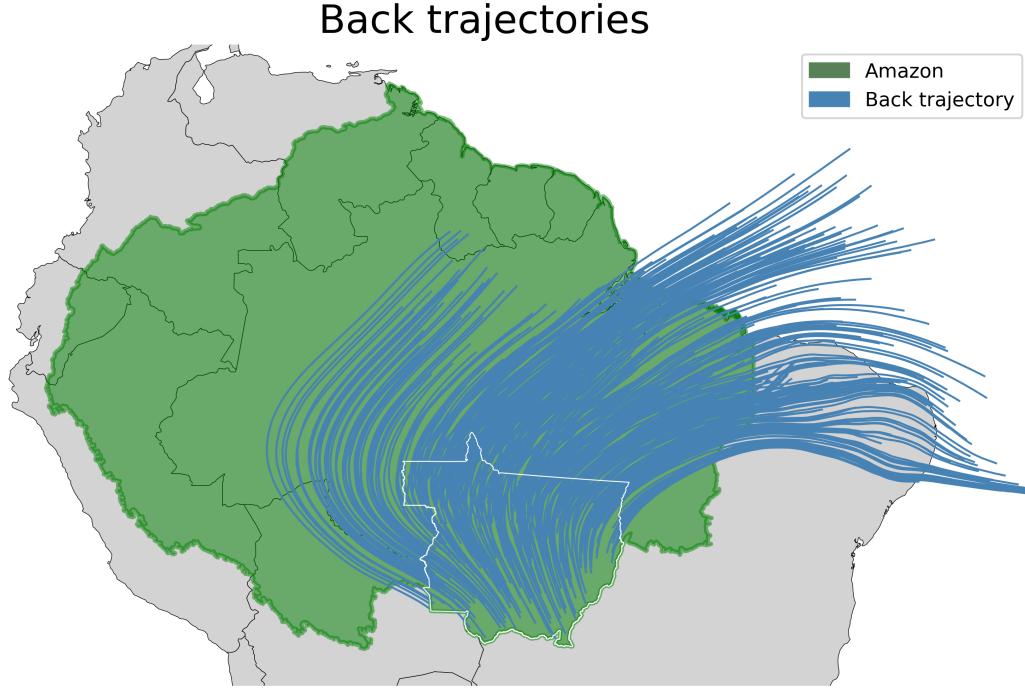


Figure 3: This figure shows a sample of 400 5-days trajectories that arrive at some locations in Mato Grosso.

0, and then apply an average filter to match the resolution of the Copernicus (2017) data to compute h_o .

Although I use binary forest data (forest or deforested), the high resolution provides significantly more granularity compared to LAI data available at 5 km resolution (Lemieux and Luquire, 2021)⁶. Note, even though the forest data is binary, the required aggregation process for the climate model generates a continuous variable of forest cover status. I set θ , the parameter governing the weight of upwind forest by distance, to $\theta = 0.99$. This choice closely mirrors Spracklen et al. (2012) – which does not weight distance – while providing a straightforward approach to ensure that deforestation in one location does not produce an indefinite significant cascading effect.⁷

For every location o , I compute the average annual rainfall and temperature, and upwind

⁶For comparison: a single pixel of 5 kilometers contains around 28,000 pixels of 30 meters.

⁷One way to validate this calibration is by examining the fraction of precipitation originating in the Amazon for different regions in South America. Regions located in the southern part of the continent—approximately 2,500 to 3,500 km from the center of the forest—derive between 5% and 15% of their rainfall from the Amazon (see Costa et al. (2021) and Zemp et al. (2014) for a review). In contrast, regions within the Amazon receive around 50% of their rainfall from local forest sources. Such observations would not be possible with $\theta < 0.98$, as the distance decay would be too strong. It is also worth noting that larger values of θ make it more challenging to satisfy the sufficient condition in Lemma 2 for contraction. By selecting $\theta = 0.99$, I aim to strike a balance: this choice is conservative in terms of the theoretical implications while remaining consistent with empirical observations.

forest.

Additional variables.

Transportation cost data is from Araujo et al. (2020). This dataset combines: (1) a transportation network encompassing paved and unpaved federal and state roads, waterways, and ports; (2) freight cost data. This data is used to build an application that computes the transportation cost from each location in my data and the most cost-effective port with access to international markets. **Soil data** are from the Brazilian Agricultural Research Corporation. This comprises the best available data on soil for the entire country, with high-resolution classification of 40 soil types, developed through field observations and expert classifications.⁸. Finally, terrain slope data is from Jarvis et al. (2008). For each location o , I compute the average transportation cost, slope, and share of each soil type.⁹

3.2 Estimation, Identification and Results.

Rainfall and Temperature Effects.

I begin by establishing that upwind forest indeed causes more rainfall and milder temperatures. This is an important step to show that the variation of upwind forest indeed operates through the effects on climate, since in the land use model upwind forest enters in a reduced form way.

As I have yearly panel data on rainfall, temperature, and upwind forest – from 1985 to 2018 – the empirical analog of Expressions 3 and 4 are given by:

$$R_{oy} = \gamma_o + \gamma_y + \gamma_R H_o + X_{oy}\beta + \epsilon_{oy} \quad (9)$$

$$T_{oy} = \gamma'_o + \gamma'_y + \gamma_T H_o + X_{oy}\beta' + \varepsilon_{oy} \quad (10)$$

Where y denotes year, γ_o and γ_y are location and year fixed effects, and X_{oy} denotes controls. Specifically, I control for total distance of the back trajectory within continental land, accounting for the fact that longer trajectories will mechanically have higher upwind

⁸At the most aggregated level the data has soils classified in Alisol, Argisol, Acrisol, Cambisol, Chernozem, Spodosol, Gleysol, Latosol, Luvisol, Neosol, Nitisol, Organosol, Planosol, Plinthosol, Vertisol

⁹Why don't I use the FAO GAEZ data on crop suitability? The FAO data combines soil and climate information, meaning that climate variables are already embedded in their crop suitability measure. As a result, attempting to estimate the impact of upwind forest while holding the FAO suitability measure constant would be conceptually inconsistent. I also note that the soil data I use provides greater detail than the FAO GAEZ soil data.

forest. In this specification the variation identifying the parameters is a combination of year-to-year changes in deforestation and atmospheric trajectories. However, unobserved factors could still affect both deforestation and climate outcomes. Although the variation in H_o relies on long trajectories (primarily reflecting deforestation far from location o), there remains a possibility of reverse causality. To address this, I create an instrument for the upwind forest. By fixing the forest cover scenario in 1985 (the first year of data) and calculating upwind forest each year based only on changes in atmospheric trajectories, I ensure that the instrument is exogenous to deforestation.

It is important to understand the source of the correlation between upwind forest and potential upwind forest. The sign of this correlation is determined by two opposing forces: (1) cross-sectional variation which increases the correlation, as more potential upwind forest implies greater general exposure to the forest (2) temporal variation which decreases the correlation, as more potential upwind forest implies that deforestation throughout the years will have a higher impact on changes of upwind forest. Consequently, in a specification with fixed effects, the correlation between upwind forest and potential upwind forest can be either positive or negative, depending on the time window used for estimation, which determines the dominant force. Regardless, the variation in the instrument itself is driven solely by changes in atmospheric circulation. Additionally, established mechanisms linking transpiration and humidity transport (Salati et al., 1979; Spracklen et al., 2012) support a causal interpretation of my estimates. Additionally, established mechanisms linking transpiration and humidity transport (Salati et al., 1979; Spracklen et al., 2012) support a causal interpretation of my estimates.

