Does Conservation Work in General Equilibrium?*

Verónica Salazar Restrepo† and Gabriel Leite Mariante‡

December 8, 2023

PRELIMINARY VERSION: This draft is edited frequently. See latest version HERE.

Abstract

Deforestation and the subsequent use of deforested land for agricultural activities account for roughly 20% of the global CO2-equivalent emissions in the past two decades. Despite the global scope of the consequences of deforestation, public policies and private initiatives to reduce deforestation are often spatially targeted: they intensify environmental protection in specific ecosystems, making agricultural land scarcer. While potentially effective at a local level, their global effectiveness may be attenuated in general equilibrium, due to resulting increases in the demand for agricultural land in non-targeted areas, i.e. deforestation leakage. To quantify leakage, build a quantitative spatial equilibrium model of the Brazilian economy where agricultural land is the output of a costly process of deforestation, firms produce goods that are differentially land-demanding, and there is costly trade and migration. Our main findings are that (i) targeting the regions with highest deforestation levels can be an effective tool to curb aggregate deforestation in Brazil, and (ii) leakage increases significantly when considering a longer time-horizon. After one year, 2-3% of the deforestation reductions are outdone by leakage. Simulating the model forward for 10 years, this number goes up to 10%. The relatively small leakage is driven by agricultural intensification, including more crop farming, increased worker and cattle density per pasture, and shifts of production towards more productive regions.

*For their generous mentorship, we thank Oriana Bandiera, Gharad Bryan, Robin Burgess, and Daniel Sturm. For all their comments and questions, we thank seminar audiences at LSE's Development Economics Work in Progress, LSE's Energy & Environmental Economics seminar, LSE's International Trade Workshop, and conference audiences at STEG-CEPR 2023, LSE's Environment Week 2023, Bocconi (EMUEA 2023), Cyprus (EAERE 2023), and Oxford (OxDev 2023). For their time and insightful discussions, we thank Clare Balboni, Tim Doberman, Matthias Doepke, James Fenske, Douglas Gollin, Allan Hsiao, Juliana Jaramillo Echeverri, Xavier Jaravel, Per Krusell, Isabella Manelici, Pol Simpson, Xavier Tierney, Julien Wolfersberger.

†Department of Economics, London School of Economics. E-mail: v.salazar-restrepo@lse.ac.uk.

‡Department of Economics, London School of Economics. E-mail: g.leite-mariante@lse.ac.uk.

1 Introduction

Tropical deforestation is among the human activities with the highest environmental impact. Forest clearing and the subsequent land use for agricultural activity is responsible for about one-fifth of global CO2-equivalent emissions of the past two decades. In addition to direct instantaneous carbon emissions, deforestation permanently destroys carbon sinks, extinguishes wildlife, degrades native soil, alters weather patterns, and negatively impacts the livelihoods of millions of humans who live in communities that depend on forest dwelling.

Given its enormous environmental damage, combating deforestation is a key component of emission reduction pathways in the global fight to mitigate the effects of climate change, as established by the Intergovernmental Panel on Climate Change (IPCC) (IPCC, 2019)[1]. Nevertheless, strategies to mitigate deforestation are excluded from major green finance mechanisms, such as the United Nations Clean Development Mechanism (CDM), due to concerns about leakage of locally targeted policies into untargeted areas - i.e. the reduction of deforestation in a specific area might be outdone by increases in deforestation elsewhere, undermining the global effects of local policies[2].

Since approximately three-quarters of global deforestation is driven by agriculture, quantifying general equilibrium effects of anti-deforestation policies requires a model of how agriculture is redistributed across space in response to local policies. To that end, we build a quantitative model that explicitly embeds deforestation as an economic sector that supplies land as a factor of production for agriculture. Beyond the direct policy relevance of assessing the global efficiency of localised policies, our framework can guide the understanding of broader trade-offs in environmental policy (promoting conservation vs curtailing agricultural production), and the value of mitigating global externalities of localised forest-cutting actions[3].

Deforestation in Brazil

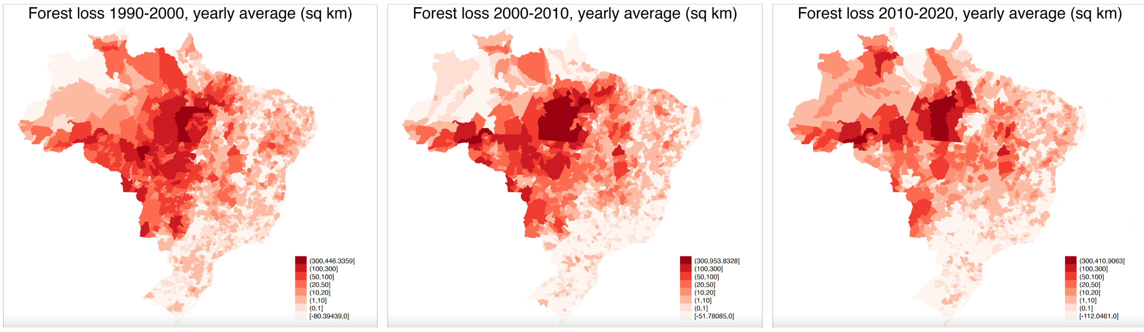
With over 5 million km^2 of rainforest area (MapBiomas Collection 6.0), Brazil accounts for about one third of all remaining rainforests and 13% of all forests in the planet, far more

[1]In its 2019 Special Report on Climate Change and Land, the IPCC mentions the word "deforestation" a total of 493 times, and states that *"Reducing deforestation and forest degradation lowers GHG emissions, with an estimated technical mitigation potential of 0.4–5.8 GtCO2/yr"*

[2]The UN defines deforestation leakage as *"The unexpected loss of anticipated carbon benefits due to the displacement of activities in the project area to areas outside the project, resulting in carbon emissions."*. In other words, locally targeted policies might not decrease overall deforestation, but rather displace it to untargeted areas.

[3]This framework could be used, for example, to implement a compensation policy for specific regions or for the country as a whole, a policy that has been suggested by a number of researchers and politicians as the only viable way to sustainably avoid deforestation (Brner and Wunder, 2008).

Figure 1: Average yearly forest loss by municipality



Note: This figure shows average forest loss in square kilometers across all 5,568 Brazilian municipalities over three 10-year periods: 1990-2000, 2000-2010, 2010-2020. Municipalities are coloured according to total forest loss over each time period, with darker colours indicating higher forest loss.

than any other country. It also accounts for a third of yearly tropical deforestation, with an average net yearly forest loss of about 20,000 km^2 between 1985 and 2020[4] (roughly equivalent to the size of Wales, or the American state of Connecticut). As such, outcomes of policies aiming at reducing deforestation in Brazil will have significant impact in the global forest coverage and, consequently, carbon emissions.

Deforestation has played a central role in the Brazilian economy over the past 35 years as an input for the country's agricultural sector, currently responsible for approximately one-quarter of the Brazilian GDP (CNA and CEPEA, 2023). Since 1985, the first year for which reliable high resolution satellite data on land use is available, total agricultural area has increased by more than 45%, or 840,000 km^2 (MapBiomas Collection 6.0), of which 92% were previously forests and 8% were previously tropical savannas. As illustrated in Figure 1, this activity is highly spatially concentrated along the so-called "arc of deforestation", in the fringe the Amazonian region, especially in the states of Pará, Mato Grosso and Rondônia. The large municipality of São Félix do Xingu, in the state of Pará, with a total area of around 84,000 km^2 (roughly the size of Austria), is an excellent illustrative example. In 1985, it had a forest cover of over 80,000 km^2, equivalent to more than 95% of its total area, and agricultural activity occupied only 410 km^2 (0.5%). In 2021, it had just under 63,000 km^2 (under 75% of its area) of forest cover, and over 19,000 km^2 (well over 20%) of agricultural land, most of which pastures.

Over the past three decades, the Brazilian government has enacted a wide array of policies aiming at tackling deforestation. These policies include country-level actions, such as changes in the national legislation (e.g. Forest Code) or the use of satellite monitoring

[4] The gross loss of primary forest has been higher, over 25,000 km^2 per year, but there is considerable amount of forest regrowth.

(SIVAM - from the 90s), but also localised policies aiming at protecting specific areas of the Brazilian territory. The latter group, which is the focus of our study, has typically been implemented either as increased enforcement in areas with particularly high deforestation, or as the designation of specific areas in the territory as conservation units or protected indigenous land. Such localised policies have been the subject of past evaluation which has generally found them to be successful in the targeted areas (Assunção and Rocha, 2019).

Typically, studies evaluating such policies use an event-study approach to measure their effectiveness in the targeted regions vis-à-vis comparable non-targeted regions (Assunção and Rocha, 2019). However, this approach does not account for potential relocation of deforestation activities into non-targeted areas, which can significantly attenuate the effect of such policies on overall country-wide deforestation. To illustrate this idea, suppose that regions targeted by a certain policy see a decrease of 5,000 km^2 in their annual deforestation, whereas comparable non-targeted regions see an increase of 5,000 km^2 in annual deforestation rates – i.e. there is perfect relocation. A simple difference-in-differences analysis would suggest that the policy decreased deforestation by 10,000 km^2 when in fact it only changed where it occurred, and the global effect is zero. This issue, referred to in the climate policy literature as leakage, is often discussed (Pfaff and Robalino, 2017) but rarely measured when evaluating specific anti-deforestation policies[5].

The Model

Our model considers a multi-region economy with two main sectors: agriculture and non-agriculture. The agricultural sector is further split into different types of crops and pastures that demand different amounts of land and labour as factors of production, whereas the non-agricultural sector has only labour as input. Crucially, we model deforestation as an intermediate sector which endogenously supplies land as a factor of production for agriculture. Each location differs in their sectoral productivity (agricultural, non-agricultural and deforestation), amenities and trade links. We allow workers to migrate between regions subject to frictions, and goods to be traded across regions subject to iceberg costs.

In the model, a local anti-deforestation policy is modelled as an exogenous negative shock to land supply in a targeted region. Leakage happens due to relocation of agricultural activity via the markets for goods and labour. In the goods market, a local decrease in supply of land will decrease the local supply of agricultural goods. Consumers then substitute these

[5] It has become increasingly common to measure spillovers by looking at the changes in land use in the vicinity of protected areas, for example. While a valuable empirical exercise, it might not appropriately account for global spillovers, as the most suitable substitute for the land being protected may be in an entirely different part of the country - in this approach, the choice of which areas can be the subject of spillovers needs to be made *ex-ante* and is arbitrary.

goods with non-agricultural goods, or with agricultural goods produced elsewhere, increasing demand for land in other regions, which creates leakage. Analogously, in the labour market, a decrease in supply of land will decrease the local demand for agricultural labour – workers will then partly change sectors, and partly migrate increasing the supply of agricultural labour elsewhere, which also creates leakage. The extent to which reductions in forest loss are outdone by leakage ultimately depends on the elasticity of demand for agricultural goods, and on the substitutability of agricultural land across space.

Literature

Our paper contributes to the growing literature on the evaluation of anti-deforestation policies (Burgess et al., 2012; Jayachandran et al., 2017; Szerman et al., 2022). Existing evidence suggests that command-and-control policy instruments have played a crucial role in the slowdown of deforestation in Brazil, lowering it by as much 56% from what if would have been in the absence of policies that started to be implemented around 2004 (Assunção et al., 2015; Burgess et al., 2019).

We contribute to this strand of the literature by: (i) disentangling the many driving forces of deforestation (regional comparative advantage in agricultural activities, market access, residential amenities, and deforestation productivity), and (ii) quantifying general-equilibrium spillovers, whereby the reduction of deforestation in one area is accompanied by its increase elsewhere. Pfaff and Robalino (2017) summarise the theory and evidence on the spillovers of conservation programs which, at the time, lacked any structural quantitative analysis of spillovers due to general equilibrium effects. Instead, it looked at spillovers to neighbouring areas outside conservation zones.

A more recent strand of the trade and IO literature looks at deforestation in general equilibrium (Souza-Rodrigues, 2018; Hsiao, 2021; Domiguez-Iino, 2021) by drawing on foundational insights from theoretical models of trade and environmental conservation (Copeland and Taylor, 2004). A related literature analyses how comparative advantage in agricultural activities shapes the spatial distribution of different land uses, generally without considering deforestation explicitly (Cui, 2020; Pellegrina and Sotelo, 2021).

We contribute to this literature by adding labor and migration to general equilibrium models of land use change, and by explicitly linking land use changes with patterns of economic growth and structural change in space (Eckert and Peters, 2022; Farrokhi and Pellegrina, 2020; Bustos et al., 2016; Herrendorf et al., 2014; Boppart, 2014) .

Methodologically, our model expands on the spatial equilibrium model proposed by Eckert and Peters (2022). We build on their framework by (i) adding a deforestation sector that endogenously produces new agricultural land, and (ii) considering different agricultural sectors that can have heterogeneous shares of land and labour.

The rest of this paper is structured as follows. Section 2 describes the institutional context and estimates the local effects of existing targeted conservation strategies with a reduced form approach as motivating evidence. Section 3 describes the model along with the estimation and calibration of the various model parameters. Section 4 presents the results of the counterfactual analysis. Section 5 concludes.

2 Local Effects of Targeted Conservation

Existing research suggests that Brazil's approach to tackling deforestation over the first two decades of the 21st century has been successful in reducing country level deforestation, (Burgess et al., 2019; Assunção et al., 2015) . During this period, a vast array of policies were implemented by the federal government. Some of these are "blanket-level" policies affecting the entire Brazilian territory, such as changes in laws increasing legal limits for deforestation in private land (DOU, 2012) and the adoption of a unified *Action Plan for the Prevention and Control of Deforestation in the Legal Amazon* (PPCDAm). Others, however, were locally targeted policies, which raises the concerns of possible leakage: has the forest loss avoided in these areas simply moved somewhere else?

Among the local anti-deforestation policies enacted by the Brazilian government, two have been especially prominent. The first is the establishment of "Priority Municipalities" policy, enacted by the Brazilian government in 2007 (DOU, 2007). A set of municipalities in the Amazon region where deforestation rates were among the highest in the country were selected to be subject to extra enforcement actions. In a first round, 36 municipalities accounting for around 45% of the previous year's deforestation were included in the priority list, which has been updated on a yearly basis ever since. Selected municipalities were subject to increased law enforcement activities such as fines, embargoes on private farms, political agreements with local leaders, and credit incentives from the federal government. After the start of the policy, yearly deforestation in targeted areas significantly decreased (Assunção and Rocha, 2019). The second prominent example of local conservation policies is the continuous establishment by the Brazilian Government of specific areas where deforestation is completely banned. Such areas are established either for wildlife and biodiversity conservation – the so called Unidades de Conservação (Conservation Units), or for preservation of land that has been traditionally inhabited by native indigenous people – the so called Territórios Indígenas (Indigenous Territories). These policies began with the country's return to democracy in 1985 and, as of 2022, a total area of approximately 3,500 km^2 has been granted one of these two status. Due to their practical similarities, both types of policies will henceforth be referred to as Protected Areas. Unlike the Priority Municipalities policy, Protected Areas typically have high forest coverage and little deforestation, and the goal is to conserve the natural biome rather than to crack down on existing degradation.

Before turning to a quantitative analysis of leakage in place-based policies, it is important to discuss the reasons why, given the local effectiveness of targeted policies, one does not simply extend these policies to cover the entire country which would, by construction, end deforestation (and, trivially, leakage) altogether. Firstly, the enforcement of conservation policies is costly for public funds – research estimates that, in order to place 80% of the Brazilian Amazon under some form of currently existing policy, the federal government would need to spend at least 1.7 Billion USD per year (da Silva et al., 2022). Secondly, from a social welfare perspective, the conversion of forested land into agriculture generates private profits, which could result in a non-zero optimal level of deforestation from the perspective of Brazil. Thirdly, and related to the private optimal argument, there is a political cost of restricting deforestation activities, since the proceeds of those are often captured by local political elites. Fourthly, from a public revenues perspective, modern agriculture that results from forest removal is more easily measured and, consequently, taxed, than alternative sustainable economic activities

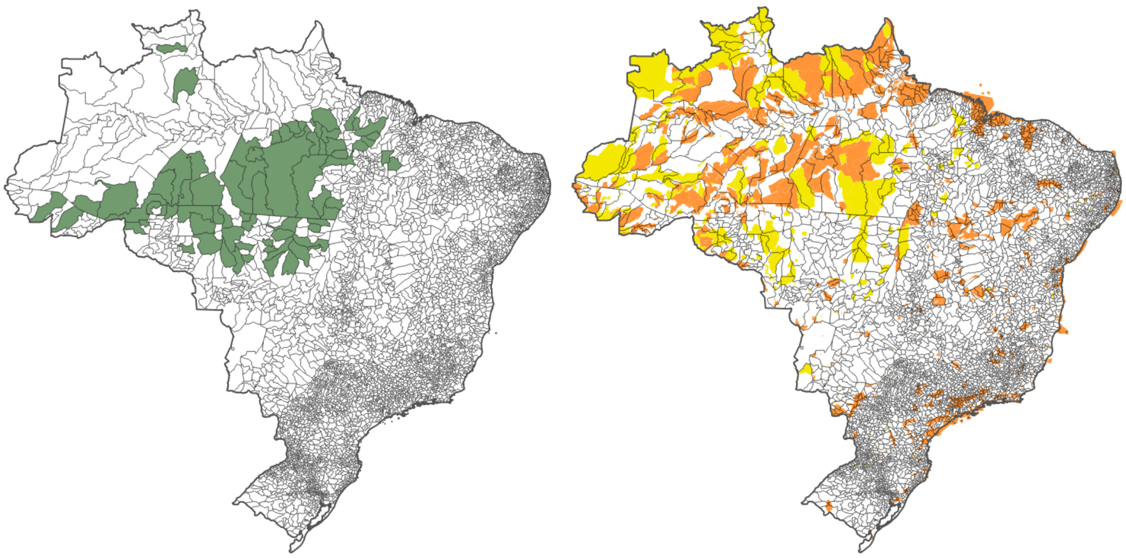
Reduced-form evidence from studies evaluating the effects of place-based conservation policies in Brazil suggest that they have been effective in decreasing deforestation in targeted areas. However, in order to capture the true effect of a local policy relative to a counterfactual in which this policy was never implemented, reduced-form methods rely heavily on the stable unit treatment value assumption (SUTVA) - i.e. absent treatment, non-treated areas would have behaved exactly as they did in the presence of treatment elsewhere. In the case of a conservation policy that can be subject to leakage, this assumption might not hold, as shown next.