Table 1 shows the results for rainfall and temperature. Upwind forest has a positive effect on rainfall, as shown in columns (1) and (2). The difference between the OLS and the IV estimates suggests a potential downward bias in the OLS estimates, with the IV estimate being more than 3 times greater than the OLS estimate. Quantitatively, a one standard deviation increase in upwind forest increases average rainfall by 10% [$12.72 \times 0.0412 / 4.77$]. Additionally, a one standard deviation in upwind forest accounts for 60% of the standard deviation of rainfall [$12.72 \times 0.0412 / 0.88$], a figure in line with typical estimates of the proportion of rainfall attributed to forest cover (Zemp et al., 2017). Table 1 also shows that the average temperature in Mato Grosso has a low standard deviation. This stability is one reason for the region's high agricultural productivity. However, I still estimate a precise effect of upwind forest on temperature. Quantitatively, a one standard deviation increase in upwind forest reduces temperatures by $0.12^\circ C$.

Having established that upwind forest indeed impacts rainfall and temperature in a significant way, I now turn to the estimation of land use decisions.

Table 1: Upwind Forest on Rainfall and Temperature

	OLS Precipitation (mm) (1)	IV Temperature (°C) (2)	OLS Temperature (°C) (3)	IV Temperature (°C) (4)
Upwind forest	0.0119*** (0.0012)	0.0412*** (0.0014)	-0.0061*** (0.0006)	-0.0096*** (0.0019)
First-stage				
Upwind constant forest		0.8931*** (0.0025)		0.8931*** (0.0025)
F stat		48,221		48,221
Year FE	✓	✓	✓	✓
Pixel FE	✓	✓	✓	✓
Land	✓	✓	✓	✓
Observations	39,732	39,732	39,732	39,732
Cluster	1,204	1,204	1,204	1,204
Mean Dep.	4.77	4.77	25.36	25.36
Std Dep.	0.88	0.88	0.88	0.88
Mean Ind.	28.97	28.97	28.97	28.97
Std Ind.	12.72	12.72	12.72	12.72

This table shows the estimates of Expressions 9 and 10. All specifications include location and year fixed effects, and control for upwind exposure to continental land (length of the trajectory within continental land). Columns (1) and (3) show OLS estimates of upwind forest on rainfall and temperature. Columns (2) and (4) show IV estimates of upwind forest on rainfall and temperature, where the instrument is the upwind forest fixing forest cover in 1985. The 'Upwind constant forest' row shows the results of the first-stage, that is a regression of the upwind forest on upwind forest with forest as 1985. Standard errors are clustered at the location level. The bottom four rows shows descriptive statistics of the dependent (rainfall and temperature) and independent (upwind forest) variables. *** p-value < 0.01

Land Use Decisions.

To estimate the parameters that govern land use decisions, take the logarithm of Expression 7 and use the forest choice as the outside option to obtain the usual regression from the discrete choice model:

$$\log \frac{P_o(c)}{P_o(F)} = \alpha_c H_o + X_o \beta_c - k_f + \xi_o^c - \xi_o^c \quad (11)$$

As my model is static but I have panel data, some adaptations are necessary. I use data on land use for the years of 2008 and 2017. The idea is to take long differences since I am not investigating the impact of weather on land use decisions but rather of climate. One adaptation is to allow H_o to be measured as the cumulative upwind forest within a given interval. Specifically, I measure H_o in 2017 as the average upwind forest between 2009 and 2017 while I measure H_o in 2008 as the average upwind forest before 2008. The rationale is that it is the upwind forest of the entire past that determines the land use decision I observe today, such that changes in land use decisions are determined by changes in upwind forest during the time frame in which I measure land use changes.

The main parameters of interest are the α_c parameters - one for each agricultural choice - which capture the effect of upwind forest on the return of each land use choice. It is difficult to argue that this specification can be estimated by OLS, as it may also suffer from reverse causality and the upwind forest could be correlated with various unobservables. For example, locations in the core of the forest are likely to have a higher upwind forest but are also more isolated, with less infrastructure. One possible approach is to model these factors, as in Araujo et al. (2020) and Domínguez-Iino (2021). I take a different approach and estimate my model in two stages. First, I take long differences to deal with fixed factors and build an instrumental variable to estimate α_c . Second, once α_c is estimated, I select and estimate the remaining factors using a predictive approach. I further discuss the rationale and implications of this approach as I present the results. Taking first differences I arrive at:

$$\Delta \log \frac{P_o(c)}{P_o(F)} = \alpha_c \Delta H_o + \underbrace{\tilde{\xi}_o}_{\Delta X_o \beta_c + \Delta \xi_o^c} \quad (12)$$

The first difference already eliminates all factors that are constant over time. This includes, for example, soil characteristics, and within our ten-year window (2008 to 2017), this also includes transportation cost. But of course there could be other factors contained in $\tilde{\xi}_o$ that may correlate with the upwind forest measure and affect land use decisions. To resolve this identification problem, I need an instrument. The instrument I build is the same as the one used in the climate model, that is, it is the potential upwind forest derived from the

annual atmospheric trajectory variation while holding deforestation constant as of 1985, the first year for which I have data. This does not imply that I am using similar variation to estimate two different mechanisms; rather, the climate model was estimated to show that the mechanism I propose - that of upwind forest affecting climate - is indeed the mechanism through which upwind forest can affect land use decisions.

Table 2 shows the results of this estimation. With four possible land use choices - pasture, single cropping, double cropping, and forest - I estimate 3 coefficients ($\alpha_{pasture}$, α_{double} , α_{single}). The coefficient for the pasture choice is largely insignificant, that is, changes in climate do not affect the return of pasture. This is expected given the resilience of cattle and the fact that pasture grazing in Brazil is a low-productivity activity, averaging around one cattle head per hectare, which is well below the support capacity of pastureland in Brazil (Strassburg et al., 2014). The other two parameters are positive and statistically significant, showing that the return of single and double cropping increase with higher upwind forest. The IV coefficients are significantly greater than the OLS coefficients, showcasing that unobservable factors - such as the core versus edge of the forest dynamics - may significantly bias the estimates.

In the estimation, I include a constant and upwind exposure to continental land (length of the trajectory within continental land). I do not control for climate variables since, as established by the climate model results, they are bad controls (Cinelli et al., 2024). The remaining variables identified as important in the literature, such as soil and transportation cost, are eliminated by first differencing the regression. What could invalidate my results is the existence of some unobservable factor that simultaneously affect land use decisions and atmospheric circulation. However, atmospheric circulation is determined by temperature differences between the Equator and the Poles and the rotation of the Earth (the Coriolis effect), that is, it is exogenous that what happened in my study area between 2008 and 2017.¹⁰

Establishing the causal effect of upwind forest on land use decisions is not enough to understand counterfactuals. This is because of two additional challenges: first, each α_c gives a *ceteris paribus* effect, and we know from the model that farmers will re-optimize, arriving at a new land use equilibrium after a change in upwind forest; second, the level of the returns (π_o^c) matters for the counterfactual given the functional form of the choice probabilities. Thus, I need to recover the remaining factors that determines land use returns. With α_c

¹⁰There is a study on how the rainforest can affect atmospheric transport: The Biotic Pump Theory (Makarieva and Gorshkov, 2007) approaches this question from a theoretical perspective. As far as I know, this theory is not a consensus because it lacks empirical evidence. Nonetheless, the Biotic Pump Theory would apply in a context of extensive deforestation, and therefore, even if it were true, it is unlikely to have an effect in my setting.