2.1 Descriptives

Figures 2 and 3 illustrate the evolution of place-based anti-deforestation policies in Brazil between 1985 and 2020. From Figure 2, we can see that both policies are largely focussed on areas in Amazonian region. The Priority List (in green) targets exclusively municipalities within the Brazilian Legal Amazon, most of which located in the so called "deforestation arc" covering the south of the states of Amazonas and Pará and the north of the state of Mato Grosso. Although the Protected Areas are somewhat more spread across the country[6], they are still spatially concentrated in the Amazon biome, north of the municipalities in

[6] Mainly due to territories historically occupied by indigenous peoples in the Northeast and Southeast regions

Figure 2: Spatial distribution of Priority List municipalities (green), Conservation Units (orange), and Indigenous Territories (yellow) as of 2021.



Note: This figure illustrates the spatial targeting of the three relevant types of protected areas implemented by the Brazilian government over the past decades. The left panel shows, in green, the municipalities that have been added to the Priority List since its start in 2008. The right panel shows the location of protected areas (which do not necessarily coincide with municipal borders) according to its type: orange areas are Conservation Units, yellow areas are demarcated Indigenous Territories

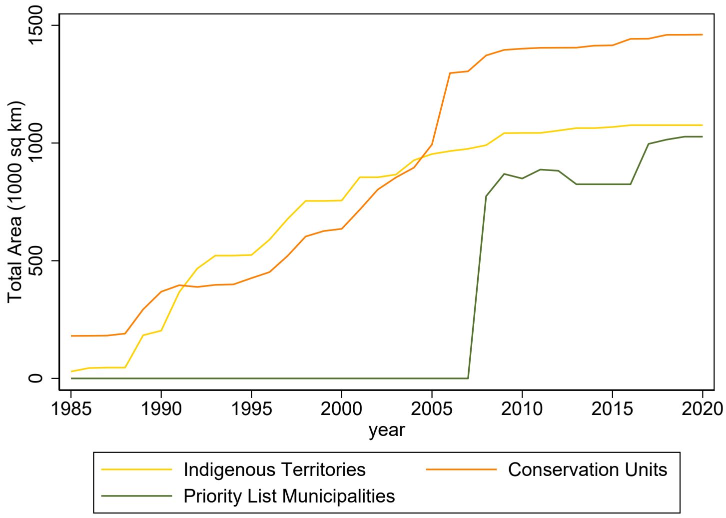
the Priority List, in areas with more untouched forest cover. In Figure 3, we can see the temporal evolution of these policies. Protected areas have been gradually established since the country's redemocratisation in 1985, whereas the Priority List policy was created in 2007, with most municipalities being added to the list in that same year.

Figure 4 below illustrates trends in forest change in Brazil according to their status as protected areas in 2020: indigenous territories, conservation units, and the rest of country. It is clear that, between 1985 and 2020, the bulk of forest loss happened outside protected areas. Throughout the period, forest cover in indigenous territories and protected areas have remained above 90% and 80% respectively, compared to those outside of them, which have gone from above 60% to under 50%.

However, this fact alone is not convincing evidence that they have worked as an effective halt on deforestation. The first concern with this simple comparison is the endogeneity of their placement - this is often referred to in policy circles as the question of "additionality". The second concern when evaluating the effectiveness of any local conservation policy is the possible presence of spillover effects - in particular the question of "leakage".

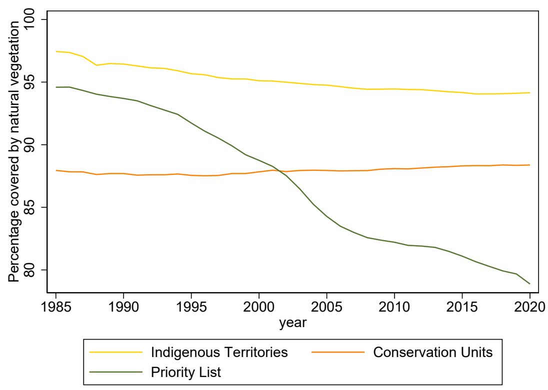
Regarding endogeneity, the bias could be positive or negative. These territories may have been established in places without good soils for agriculture, far from roads and markets,

Figure 3: Evolution of spatially targeted conservation policies



Note: This figure illustrates the evolution over time of the total area under different types of locally targeted policies by the Brazilian government. The yellow line represents the total area of demarcated Indigenous Territories, the orange line represents the total area under Conservation Units, and the green line represents the total area of all municipalities included in the Priority List policy.

Figure 4: Area covered in natural vegetation under different conservation policies



Note: This figure illustrates the evolution over time of the territorial coverage in natural vegetation, as a share of total area, for regions under the different types of locally targeted policies implemented the Brazilian government. The yellow line represents the evolution of natural vegetation coverage in Indigenous Territories, the orange line represents represents the evolution of natural vegetation coverage in Conservation Units, and the green line represents the eevolution of natural vegetation cover in municipalities from the Priority List.

or with geographic conditions that make them harder to deforest. But it could also go the opposite way, if environmental agencies respond to environmental threats by establishing protected areas and indigenous peoples decide to formalise protection as a reaction to settlers arriving and cutting down the forests in their land. It could go one way in places where the local politicians and bureaucrats are more aligned with the interests of agricultural businesses, and the other way where they are operating with the goal of stopping deforestation. In order to evaluate the effectiveness of the establishment of protected areas and indigenous territories in halting deforestation, we exploit their staggered establishment and look at places near their borders with a regression discontinuity design.

Most of the area under indigenous territories and protected areas has been assigned after 1985, which is when our land use data, taken begins. The graph below shows how the area under each of those property right regimes has changed over time.

The blacklist started in 2008. It is a subset of municipalities in the Legal Amazon that were chosen for their strategic importance in halting deforestation. The total deforestation levels the five years before are good predictors of blacklist status, although they do not seem to exactly determine blacklisting. Figure 3 illustrates how the blacklist has changed over time. In 2008, 36 municipalities were added to the list, covering a total area of approximately 800,000 square kilometers. Since then, the total number of blacklisted municipalities has not changed vary much, remaining around 40. The total area has increased to around 1,000,000 square kilometers. Some municipalities have entered and some have exited the blacklist. Upon exit, previously blacklisted municipalities acquire the status of "monitored". Since 2008, 20 municipalities have exited the blacklist and 3 of those have returned. There have been 2 periods of significant change in the blacklist. First, the period 2008-2012, with most of the action concentrated in 2008. Second, the period 2017-2020, with most of the action concentrated in 2017, when 9 municipalities covering over 200,000 square kilometers were added. The second period reflects a shift of deforestation further north, towards the Amazon river.

2.2 Econometric specification

2.2.1 Priority List

To evaluate the effectiveness of placing municipalities in a Priority List, we follow (Assunção and Rocha, 2019) and conduct a standard two-way fixed effects event study around the year in which municipalities were added to the priority list. Our baseline specification is as follows:

$$Defor_{rt} = \delta_t + \gamma_r + \sum_{\tau=-N_L}^{N_F} \beta_\tau Priority_{r,t-\tau} + \epsilon_{rt} \tag{1}$$

Where $Defor_{rt}$ is log-deforestation observed in municipality r at year t, δ_t are year fixed-effects, γ_r are municipality fixed effects, $Priority_{r,t-\tau}$ is a dummy variable equal to one if municipality r has been added to the priority list exactly τ years ago, and β_τ are our coefficients of interest. We consider event-times between 7 years before (N_L) and 10 years after (N_F) policy implementation. In order to address potential issues caused by staggered treatment adoption in two-way fixed effects event studies, we keep in our sample only municipalities that have either been added to the Priority List in 2008 or have never been added to the list, excluding municipalities added to the priority list after 2008.

Validity – The validity of our event-study approach relies on the assumption of parallel trends, i.e., treated and untreated municipalities followed parallel trends in deforestation rates in the years leading up to 2008, which implies $\beta_\tau = 0 \; \forall \; \tau \; \in [-N_L, 0)$. Given that the Priority List policy was explicitly targeted at municipalities considered "deforestation hotspots", it is plausible that treated municipalities followed different trends by design. In this case, the estimates resulting from a standard event-study approach may be downwards biased. To deal with this issue, we complement our analysis by estimating a difference-in-differences equation with synthetic controls, where each treated municipality is matched to a linear combination of untreated municipalities that results in equal pre-treatment trends, as per the equation below:

$$Defor_{rt} = \delta_t + \gamma_r + \beta_\tau Priority_{rt} + \epsilon_{rt} \tag{2}$$

2.2.2 Protected Areas

Unlike the Priority List policy, the establishment of protected areas (both Conservation Units and Indigenous Territories) does not necessarily coincide with municipal borders. Hence, we split the Brazilian territory in 100, hexagons with 10km width, and classify each of them as being or not part of a protected area if at least 50% of it's surface falls within the demarcated boundaries.

We use the hexagon-level data to estimate a Regression Discontinuity Design around the border of the conservation unit. Our baseline specification is:

$$ForestArea_{rt} = \delta_t + \gamma_m + f(Dist_{rt}) + \beta D_{rt} + \epsilon_{rt} \tag{3}$$

Where $ForestArea_{rt}$ is total forest area in logs observed in hexagon r at year t, δ_t are year fixed-effects, γ_m are municipality fixed effects, $f()$ is a continuous function, $Dist_r t$ is the running variable measuring the distance in km to the border of the nearest conservation unit where negative (positive) values mean that the hexagon falls inside (outside) the protected area, and $D_r t$ is a dummy variable indicating whether hexagon r falls within a protected area at time t. The coefficient of interest in β, identifying the effect on deforestation of being inside the conservation area. We use forest area instead of deforestation as an outcome because protected areas are typically established in regions with very low levels of deforestation, so the numeric interpretation of the effect on forested area is clearer. We follow Calonico et al. (2014) and implement optimal bandwidth selection to choose the maximum distance to the demarcated conservation border to be considered.

Validity − The validity of our RD design relies on the assumption of continuity of the outcome variable with respect to the running variable in the absence of treatment. In other words, in the absence of the protected areas, deforestation does not see a discontinuous spatial jump at $D_{rt} = 0$. To test this assumption, we estimate the same Regression Discontinuity design considering only hexagons located in future protected areas, but before they were established.

SUTVA violation

As previously discussed, an assumption required for the validity of both reduced-form methods discussed above is SUTVA. Consider the treatment D_i a binary variable equal to 1 if a place is subject to a conservation policy, 0 otherwise. Assume that the potential outcome of i depends on two factors, its own conservation status, and the price of land, which depends on the conservation statuses in all regions $Y_i(D_i, P(\vec{D}))$. Even if treatment was as good as randomly assigned, but SUTVA did not hold, we would be estimating a combination of the desired treatment effect and leakage as show below

$$\begin{aligned}
\widehat{\beta} \rightarrow & \mathbb{E}\big[Y_i(1, P(\vec{D}_{policy}) - Y_i(0, P(\vec{D}_{policy})\big] \\
& = \underbrace{\mathbb{E}\big[Y_i(1, P(\vec{D}_{nopolicy}) - Y_i(0, P(\vec{D}_{nopolicy})\big]}_{\text{Pure effect of policy on treated}} \\
& + \underbrace{\mathbb{E}\big[Y_i(1, P(\vec{D}_{policy}) - Y_i(1, P(\vec{D}_{nopolicy})\big]}_{\text{Leakage on treated}} \\
& - \underbrace{\mathbb{E}\big[Y_i(0, P(\vec{D}_{policy}) - Y_i(0, P(\vec{D}_{nopolicy})\big]}_{\text{Leakage on untreated}}.
\end{aligned}$$

A reduced-form estimate, therefore, would include both the desired ATE on the Treated and potential leakage onto untreated areas. Therefore, it requires an assumption about the nature of the second right-hand side term in the equation above. Typically, the implicit assumption is that there is no leakage, i.e. $Y_i(d, P_1) = Y_i(d, P_2)$, which results in an overestimation of the true treatment effect.

Even if leakage is explicitly considered, a reduced-form estimate requires assuming some arbitrary *ex-ante* structure determining which regions are prone to leakage, and which regions can be considered a "pure control". One common approach is to assume that neighbouring areas are prone to leakage, whereas places further away from treatment are not affected in any way. Whereas this might be plausible from a purely spatial perspective, it has two problems: (i) it still requires some arbitrary distance cut-off that separates areas that are susceptible to leakage from areas that are not, and (ii) it ignores other spillover mechanism, such as price changes, transportation connections and local amenities that might also be important determinants of the location of spillovers.

2.3 Results

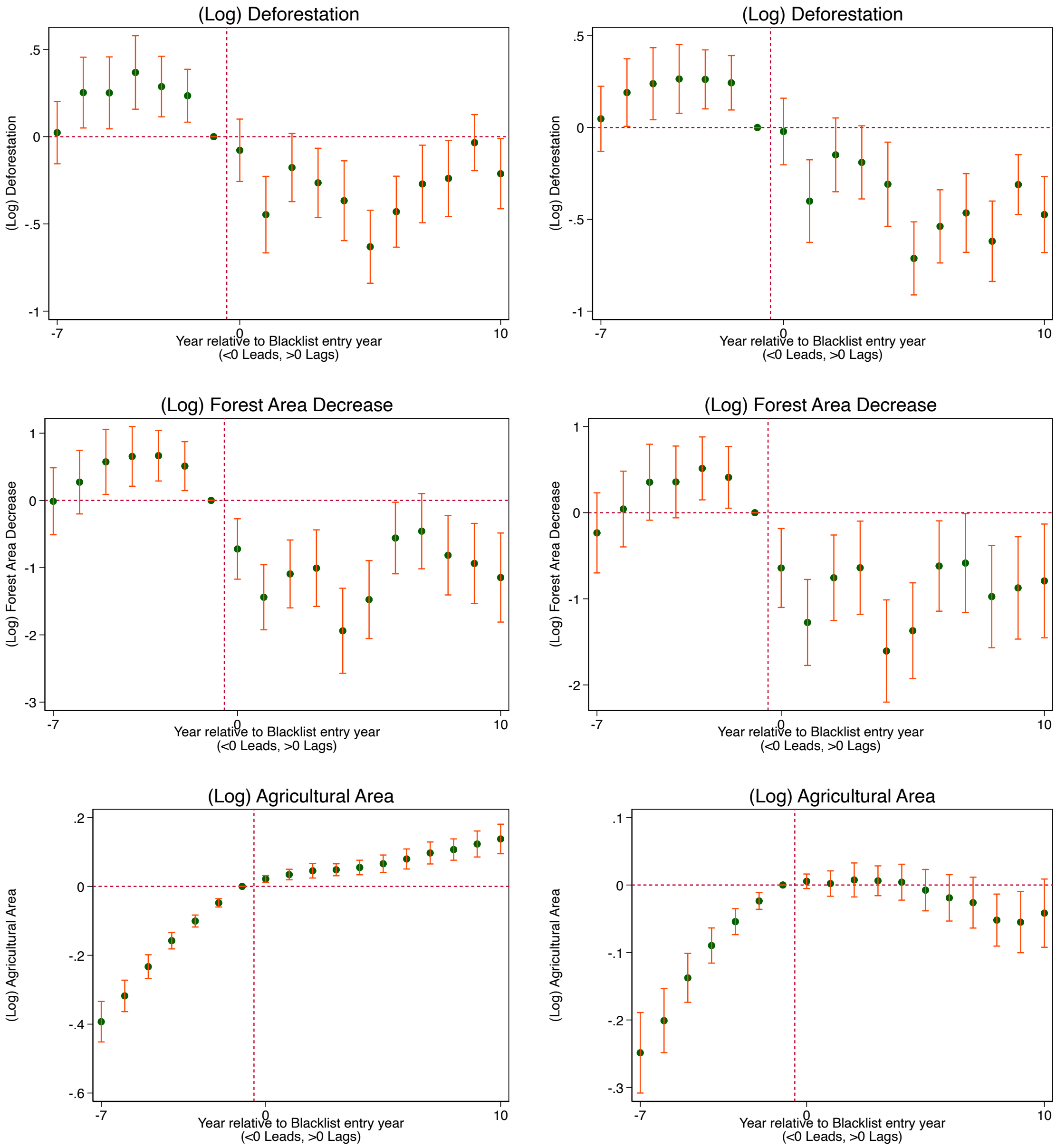
2.3.1 Priority List

Figure 5 below shows a visual representation of the event study estimates. Panels (a) and (b) show the effect on deforestation, panels (c) and (d) show the effect on decrease in total net forest area decrease (including naturally regrown forest), and panels (e) and (f) show the effect on total agricultural area. The left panels (a), (c) and (e) consider the entirety of Brazil, whereas the right panels (b), (d) and (f) consider only the Brazilian Legal Amazon.