Table 2: Upwind Forest on Land Use

	OLS log ratio of probabilities (1)	IV (2)
Upwind forest		
pasture	-0.0015 (0.0009)	-0.0044* (0.0026)
double	0.0058* (0.0036)	0.0496*** (0.0077)
single	0.027*** (0.0039)	0.0708*** (0.0084)
Observations	1,428	1,428
Clusters	1,428	1,428
		First stage
Δ Upwind forest constant		-1.488*** (0.093)
F stat		71
Observations		476
Clusters		476
Length	✓	✓

This table shows the estimates of Expression 12. All specifications include a constant and control for upwind exposure to continental land (length of the trajectory within continental land). Column (1) shows OLS estimates of upwind forest on the log of probabilities ratio. Columns (2) shows IV estimates of upwind forest on the log of probabilities ratio, where the instrument is the upwind forest keeping forest cover constant at 1985. Standard errors are clustered at the location level, which in this case of first-differences, is the heteroskedastic robust standard error. p-value *** < 0.01, * < 0.05, * < 0.1.

estimated, denoted by $\hat{\alpha}_c$, I rewrite my regression as:

$$\log \frac{P_o(c)}{P_o(F)} - \hat{\alpha}_c H_o = X_o \beta_c - k_f + \xi_o^c - \xi_o^c \quad (13)$$

One approach would be to explicitly model the right-hand side as in Araujo et al. (2020) and Domínguez-Jino (2021). This is usually done using some measure of crop suitability - most commonly the FAO GAEZ data - which would keep fixed climate conditions and thus is not suitable for my setting. For my counterfactual, the parameter of interest is α_c , while the other parameters β_c are used only to determine the initial level of the choice probabilities, that is, I am not interested in any particular parameter within β_c . Therefore, my approach is a predictive one: within a set of possible variables, I select those with the best out-of-sample fit, that is, the model that best predicts $\log \frac{P_o(c)}{P_o(F)} - \hat{\alpha}_c$ out of sample. This approach limits the counterfactuals I can perform to those related to H_o , but that is precisely the main point of this paper.¹¹

In this predictive step, I use data for 2017, as it is the last year of the data and represents the equilibrium I will perturb in my counterfactual. I split my sample by location into 5 folds to assess the predictive power of the model through a 5-fold cross-validation exercise. The variables that performed best in this exercise were latitude and longitude, transportation cost, terrain slope, and the share of nine different types of soil. The average out-of-sample R^2 across the five folds is 0.97, indicating very high predictive power and highlighting that this simple model effectively distinguishes different probabilities of land use choices. Table A.1 in the Appendix shows the coefficients of the predictive model.

I can now proceed to my counterfactual exercise.

3.3 Counterfactual: Deforesting the Xingu

The Indigenous Territories of the Xingu basin (Figure 2), located within the Amazon Basin in Brazil, represent one of the earliest initiatives in indigenous land protection.¹² The Indigenous Park of Xingu, a portion of the basin, was established in the 1960s as Brazil's first indigenous territory. The creation of the Xingu Park marked an important moment in Brazil's history, as it set a precedent for the protection of indigenous lands and cultural heritage in the Amazon, leading to the establishment of other indigenous territories across

¹¹An alternative approach would be factor models, allowing for interactive fixed effects (Moon and Weidner, 2015). Unfortunately, this approach does not fit well in my application because I need to be able to extrapolate the parameters from a set of pixels (private land) to another set of pixels (those on the Xingu), and I cannot extrapolate fixed effects.

¹²This region encompasses the territories of Kayapó, Baú, Menkagnoti, Badjônkôre, Panará, Terena Gleba Iriri, Kapót, Nhore, Capoto/Jarina, Wawi, Xingu, Batovi, and Pequizal do Naruvôtu.

the region. The Xingu Indigenous Territories spans around 141 thousand square kilometers, home to more than 30 indigenous groups, including isolated tribes.

Despite its historical significance and legal protections, the Xingu Indigenous Territories face mounting threats that endanger both the environment and the indigenous communities residing there. Degradation driven by illegal logging and gold mining, and edge deforestation driven by agriculture and cattle ranching, continues to encroach upon the territory, heavily impacting indigenous welfare and the environment. Additionally, projects such as hydroelectric dams and highway construction pose risks to the ecological balance and the traditional livelihoods of Xingu's indigenous peoples.

In my application, I generate a scenario where the Indigenous Territories of the Xingu River Basin would be opened to deforestation and the expansion of the agricultural frontier. Specifically, I simulate a scenario where the protection status of the territory is rolled back. To do this, I apply the estimated coefficients from data on private land to the Xingu region to generate a counterfactual deforestation scenario in the territory. Leveraging the contraction result of Lemma 2, I use the following algorithm to derive a new equilibrium for this economy:

1. Set initial choice probabilities – in my specific application I set probabilities as those fitted by the model, with pixels inside Xingu having forest probability of one
2. Perform a counterfactual exercise that will change choice probabilities – in my specific application, I change the choice probabilities of pixels inside the Xingu from 1 to the fitted probabilities from the model
3. Compute the change in upwind forest caused by the change in choices probabilities
4. With the new upwind forest compute new choice probabilities
5. If old and new choice probabilities are equal, stop; otherwise return to step 3.

After rolling back the protection status of the Xingu, 47% of its territory is converted to agricultural land, equivalent to 66 thousand squared kilometers of deforestation. This is already an equilibrium outcome, that is, after allowing for deforestation in the Xingu, locations inside and outside the Xingu will adapt moving to a new equilibrium. This substantial deforestation in the Xingu has spillover effects on areas beyond its borders. Given that changes in land use inside and outside of the Xingu are very different and to highlight the externality effect, in the figures that follow I focus on the areas outside of Xingu.

Figure 4A shows the change in forest choice probability between the old and new equilibrium. I interpret these changes as farmers abandoning agricultural activities. In total 10

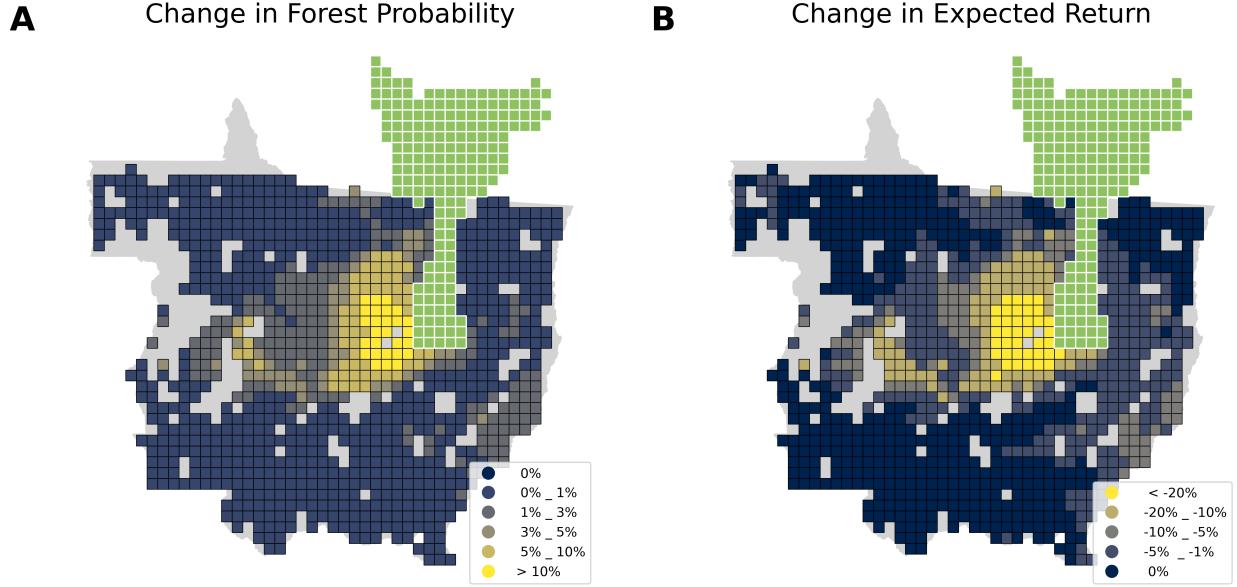


Figure 4: These maps show: (A) the change in forest choice probability between the two equilibria, I interpret this increase in forest probability as farmers abandoning agricultural activities; (B) the change in expected return between the two equilibria. Notice the stark spatial heterogeneity that is determined by atmospheric circulation. Pixels in green represent areas within the Xingu, where values are omitted to better display heterogeneity outside the Xingu region. Gray pixels indicate other protected areas.

thousand squared kilometers would be abandoned, but there is considerable spatial heterogeneity in the externality caused by upwind deforestation which is determined by atmospheric circulation. Thus, the determination of winners and losers of a policy that rolls back the protection status of the Xingu depends on patterns of atmospheric circulation. Notably, the biggest losers in this scenario are located in the center of Mato Grosso, the most productive region in the state. This exercise highlights the importance of the protection of the Xingu for the economy of one of the most important agricultural hubs in the world.

To translate this change in forest probability into loss of returns, Figure 4B shows the change in expected returns from land,¹³ which can be interpreted as a loss of land value. The spatial heterogeneity is very similar, with the center of Mato Grosso losing between 10% and 20% of expected return, that is between 10% and 20% of land value.

Ending the protection status of the Xingu has substantial effects on regional land use, which in turn affects climate. Figure 5 shows these impacts on rainfall and temperature in the new equilibrium. Some locations in northern Mato Grosso, those closest to the Xingu, lose more than 1 mm of rainfall and face an increase in temperature of more than 0.3 °C, with some locations approaching a rise of 0.5 °C. These climate changes are the mechanisms that drive the significant decline in expected land returns following extensive

¹³Given the structure of the model, expected return can be written as $\log \sum_{n \in \{C, F\}} e^{\pi_n}$

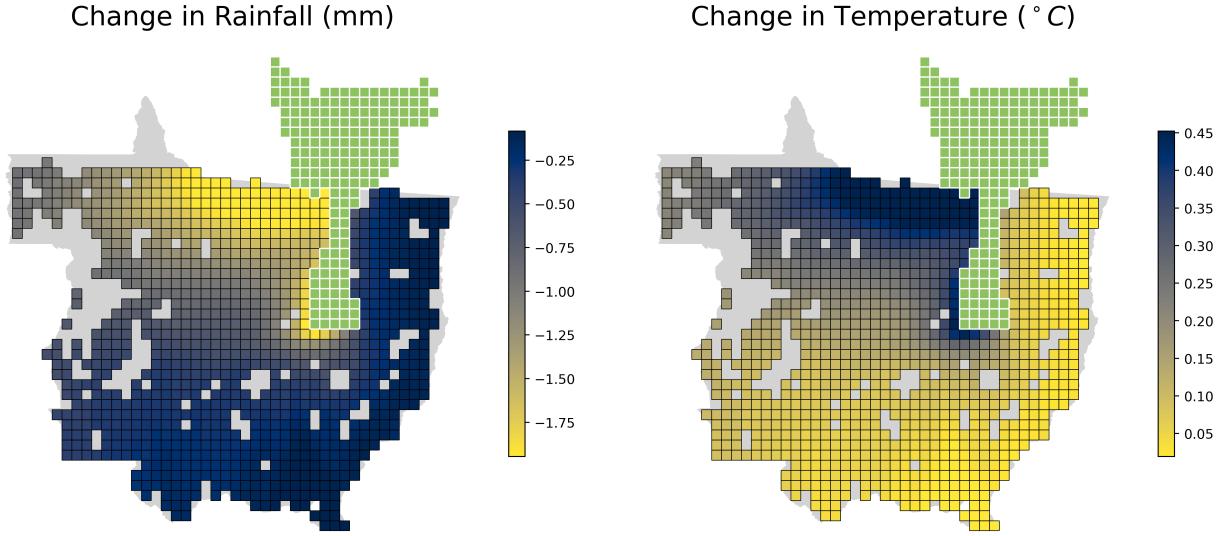


Figure 5: These maps show the impact of the new equilibrium, after rolling back the protection status of the Xingu, on rainfall (A) and temperature (B)

deforestation in the Xingu. Notice that the spatial patterns of the impacts in climate are exactly the same, since the variation used to compute them is the same, that of changes in upwind forest. That is why I use upwind forest directly as the main variable of interest in the discrete choice model of land use.

Note that the regions experiencing the most significant climate impacts are not necessarily those with the greatest changes in land value. A key mismatch appears when comparing Figures 4 and 5, particularly in northern Mato Grosso. While the extreme north of Mato Grosso faces substantial climate alterations, this area is less productive than the state's central region. Consequently, an identical climate change yields a larger impact on land values in central Mato Grosso.

After removing Xingu's protection status, 66,000 square kilometers of new agricultural land enter the market, leading to climatic changes that reduce rainfall and raise temperatures. This environmental shift causes a substantial portion of previously productive land — approximately 10,000 square kilometers — to be abandoned. What is the net result?

Naturally, this question must be approached with significant caveats. The net result presented here does not account for additional externalities, which I discuss in Section 4, nor does it consider the severe impacts that revoking Xingu's protection would impose on Indigenous communities. Instead, this result is constrained to the agricultural sector and the flying rivers mechanism alone. Nevertheless, such an exercise is valuable as it provides a lower bound on the bias that could arise if a researcher overlooks the endogeneity of climate in land-use policy counterfactuals.

Summing the increase in expected return within the Xingu region with the decrease in expected return outside the Xingu reveals that the externality offsets 40% of the increase in expected return from newly available land in the Xingu. That is, the decrease in expected return outside the Xingu equals 40% of the gain inside the Xingu. Thus, ignoring climate endogeneity leads to a substantial overestimation of the benefits of deforestation, or equivalently, a severe underestimation of its costs.