A few clear patterns emerge: firstly, the β_τ coefficients for $\tau > 0$ are negative and generally significant for deforestation and net forest area decrease, suggesting that the policy had an effect in reducing deforestation locally, in line with the findings from Assunção and Rocha (2019). Secondly, the β_τ coefficients for $\tau < 0$ are positive and generally significant for deforestation for deforestation and net forest area decrease, suggesting the existence of pre-trends: before being added to the priority list, targeted municipalities had a stronger evolution in deforestation than non-targeted. Thirdly, there is a very clear upwards trend in agricultural area before the implementation of the policy, suggesting that targeted areas are those which were increasingly converting forest area into agricultural land prior to being included in the priority list.

The existence of pre-trends violates the validity assumption of the event-study design, which means that our estimates for the policy effect are likely biased. To deal with that, we conduct a synthetic differences-in-differences estimation. The results are qualitatively in line with a visual inspection of the event studies: the point estimates (standard errors) are 1.09% (0.11) for the policy impact on annual deforestation and -0.11% (0.03) on agricultural

Figure 5: Event-study estimate of the impact of the Priority List policy on various outcomes

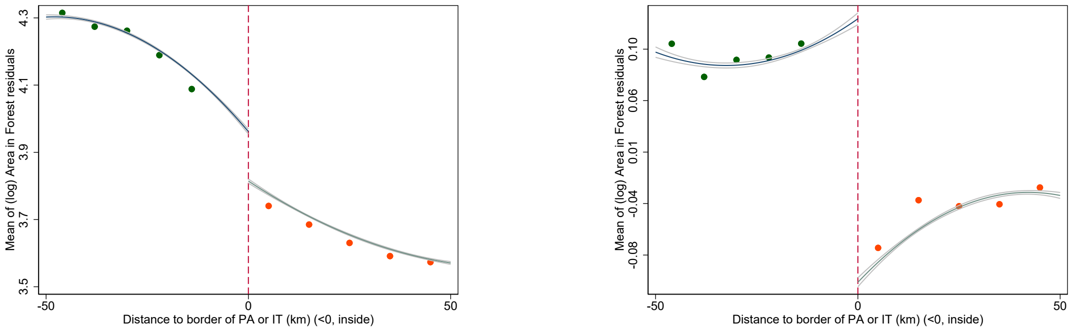


Note: This figure shows the β_τ estimates from the event-study specified in Equation 1 on three outcomes of interest: deforestation (top row), forest area decrease (middle row - defined as the difference between deforestation and forest regrowth), and total area in agriculture (bottom row). The three panels on the left consider all 5,568 munipalities in Brazil, whereas the three panels on the right consider only the municipalities from the Brazilian Legal Amazon, defined by the government as comprising all seven states from the Northern region of Brazil, plus the states of Maranhao and Mato Grosso.

area in the ten years following the policy implementation.

2.3.2 Protected Areas

Figure 6: Discontinuity in forested area around borders of protected areas.



(a) No controls

(b) Municipality and Year FE's

Note: This figure illustrates the discontinuity in forested area around the borders of protected areas. Each dot represents the mean total log-forested area within bins of 10km around the border of all protected areas, from 50km inside of 50km outside. The x-axis shows distance to the border, with negative (positive) values indicating the area inside (outside) the protected area. The left-hand side panel illustrates the simple raw averages. The right-hand panel illustrates averages of the residuals of a regression of log-forested area on municipality and year fixed effects.

Figure 6 illustrates our regression discontinuity approach by showing average forest area in bins of 10km around the border of a protected area. Panel (a) on the left shows the simple RD specification with a quadratic fit and no fixed-effects, whereas panel (b) on the right shows the residualised means on municipality and year fixed-effects. Both specification show a clear jump in forested area of around 0.2 log-points at the demarcated border, suggesting that the establishment of protected areas does succeed in locally protecting forest cover.

Table 1 show the estimated β from Equation 3 when adding different sets of fixed-effects. The resulting estimates show an increase of between 0.14 and 0.24 log-points in the forest coverage at the border of a protected area, in line with Figure 6.

In order to test the validity of our regression discontinuity design, we investigate the plausibility of the smoothness assumption by comparing what happens at the border a future protected area before it is established. Figure 7 illustrates this comparison. Panels (a) and (b) show, respectively, average forest cover around the border of a protected area before and after its implementation. Panels (c) and (d) show the exact same comparison, but now including year and municipality fixed-effects. Table 2 below shows the corresponding estimates for the RD parameter β.

The raw comparison from Figure 7 panel (a), as well as estimates from Table 2 columns (1) and (2) shows that, before the implementation of a protected area, the area with for-

	(Outcome: log-forested area)		
	(1)	(2)	(3)
Estimated Gap (PA or IT)	0.1430***	0.2396***	0.2378***
	[0.0425]	[0.0343]	[0.0356]
Quad spline	Yes	Yes	Yes
Year FE	Yes	No	Yes
Mun FE	No	Yes	Yes
Year X mun. FE	No	No	Yes
R2	0.0691	0.6081	0.6259
Observations	2.16e+06	2.16e+06	2.16e+06

Note: *p<0.1, **p<0.05, *** p<0.01.

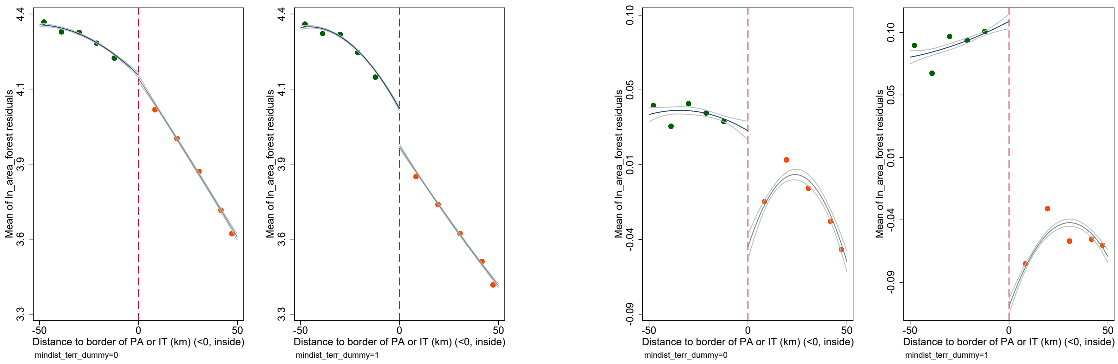
Table 1: Regression discontinuity estimates for the effect of Protected Areas on forested area

	5+ years before		10+ years before	
	(1)	(2)	(3)	(4)
Estimated Placebo Gap	0.0387	0.0678	0.2509***	0.0777**
	[0.0620]	[0.0450]	[0.0583]	[0.0336]
Quad. spline	Quad. spline	Quad. spline	Quad. spline	Quad. spline
Year FE	Yes	Yes	Yes	Yes
Year X mun. FE	No	Yes	No	Yes
Years before	10+	10+	5+	5+
R2	0.0801	0.6756	0.0513	0.6742
Observations	3.22e+05	3.22e+05	2.82e+05	2.82e+05

Note: *p<0.1, **p<0.05, *** p<0.01.

Table 2: Placebo Regression Discontinuity estimates around the border of Protected Areas before their establishment

Figure 7: Discontinuities in forested area around borders of Protected Areas, before and after the introduction of the policy



(a) No controls

(b) Municipality and Year FE's

Note: This figure compares the discontinuity in forested area around the borders of protected areas, before and after the implementation of the policy. Each dot represents the mean total log-forested area within bins of 10km around the border of all protected areas, from 50km inside of 50km outside. The x-axis shows distance to the border, with negative (positive) values indicating the area inside (outside) the protected area. The two left-hand side panels illustrates the simple raw averages, pre- and post-policy. The right-hand panel illustrates averages of the residuals of a regression of log-forested area on municipality and year fixed effects, pre- and post-policy.

est coverage evolved smoothly around the border of the future conservation area. When controlling for year and municipality fixed effects, we see evidence of some prior discontinuity (Figure 7 panel (c)), but significantly smaller than the discontinuity after the policy implementation.