I have emphasized that the economics literature on deforestation has largely overlooked the endogeneity of climate. In contrast, the natural sciences have addressed this endogeneity, but typically lack a framework for modeling human behavior, which limits discussions on adaptation. In my model, adaptation occurs through land use choice, allowing shifts such as moving from soybeans to pasture grazing, as pasture is less sensitive to climate variability (see Table 2).

To evaluate the importance of considering adaptation, I compare, for locations outside the Xingu, a weighted return of land in the new equilibrium $\sum_{n \in \{F,C\}} P_o(n)(\pi_o^n + \gamma)$ with a weighted return of land where returns are as in the new equilibrium but farmers' choices remain as in the old equilibrium $\sum_{n \in \{F,C\}} P_o^{old}(n)(\pi_o^n + \gamma)$.¹⁴ This comparison allows me to compute the loss a farmer would incur if they deviated from optimal behavior and retained their original land use decision. The sum of utilities in the scenario without adaptation is 9% lower than in the scenario with adaptation. Such an aggregate effect, nonetheless, mask important spatial heterogeneity. Figure 6 presents the distribution of weighted returns under these two scenarios, with and without adaptation. Since returns are measured in utilities, which are difficult to interpret directly, I normalize the weighted returns by the standard deviation in the old equilibrium. Adaptation significantly impacts the distribution of weighted returns in the economy, particularly the lower tail. Therefore, a model that accounts only for changes in climate to mechanically compute changes in agricultural productivity would likely overestimate the adverse impacts of deforestation.

4 Caveats

As with any research paper, mine has limitations. Below I summarize the ones I find most important and most promising for future research.

Other effects of deforestation. My model does not consider other externalities of deforestation besides the rainfall/climate mechanism and thus is not suitable to be used as a stand-alone approach for cost-benefit analysis of deforestation. An important externality

¹⁴ γ is the Euler–Mascheroni constant.

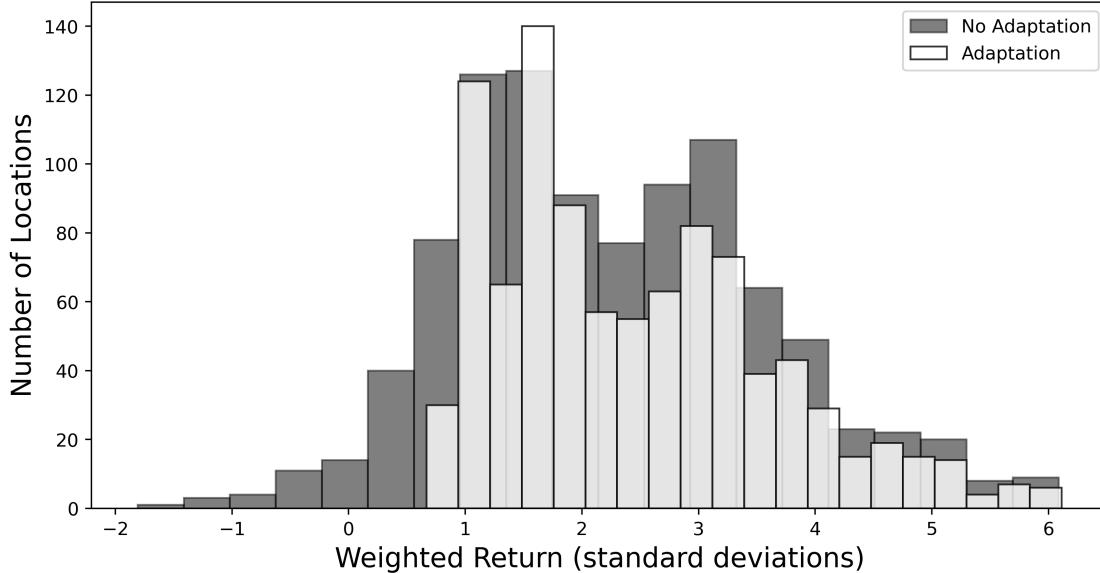


Figure 6: This figure shows two distributions of weighted returns. The white distribution, given by $\sum_{n \in \{F,C\}} P_o(n)(\pi_o^n + \gamma)$, represents the weighted return considering both returns and optimal choice probabilities in the new equilibrium after the deforestation in Xingu. The gray distribution, given by $\sum_{n \in \{F,C\}} P_{o^{\text{old}}}(n)(\pi_o^n + \gamma)$ represents the weighted return considering returns as in the new equilibrium after the deforestation in Xingu but holding choice probabilities fixed as in the old equilibrium. The difference between these two distributions shows the importance of accounting for adaptation.

of deforestation is the release of greenhouse gases. The 66 thousand square kilometers of deforestation in the counterfactual Xingu region is equivalent to the emission of 3.2 billion tons of CO_2 . With a *very* conservative social cost of carbon of US\$5 ton/ CO_2 , this amounts to US\$ 16 billion of externality cost.¹⁵ The same mechanism of deforestation affecting climate has effects on the energy sector as well, since hydropower energy is the main source of electricity in Brazil (see Araujo (2024)). The forest fragmentation in the Xingu in the counterfactual scenario could cause an abrupt decrease in biodiversity (Ferraz et al., 2003), generating a cost that economics is still struggling to measure. Deforestation makes extensive use of fires and thus creates pollution that affects human health (Rangel and Vogl, 2019; Rocha and Sant’Anna, 2022). Deforestation is also connected with illegal activities, violence, and conflict (Chimeli and Soares, 2017; Fetzer and Marden, 2017).

Dynamics. The deforestation literature presents static models both with cross-sectional (Souza-Rodrigues, 2019) and panel data (Domínguez-Iino, 2021), estimated dynamic models (Scott, 2014; Araujo et al., 2020; Sant’Anna, 2021) and calibrated dynamic general equilibrium models (Farrokhi et al., 2023). Dynamic models of land conversion, that nest static

¹⁵The value of 48,510 converts deforestation in km^2 to released tons of CO_2 (Technical Note n° 2093/2018-MMA from the Amazon Fund). US\$ 5 ton/ CO_2 is the carbon cost considered in the same note and used in the Amazon Fund.

models, show that dynamics lead to higher elasticities of deforestation (Scott, 2014; Araujo et al., 2020), but I cannot derive whether this would still be true in a setting with externalities, since a no spillover assumption is necessary to get rid of expectations in the estimation step proposed by Scott (2014). Additionally, multiplicity of equilibrium is likely to be a problem in such a setting. An important next step is to understand whether the contraction result presented in this paper can be extended to dynamic settings.

Tipping points. It has been hypothesized the existence of a point of no return in Amazon deforestation, that is, a level of deforestation from which the rainforest cannot create the humidity necessary for its own stability (Nobre et al., 1991; Sampaio et al., 2007; Gatti et al., 2021; Araujo et al., 2023; Flores et al., 2024). With a potential abrupt forest dieback, any model that does not account for it will largely underestimate the cost of deforestation. This is connected with the discussion on dynamics in the previous paragraph, but more complicated, since to incorporate tipping points it would be necessary to model expectations with an unknown probability of an abrupt state transition. However, as long as the deforestation in the counterfactual scenarios is not enough to trigger a tipping point, my model still offers a way to quantify externality effects of deforestation.

5 Conclusion

In this paper, I develop a model of land use in which climate is endogenous, focusing on changes in rainfall patterns caused by deforestation. I derive conditions for this model, which features a network of externalities, to have a unique equilibrium. As an application, I estimate the model with climate data from the Amazon Rainforest and land use data from the Brazilian state of Mato Grosso, one of the most important agricultural hubs in the world. Deforestation, as a result of land use expansion, causes a decrease in rainfall and an increase in temperature, which in turn affects land use decisions. An integrated model allows me to compute policy counterfactuals in an equilibrium setting, where land use choices and climate are consistent with each other.

I then explore a counterfactual where farmers are allowed to produce and deforest inside the Indigenous Territories of the Xingu River Basin. Allowing farmers into the Xingu region results in 66 thousand square kilometers of deforestation or 47% of the region. In this counterfactual, precipitation in Mato Grosso would decrease by more than 1 mm, and temperature would increase by more than 0.3 °C in the central and northern regions of Mato Grosso. In a model without endogenous climate, all of these effects would be absent, underscoring how far off are models that take climate as given. In such models, the costs of

deforestation are severely underestimated. I also show that farmers' re-optimization through changes in agricultural activities is an important channel of adaptation. Ignoring these effects, as often done in the natural sciences, would lead to an overestimation of the costs of deforestation through this mechanism of flying rivers. Thus, while economics has underestimated the costs of deforestation, natural sciences have overestimated it. The approach in this paper offers a way to reconcile the knowledge that has been generated by both fields.

My results highlight a way to integrate the interaction of human behavior and the natural environment into a single consistent model. The mechanism of deforestation affecting rainfall and other climate outcomes has consequences for different regions – such as Africa, Southeast Asia, and other parts of South America – and for different sectors – such as energy and water supply for urban centers. Each one of these consequences is a topic for future research.

References

- Araujo, R. (2024). The value of tropical forests to hydropower. *Energy Economics*, 129:107205.
- Araujo, R., Assunção, J., Hirota, M., and Scheinkman, J. A. (2023). Estimating the spatial amplification of damage caused by degradation in the amazon. *Proceedings of the National Academy of Sciences*, 120(46):e2312451120.
- Araujo, R., Costa, F., and Sant'Anna, M. (2020). Efficient forestation in the brazilian amazon: Evidence from a dynamic model.
- Assunção, J., Gandour, C., and Rocha, R. (2013). Deterring deforestation in the brazilian amazon: environmental monitoring and law enforcement. *Climate Policy Initiative*, 1:36.
- Assunção, J., McMillan, R., Murphy, J., and Souza-Rodrigues, E. (2019). Optimal environmental targeting in the amazon rainforest. Technical report, National Bureau of Economic Research.
- Balboni, C., Berman, A., Burgess, R., and Olken, B. A. (2023). The economics of tropical deforestation. *Annual Review of Economics*, 15(1):723–754.
- Bayer, P., Kozics, G., and Szőke, N. G. (2023). Best-response dynamics in directed network games. *Journal of Economic Theory*, 213:105720.
- Bramoullé, Y., Kranton, R., and D'amours, M. (2014). Strategic interaction and networks. *American Economic Review*, 104(3):898–930.
- Burgess, R., Costa, F., and Olken, B. A. (2019). The brazilian amazon's double reversal of fortune.
- Bustos, P., Caprettini, B., and Ponticelli, J. (2016). Agricultural productivity and structural transformation: Evidence from brazil. *American Economic Review*, 106(6):1320–65.
- Chimeli, A. B. and Soares, R. R. (2017). The use of violence in illegal markets: Evidence from mahogany trade in the brazilian amazon. *American Economic Journal: Applied Economics*, 9(4):30–57.
- Cinelli, C., Forney, A., and Pearl, J. (2024). A crash course in good and bad controls. *Sociological Methods & Research*, 53(3):1071–1104.

Copernicus, C. C. S. (2017). Era5: Fifth generation of ecmwf atmospheric reanalyses of the global climate.

Costa, M. H., Borma, L., Brando, P. M., Marengo, J. A., Saleska, S. R., and Gatti, L. V. (2021). Chapter 7: Biogeophysical cycles: Water recycling, climate regulation. In Nobre, C., Encalada, A., Anderson, E., Roca Alcazar, F. H., Bustamante, M., Mena, C., Peña-Claros, M., Poveda, G., Rodriguez, J. P., Saleska, S., Trumbore, S., Val, A. L., Villa Nova, L., Abramovay, R., Alencar, A., Rodríguez Alzza, C., Armenteras, D., Artaxo, P., Athayde, S., Barretto Filho, H. T., Barlow, J., Berenguer, E., Bortolotto, F., Costa, F. A., Costa, M. H., Cuvi, N., Fearnside, P. M., Ferreira, J., Flores, B. M., Friari, S., Gatti, L. V., Guayasamin, J. M., Hecht, S., Hirota, M., Hoorn, C., Josse, C., Lapola, D. M., Larrea, C., Larrea-Alcazar, D. M., Lehm Ardaya, Z., Malhi, Y., Marengo, J. A., Melack, J., Moraes, R. M., Moutinho, P., Murmis, M. R., Neves, E. G., Paez, B., Painter, L., Ramos, A., Rosero-Peña, M. C., Schmink, M., Sist, P., ter Steege, H., Val, P., van der Voort, H., Varese, M., and Zapata-Ríos, G., editors, *Amazon Assessment Report 2021*. United Nations Sustainable Development Solutions Network, New York, USA.

Costa, M. H., Borma, L., Brando, P. M., Marengo, J. A., Saleska, S. R., and V., G. L. (2012). Biogeophysical cycles: Water recycling, climate regulation. *Science Panel for the Amazon.*, Chapter 7.

Cruz, J.-L. and Rossi-Hansberg, E. (2024). The economic geography of global warming. *Review of Economic Studies*, 91(2):899–939.

Dell, M., Jones, B. F., and Olken, B. A. (2014). What do we learn from the weather? the new climate-economy literature. *Journal of Economic literature*, 52(3):740–798.

Desmet, K. and Rossi-Hansberg, E. (2024). Climate change economics over time and space. *Annual Review of Economics*, 16.

Domínguez-Iino, T. (2021). Efficiency and redistribution in environmental policy: An equilibrium analysis of agricultural supply chains.

Farrokhi, F., Kang, E., Pellegrina, H. S., and Sotelo, S. (2023). Deforestation: A global and dynamic perspective. *Working paper*, page 6.

Ferraz, G., Russell, G. J., Stouffer, P. C., Bierregaard, R. O., Pimm, S. L., and Lovejoy, T. E. (2003). Rates of species loss from amazonian forest fragments. *Proceedings of the National Academy of Sciences*, 100(24):14069–14073.