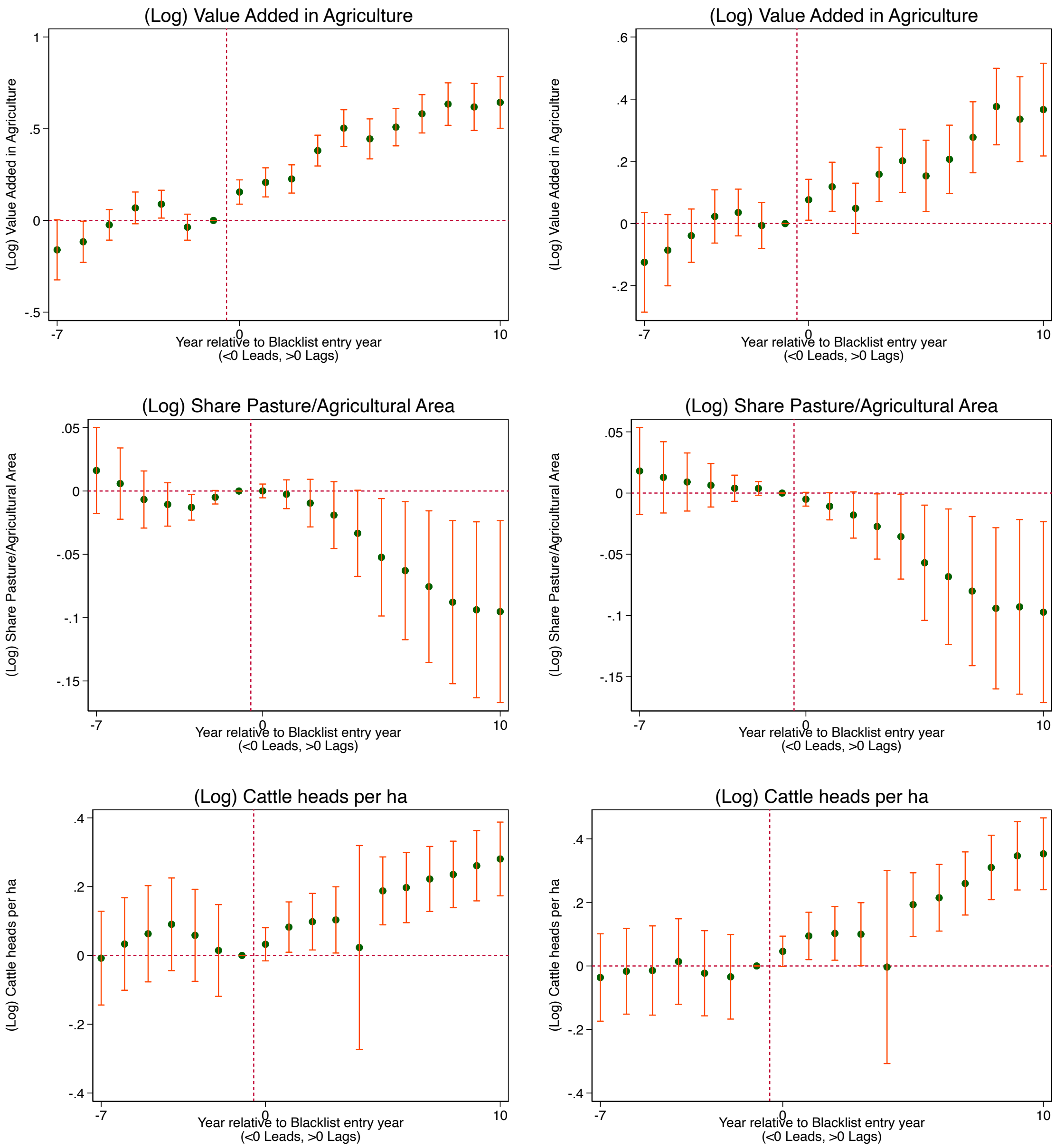
A discontinuity around the border of a future protected area before its implementation can be due to a number of factors, for example: (i) lengthy legal disputes over the demarcation of new protected areas, during which deforestation decreased before the final settlement and establishment [7] or (ii) historical occupation of the area by indigenous people whose livelihood depend on the preservation of the forest before legal demarcation of the protected territory.

Effects on other outcomes Surprisingly, municipalities that get added to the priority list see an increase in their agricultural value added as measured in (cite dataset of Contas Regionais). This seems to be primarily due to three phenomena: (1) A shift away from pasture towards crops (especially soy and maize, and (2) an increase in the number of cattle

[7] This is particularly likely in the case of Indigenous Territories. Their establishment follows a judicial process that includes several stages: study, delimitation, declaration, homologation, and regularization, that can take decades

heads per area in pasture.

Figure 8: Effects on agricultural profits and crops vs. pasture land use



Note: This figure shows the β_τ estimates from the event-study specified in Equation 1 on three outcomes of interest: log-total value added in agriculture (top row), log of the share of total agricultural area dedicated to pasture (middle row), and log of total cattle heads per hectare (bottom row). The three panels on the left consider all 5,568 munipalities in Brazil, whereas the three panels on the right consider only the municipalities from the Brazilian Legal Amazon, defined by the government as comprising all seven states from the Northern region of Brazil, plus the states of Maranhao and Mato Grosso.

3 Deforestation in Spatial Equilibrium

We build a spatial multi-sector general equilibrium model that explicitly captures the key features of the spatial distribution of agriculture, the economic forces at play in the market for land, and frictions to movement of goods and labour via internal trade and migration.

Stylised facts Our theoretical framework is motivated by observed spatial patterns in deforestation and agriculture across Brazil, as illustrated by Figures 9 and 10. From Figure 9, we can see that areas with more deforestation have lower GDP per capita[8], and lower market access[9]. From Figure 10, we can see that areas with higher deforestation have lower agriculture productivity[10] and higher share of agricultural area dedicated to pasture, which is typically much more land intensive than crop growing.

Overall, deforestation seems to happen in poorer and more remote areas, where land is dedicated to activities that are more land-intensive and less productive. There are differences in how costly it is to deforest land, as well as stark differences in agricultural productivities[11]. Given the large difference in local economic development, Brazil's continental dimensions and large heterogeneities in transportation infra-structure, goods and workers face non-negligible mobility costs. Moreover, the demand for agricultural land is not perfectly inelastic as farmers can adjust land intensity and consumers can choose between different goods depending on their relative prices.

Mechanisms for leakage In our framework, a local anti-deforestation policy is interpreted as an exogenous shock to the local supply of agicultural land. To quantify spatial leakage, we consider two key mechanism through which a negative supply shock generates increased deforestation elsewhere. The first mechanism operates through the market for agricultural goods. A negative shock to the local supply of agricultural land decreases the supply of agricultural goods. The extent to which this avoided deforestation leaks elsewhere depends on the extent to which goods are substitutable across space and can be produced by clearing forest elsewhere. The second mechanism operates through the market for labour. A local reduction in deforestation coming from a shock decreases the demand for agricul-

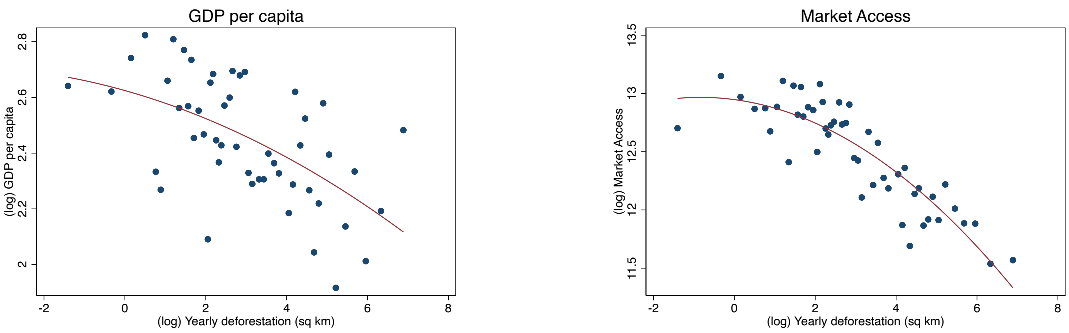
[8]Deforestation in Brazil is strikingly concentrated in space, as illustrated in Figure 1. Out of the 558 micro-regions that make up Brazil, the top 20 account for almost 40% of the total yearly deforestation. Regions with high levels of deforestation are very different ecologically and economically to the rest of Brazil.

[9]Defined by the first order approximation in (Donaldson and Hornbeck, 2016) $MA_d = \sum_o (\tau_d^o)^{\sigma-1} L_o$ where τ_d^o is the ice-berg trade cost between o and d as estimated in section 6.5

[10]Defined as agriculture revenue per hectare

[11]This is true for the agricultural sector as a whole, as well as for spatial differences in cultivation of specific crops. It is also true both when measuring productivity as revenue per hectare and revenue per worker.