- Fetzer, T. and Marden, S. (2017). Take what you can: property rights, contestability and conflict. *The Economic Journal*, 127(601):757–783.
- Flores, B. M., Montoya, E., Sakschewski, B., Nascimento, N., Staal, A., Betts, R. A., Levis, C., Lapola, D. M., Esquivel-Muelbert, A., Jakovac, C., et al. (2024). Critical transitions in the amazon forest system. *Nature*, 626(7999):555–564.
- Gatti, L. V., Basso, L. S., Miller, J. B., Gloor, M., Gatti Domingues, L., Cassol, H. L., Tejada, G., Aragão, L. E., Nobre, C., Peters, W., et al. (2021). Amazonia as a carbon source linked to deforestation and climate change. *Nature*, 595(7867):388–393.
- Golden Krone, R. E., Qin, S., Cook, C. N., Krishnaswamy, R., Pack, S. M., Bonilla, O. D., Cort-Kansinally, K. A., Coutinho, B., Feng, M., Martínez Garcia, M. I., et al. (2019). The uncertain future of protected lands and waters. *Science*, 364(6443):881–886.
- Grosset, F., Papp, A., and Taylor, C. (2023). Rain follows the forest: Land use policy, climate change, and adaptation. *Climate Change, and Adaptation (December 1, 2023)*.
- Hsiao, A. (2021). Coordination and commitment in international climate action: evidence from palm oil. *Unpublished, Department of Economics, MIT*.
- Jarvis, A., Reuter, H. I., Nelson, A., Guevara, E., et al. (2008). Hole-filled srtm for the globe version 4. available from the CGIAR-CSI SRTM 90m Database (<http://srtm.csi.cgiar.org>), 15:25–54.
- Leite-Filho, A. T., Soares-Filho, B. S., Davis, J. L., Abrahão, G. M., and Börner, J. (2021). Deforestation reduces rainfall and agricultural revenues in the brazilian amazon. *Nature Communications*, 12(1):2591.
- Lemieux, P. and Luquira, K. (2021). Data stewardship maturity report for noaa climate data record (cdr) of leaf area index (lai) and fraction of absorbed photosynthetically active radiation (fapar), version 4.
- Lima, L. (2024). Flexible demand estimation and zero market shares. Working Paper.
- Makarieva, A. M. and Gorshkov, V. G. (2007). Biotic pump of atmospheric moisture as driver of the hydrological cycle on land. *Hydrology and earth system sciences*, 11(2):1013–1033.
- Marengo, J. A., Soares, W. R., Saulo, C., and Nicolini, M. (2004). Climatology of the low-level jet east of the andes as derived from the ncep–ncar reanalyses: Characteristics and temporal variability. *Journal of climate*, 17(12):2261–2280.

- Moon, H. R. and Weidner, M. (2015). Linear regression for panel with unknown number of factors as interactive fixed effects. *Econometrica*, 83(4):1543–1579.
- Nobre, A. D. (2014). The future climate of amazonia, scientific assessment report. *Sponsored by CCST-INPE, INPA and ARA, São José dos Campos Brazil*.
- Nobre, C. A., Sellers, P. J., and Shukla, J. (1991). Amazonian deforestation and regional climate change. *Journal of climate*, 4(10):957–988.
- Nordhaus, W. (2019). Climate change: The ultimate challenge for economics. *American Economic Review*, 109(6):1991–2014.
- Nordhaus, W. D. (1991). To slow or not to slow: the economics of the greenhouse effect. *The economic journal*, 101(407):920–937.
- Pack, S. M., Ferreira, M. N., Krishivasan, R., Murrow, J., Bernard, E., and Mascia, M. B. (2016). Protected area downgrading, downsizing, and degazettement (padd) in the amazon. *Biological Conservation*, 197:32–39.
- Pendrill, F., Gardner, T. A., Meyfroidt, P., Persson, U. M., Adams, J., Azevedo, T., Bastos Lima, M. G., Baumann, M., Curtis, P. G., De Sy, V., et al. (2022). Disentangling the numbers behind agriculture-driven tropical deforestation. *Science*, 377(6611):eabm9267.
- Rafey, W. (2020). Droughts, deluges, and (river) diversions: Valuing market-based water reallocation. Technical report, Working paper.
- Rangel, M. A. and Vogl, T. S. (2019). Agricultural fires and health at birth. *Review of Economics and Statistics*, 101(4):616–630.
- Rocha, R. and Sant'Anna, A. A. (2022). Winds of fire and smoke: Air pollution and health in the brazilian amazon. *World Development*, 151:105722.
- Salati, E., Dall'Olio, A., Matsui, E., and Gat, J. R. (1979). Recycling of water in the amazon basin: an isotopic study. *Water resources research*, 15(5):1250–1258.
- Sampaio, G., Nobre, C., Costa, M. H., Satyamurty, P., Soares-Filho, B. S., and Cardoso, M. (2007). Regional climate change over eastern amazonia caused by pasture and soybean cropland expansion. *Geophysical Research Letters*, 34(17).
- Sant'Anna, M. (2021). How Green Is Sugarcane Ethanol? *The Review of Economics and Statistics*, pages 1–45.

- Schlenker, W. and Roberts, M. J. (2009). Nonlinear temperature effects indicate severe damages to us crop yields under climate change. *Proceedings of the National Academy of sciences*, 106(37):15594–15598.
- Scott, P. (2014). Dynamic discrete choice estimation of agricultural land use. *Working paper*.
- Simoes, R., Picoli, M. C., Camara, G., Maciel, A., Santos, L., Andrade, P. R., Sánchez, A., Ferreira, K., and Carvalho, A. (2020). Land use and cover maps for mato grosso state in brazil from 2001 to 2017. *Scientific Data*, 7(1):1–10.
- Souza-Rodrigues, E. (2019). Deforestation in the amazon: A unified framework for estimation and policy analysis. *The Review of Economic Studies*, 86(6):2713–2744.
- Spracklen, D. V., Arnold, S. R., and Taylor, C. (2012). Observations of increased tropical rainfall preceded by air passage over forests. *Nature*, 489(7415):282–285.
- Staal, A., Tuinenburg, O. A., Bosmans, J. H., Holmgren, M., van Nes, E. H., Scheffer, M., Zemp, D. C., and Dekker, S. C. (2018). Forest-rainfall cascades buffer against drought across the amazon. *Nature Climate Change*, 8(6):539–543.
- Strassburg, B. B., Latawiec, A. E., Barioni, L. G., Nobre, C. A., Da Silva, V. P., Valentim, J. F., Vianna, M., and Assad, E. D. (2014). When enough should be enough: Improving the use of current agricultural lands could meet production demands and spare natural habitats in brazil. *Global Environmental Change*, 28:84–97.
- Wunderling, N., Staal, A., Sakschewski, B., Hirota, M., Tuinenburg, O. A., Donges, J. F., Barbosa, H. M., and Winkelmann, R. (2022). Recurrent droughts increase risk of cascading tipping events by outpacing adaptive capacities in the amazon rainforest. *Proceedings of the National Academy of Sciences*, 119(32):e2120777119.
- Xu, J., Gao, J., de Holanda, H. V., Rodríguez, L. F., Caixeta-Filho, J. V., Zhong, R., Jiang, H., Li, H., Du, Z., Wang, X., et al. (2021). Double cropping and cropland expansion boost grain production in brazil. *Nature Food*, 2(4):264–273.
- Yang, Y., Tilman, D., Jin, Z., Smith, P., Barrett, C. B., Zhu, Y.-G., Burney, J., D'Odorico, P., Fantke, P., Fargione, J., et al. (2024). Climate change exacerbates the environmental impacts of agriculture. *Science*, 385(6713):eadn3747.
- Zemp, D., Schleussner, C.-F., Barbosa, H., Van Der Ent, R., Donges, J. F., Heinke, J., Sampaio, G., and Rammig, A. (2014). On the importance of cascading moisture recycling in south america. *Atmospheric Chemistry and Physics*, 14(23):13337–13359.

Zemp, D. C., Schleussner, C.-F., Barbosa, H. M., Hirota, M., Montade, V., Sampaio, G., Staal, A., Wang-Erlandsson, L., and Rammig, A. (2017). Self-amplified amazon forest loss due to vegetation-atmosphere feedbacks. *Nature communications*, 8(1):14681.

A Appendix

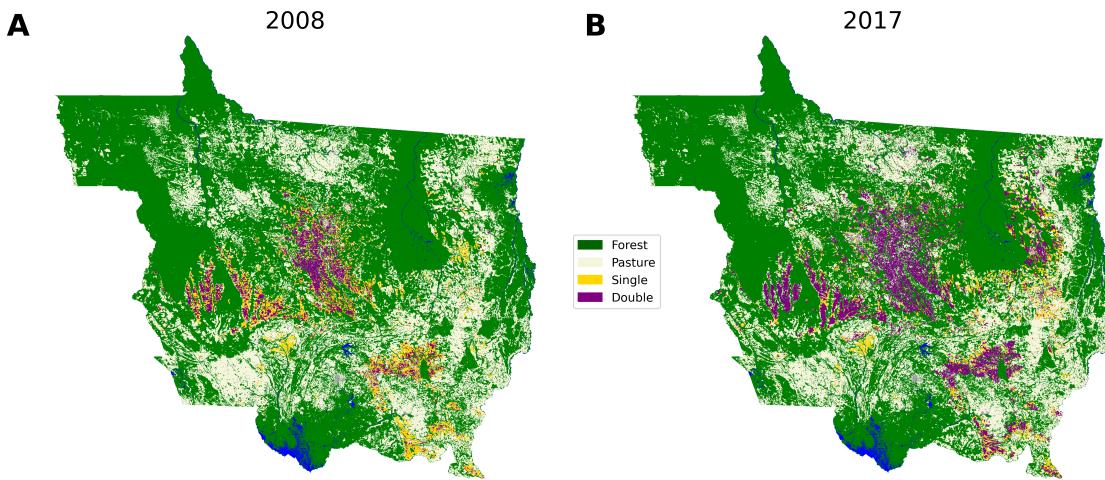


Figure A.1: These maps show the land use for the Brazilian state of Mato Grosso in 2008 and 2017. “Double” refers to a double cropping option (such as soybeans follow by corn), while “Single” refers to a single crop being produced (soybeans). Absent from the legend are blue pixels representing water and hard-to-see gray pixels representing urban areas.

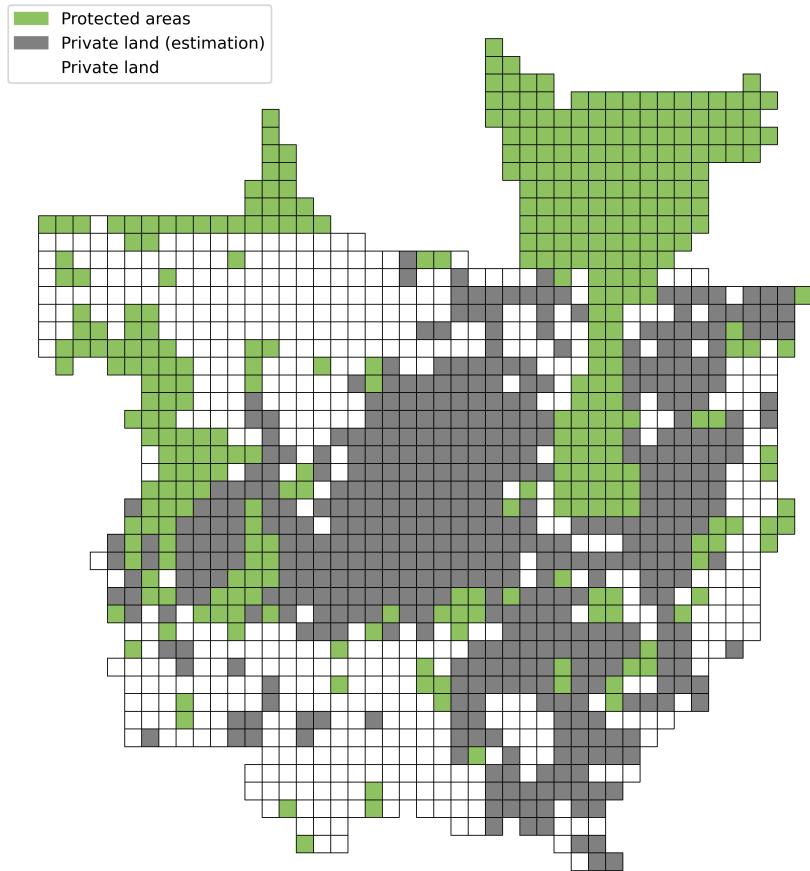


Figure A.2: This map shows the three types of pixels in our sample. Protected areas (in green) are used only in the counterfactual exercise and exclusively for the Xingu region. Private land (in white) with zero shares is not used in the estimation of the land use model but is included in counterfactual exercises. Private land without zero shares (in gray) is utilized in the estimation of the land use model and in counterfactuals.

Interaction	Variable	Coefficient	Std. Error
Pasture	Latitude	-2.986553	0.521394
	Longitude	-9.169095	1.628853
	Transport	-2.619496	0.733845
	Soil type 4	-0.127466	0.500518
	Soil type 2	-0.142535	0.313607
	Soil type 8	-0.150005	0.312569
	Soil type 1	0.800363	0.347370
	Soil type 7	0.175827	0.903351
	Soil type 6	0.067053	0.393952
	Soil type 13	0.947076	0.431755
	Soil type 14	1.022876	0.385188
	Soil type 16	-0.197015	0.308865
	Constant	-8.170254	1.448852
	Slope	-0.044121	0.035454
Single	Latitude	-21.015847	1.225425
	Longitude	217.887258	3.301841
	Transport	-10.695464	1.475721
	Soil type 4	1.269284	0.999273
	Soil type 2	-0.196944	0.696504
	Soil type 8	-0.613393	0.755327
	Soil type 1	0.123917	0.749057
	Soil type 7	0.819978	2.007220
	Soil type 6	-3.572608	0.839507
	Soil type 13	1.719118	1.018191
	Soil type 14	1.265170	0.821011
	Soil type 16	1.921517	0.715305
	Constant	163.985110	2.989359
	Slope	0.072081	0.087876
Double	Latitude	-12.944141	1.273209
	Longitude	136.871474	3.496011
	Transport	-12.545499	1.502797
	Soil type 4	-1.331484	1.021071
	Soil type 2	1.159280	0.763985
	Soil type 8	0.418201	0.839774
	Soil type 1	0.372551	0.882929
	Soil type 7	-1.316028	2.226889
	Soil type 6	-3.059022	0.975487
	Soil type 13	-0.446595	1.040595
	Soil type 14	1.615170	0.863284
	Soil type 16	3.431285	0.785935
	Constant	104.024958	3.121780
	Slope	0.029780	0.109733

Table A.1: This table shows the coefficients of the predictive step of my estimation. The predictive step has R^2 from 5-fold cross-validation of 95%, but in this table I show the in-sample coefficients, which are the ones used in my counterfactual exercise. Standard errors are clustered at the location level, which for this exercise in the cross-section is the heteroskedastic robust standard error.