Figure 9: Correlates of average yearly deforestation rate

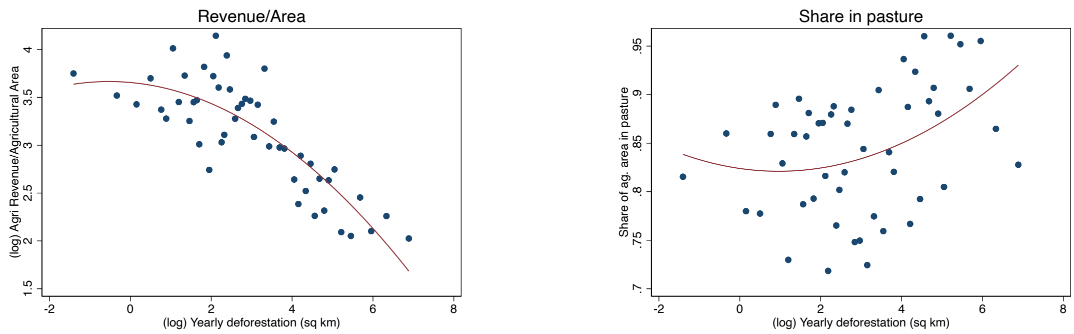


(a) log GDP per capita

(b) Market access

Note: This figure illustrates the correlation between average yearly deforestation rate and economic characteristics: log-GDP per capita (left panel) and market access (rigth panel - defined as). The dots represent the mean of the y-axis variable for 50 bins of log-yearly deforestation.

Figure 10: Correlates of average yearly deforestation rate



(a) Agricultural productivity

(b) Land intensity

Note: This figure illustrates the correlation between average yearly deforestation rate and economic characteristics: agricultural productivity (left panel - defined as the ration between agricultural revenue and agricultural area), and share of agricultural land dedicated to pasture (rigth panel). The dots represent the mean of the y-axis variable for 50 bins of log-yearly deforestation.

tural labour, changing workers' migration incentives. Inflows of workers to regions without anti-deforestation policies will increase the demand for agricultural land, raising incentives to deforest[12].

[12] A useful benchmark is a flat economy where all land is made equal, workers and goods can move freely, and demand for agricultural land is perfectly inelastic. In this case, leakage would be 100%: banning deforestation entirely in one region would have no global effect, as it would be perfectly leakeg to other regions - the demand for agricultural goods would be the sole determinant of the amount of agricultural land, and hence of the level of deforestation. Our model departs from all of those assumptions in ways that are consistent with the data.

To capture these channels, our model considers deforestation as an intermediate economic sector that supplies land as a factor of production for agriculture. Regions differ in the sectoral productivities for agricultural and non-agricultural production, as well as in their productivity in producing land via deforestation[13].

Model features We conceptualise deforestation as an investment in agricultural land that accumulates over time. Brazil is modelled as a closed economy with domestic trade and migration[14]. The model considers Brazil's 558 microrregions[15] indexed by r, which differ on their sectoral productivities, land endowments, and amenities. The economy is composed of $K+1$ sectors: K agricultural commodities that use land and labour as inputs and have different labour shares, and one non-agricultural sector that uses only labour. Additionally, we consider deforestation as a sector that uses a composite investment good in order to produce agricultural land for the K agricultural sectors. There is trade between municipalities subject to iceberg costs. Consumer preferences are non-homothetic, represented by the Price-Independent Generalized Linear preference formulation (Boppart, 2014). Final goods in each sector are a composite of regional varieties aggregated with constant elasticity of substitution σ and a final agricultural good is a composite of the various agricultural commodities aggregated with constant elasticity of substitution θ.

The model features a sequence of static spatial equilibria linked by the laws of motion of land and labour. The law of motion of land is determined by deforestation and the law of motion of labour is determined by migration and population growth.

Table 3 summarises the estimation of the various structural parameters, described in detail next.

3.1 Deforestation and the accumulation of agricultural land

Initially, each region r is endowed with L_{r0} workers, T^A_{r0} units of agricultural land area, and T^N_{r0} units of terrestrial natural ecosystems which can be converted into agricultural land.

For most of this section, we omit the time subscript and treat the equilibrium as static. However, since agricultural land accumulates, it is a time-varying quantity. Assuming a fixed forest regrowth rate ρ, at letting T^D_{rt} be the level of deforestation, agricultural land evolves

[13] Deforestation productivity can be thought of as region-specific factors that govern how suitable a particular region is for forest cutting. It can include natural factors (forest density, type of vegetation, altitute, weather patterns, geographical features), infra-structure and accessibility, the political environment and the level of enforcement of anti-deforestation laws.)

[14] See the appendix for a discussion of how international trade might change our results

[15] This is typically considered the aggregation level that is most closely associated with a local market, see for example Dix-Carneiro and Kovak (2017)

Parameter		Value	Source/Method
Preference Parameters			
β	Discount rate	0.9	
ϕ	Ag. share in price index	0.1	Eckert and Peters (2022)
ν	PIGL Preference parameter	0.5	Eckert and Peters (2022)
η	Non-homotheticity	0.506	Expenditure Survey Data (2017/18)
Elasticities of substitution			
σ	Between origins	9	Eckert and Peters (2022)
θ	Between ag. goods	2	Costinot et al. (2016), Dominguez-Iino (2023)
Production Function			
$\{\alpha_k\}_k$	Land share in ag. activities	(0.36;0.54;0.71)	From 2006 ag. census
χ_A, χ_{NA}, ι	Individual sectoral productivity shocks distribution	(2;1.6;12.8)	Alvarez (2020)
δ	Deforestation returns to scale	0.5	Two-way fixed effect of deforestation on land prices
ψ	Elasticity of deforestation TFP to natural area left	0.32	Regression derived from steady-state deforestation
ρ	Forest regrowth rate	0.003	Observed reforestation rates
Trade and Migration parameters			
$\{\mu_d^o\}_{o,d}$	Bilateral migration utilities	residuals	Migration flows 2005-2010
ϵ	Dispersion of idiosyncratic tastes	14.08	Migration flows and incomes from 2010 census
κ	Trade costs distance elasticity	0.11	State trade flows in 1999 and 2017

Table 3: Summary of estimation of model parameters

according to

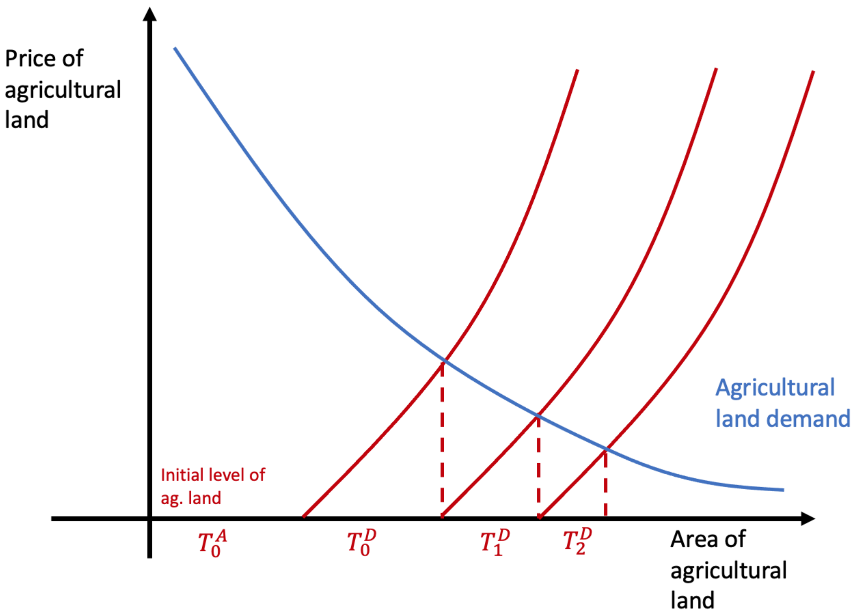
$$T^A_{rt+1} = T^A_{rt}(1-\rho) + T^D_{rt}.$$

The following simple agricultural land market graph helps visualise how land evolves over time. In this simple graph land demand is fixed, there is only one region, and the is no forest regrowth[16].

The market for deforestation. Deforestation is modelled as a costly investment in the production of agricultural land, which is a factor of production for the agricultural sector.

[16]Figure 16 in the appendix illustrate the case in with there is forest regrowth and the short-run and long-run land market equilibrium with a forest regrowth ρ and the transition dynamics.

Figure 11: The market for agricultural land over time



Note: This graph illustrates supply and demand curves for agricultural land. The blue curve represents a demand curve coming from the agricultural sector. The red curves represent the supply coming from the deforestation sector in consecutive time periods. Under the assumption of no regeneration, the equilibrium quantitiy of one period is the horizontla intercept of the supply curve in the next period.

We can think of deforesters as atomistic agents operating in a perfectly competitive market where they access forested land, pay the fixed cost of clearing it, and sell it at the price of agricultural land. We model the aggregate deforestation production function so that it has decreasing returns to scale. This motivated by the fact that, within each time period, the forest that is closer to the edge is cheaper to access and clear. This means that each additional dollar spent deforesting is less productive. The aggregate deforestation production function is such that if an amount I_r^D of the final good, bought at price p_r, is invested in deforestation, it delivers T_r^D units of agricultural land according to

$$T_r^D = Z_r^D (I_r^D)^\delta, \tag{4}$$

where Z_r^D is the region-specific "deforestation productivity", and $\delta \in (0,1)$ governs the returns to scale of the production function. The returns of each square kilometre of deforested land equal the value of agricultural land in a given region, q_r. We can interpret Z_r^D and δ as the productivity parameters that dictate the aggregate (convex) costs of deforesting, $C_r^D(T_r^D, p_r; Z_r^D, \delta) = p_r (Z_r^D)^{-\frac{1}{\delta}} (T_r^D)^{\frac{1}{\delta}}$.

Our model of the market for deforestation has two defining features. First, land in natural ecosystems is open access, which means that the future value of the forest is fully discounted[17]. Second, there is free entry of deforesters. Accordingly, they enter the deforestation market as long as there are non-negative marginal profits, so that the equilibrium level of deforestation is when marginal costs, $MC_r^D = \partial C_r^D(T_r^D, p_r; Z_r^D, \delta)/\partial T_r^D$, equal marginal revenues, $MR_r^D = q_r$,

$$T_r^{D*} = \underbrace{(Z_r^D)^{\frac{1}{1-\delta}}}_{\text{Local Factors}} \underbrace{\left(\frac{\delta q_r}{p_r}\right)^{\frac{\delta}{1-\delta}}}_{\text{Equilibrium Effects}} . \tag{5}$$

Supply side of deforestation. The curve $T_r^{D*}\big(q_r/p_r\big)$ is the region-specific deforestation supply curve: how deforestation responds to the local relative price of land. It depends on two parameters δ, the returns to scale of deforestation in one year, and Z_r^D, the location-specific deforestation productivity.

The returns to scale of deforestation δ is taken to be the same for all of Brazil, and it governs the supply elasticity of deforestation to real land prices, which equals $\frac{\delta}{1-\delta}$. That is, for a 1% increase in the value of agricultural land relative to that region's price index, deforestation goes up by $\frac{\delta}{1-\delta}$%. See section 6.2 in the appendix for the estimation of δ.

Z_r^D is the source of regional heterogeneity in deforestation costs. It reflects differences in characteristics such as, (i) local environmental conditions that influence how difficult it is to deforest, such as rainfall and temperature, (ii) the level of enforcement of anti-deforestation policies, (iii) revenues obtained from the act of deforestation itself, for example through the sale of wood, and (iv) the option value of keeping land as forest. While most of these things are difficult to observe and quantify, we correlate them to the calibrated Z_r^D's and find that the current area of unprotected forest in a region, T_r^F is a strong predictor of Z_r^D. A log-log model does a remarkably good job at describing their empirical relationship in cross-sectional data. Therefore, we let

$$Z_r^D = \overline{Z_r^D}(T_r^F)^{\psi}.$$

It is calibrated so that it perfectly explains observed differences in levels of deforestation in regions that cannot be explained by agricultural rents and market access, which are reflected by q_r and p_r respectively. In turn, prices are calibrated by inverting a spatial equilibrium model in each time period, as described in 3.6. Intuitively, we estimate the demand for agricultural land from farmers, which in turn depends on the demand for agricultural

[17]Around 50% of deforestation in the Amazon over the past few years has happened in untitled public lands. From the remaining 50%, about half has happened in rural settlements where land was and half in private properties. Although the amount of deforestation in private properties is not negligible, about 25% of the total, rights over forested land are insecure even within private property. This is due to it historically being often regarded as "unproductive" and thus subject to the ownership claims of squatters. See (Alston et al., 1999) for an in depth description of property rights in the Brazilian Amazon frontier.

products from consumers, and then, using the observed data on deforestation, we calculate the productivity of deforestation as a residual that rationalises their spatial distribution. To be more concrete, for a given level of observed deforestation, a region with lower market access and lower agricultural productivity will have a higher calibrated Z_r^D.

3.2 Technology: local demand for workers and land

In order to estimate the demand for land in each region, we first need to impose some structure on the firms operating in each region. This will determine how they demand different factors of production, land and workers, given productivities and prices. Productivities will be treated as exogenous fundamentals to be backed out from the model and prices will depend on the full (static) equilibrium which takes into account consumer preferences and trade costs.

There are four broad sectors: three agricultural sectors with varying land intensities, and non-agriculture which uses no land for production. Regions produce a differentiated variety of each of these four goods as in Armington (1969) that consumers combine with Constant Elasticity of Substitution as in Anderson (1979). The market in each of the four sectors is composed of perfectly competitive firms with constant returns to scale. The local non-agricultural goods are a product of only labour with regional productivity Z_r^{NA}, so that $Y_r^{NA} = Z_r^{NA} L_r^{NA}$. The agricultural goods, indexed by k, are a Cobb-Douglas function of land and labour with constant returns to scale and regional productivities Z_r^{Ak}, and a share of land equal to α_k, so that

$$Y_r^{Ak} = Z_r^{Ak} (L_r^{Ak})^{1-\alpha_k} (T_r^{Ak})^{\alpha_k}.$$

Goods of sector s are produced to be sold at origin prices p^{so}.

In a competitive equilibrium, the rental rate of agricultural land v_r equals the marginal product of agricultural land and the wages in each sector equal the marginal product of labour in each sector. Assuming simple adaptive expectations (i.e. agents assume the future rental rate of land equals today's) and a discount rate of β, land should be priced at its expected present value,

$$q_r = \frac{1}{1-\beta(1-\rho)} v_r.$$

Our data allows us to get agricultural revenues (see section 3.6) and use them to get v_r as

$$v_r = \sum_k \alpha_k \frac{p^{Akr} Y_r^{Ak}}{T_r^{Ak}}.$$

Wage gaps and occupational choice. In order to allow for a gap between agricultural and non-agricultural wages, as we consistently find in the data, we rely on a model like the one introduced by Lagakos and Waugh (2013) and applied to the context of Brazil in Alvarez (2020). In this model, individuals draw idiosyncratic productivities for agriculture and non agriculture from a joint distribution $F(z_i^A, z_i^{NA})$ and given their observed productivities for each sector[18]. Workers choose sector in order to maximise their wage income, so that they work in non-agriculture if and only if $z_{ri}^{NA} w_r^{NA} \geq z_{ir}^{A} w_r^{A}$. Firms set wages per efficiency unit (w_r^A, w_r^{NA}) that equal the marginal product of a worker with unit productivity in that sector. By L_r^s we refer to the total labour efficiency units in sector $s \in \{A, NA\}$ in region r, which is equal to the number of workers multiplied by the expected productivity of those who choose to work in sector s. Since the idiosyncratic productivities in each sector are not independent draws, there will be income gaps, the average wage income $(\overline{y}_r^{LA}, \overline{y}_r^{LNA})$ will not be equalised across sectors.

Estimating α_k. To estimate α_k for each agricultural commodity k, we use data from the 2006 agricultural census on area planted in each type of agriucltural activity T_r^{Ak}, workers in type of agricultural activity N_r^{Ak}, labour incomes per capita in agriculture $\overline{y}_r^{LA}$, and land value q_r.

The Cobb-Douglas functional form of the agricultural production function and the zero-profit condition together imply that

$$\frac{\alpha_k}{1-\alpha_k} = \frac{\sum_{r=1}^{R} v_r T_r^{Ak}}{\sum_{r=1}^{R} \overline{y}_r^{LA} N_r^{Ak}}.$$

Doing this we find that the land share for pastures, temporary crops, and perennials vary significantly and they are, respectively: 0.71, 0.54, and 0.36. These differences make it so that there is greater substitutability between land and labour in agricultural production as a whole. This has important implications for leakage. Consider a region where the supply of agricultural workers increases. This could be due to more migration from regions where conservation policies have been enacted. Higher supply of agricultural workers increases the demand for agricultural land. In a model with multiple agricultural sectors, however, it might also shift land use away from very land intensive activities, for example cattle ranching, towards more labour intensive ones, such as perennial crops. The extent to which this happens in equilibrium would depend mainly on consumers willingness to substitute consumption between goods and on the slope of the supply of deforestation. If deforestation

[18] The joint distribution is taken to be, as in Alvarez (2020), a Frank copula or two Frechet distributions with shape parameter χ^A and χ^{NA} and correlation ι

is very cheap and consumers very reluctant to substitute towards less land-intensive goods, cattle-ranching will remain preferable and the inflow of workers will lead to much deforestation. By contrast, in a region where deforestation is very costly and consumers readily substitute beef with soy, the increased supply of labour will mean farmers may opt to convert pastures to soy fields and demand very little deforestation.

3.3 Preferences: demand for goods

Having characterised the factors that determine the demand for agricultural land given the prices of agricultural goods, we now turn our attention to the preferences that governs the demand side of agricultural goods markets.

Non-homothetic preferences between agricultural and non-agricultural goods. Consumers have PIGL preferences as in Boppart (2014) over agricultural and non agricultural goods such that the expenditure share in agricultural goods of a consumer with expenditure e equals

$$\vartheta^A(e, \vec{p}) = \phi + \nu \underbrace{\left(\frac{p^A}{p^{NA}}\right)}_{\text{relative price effect}} \underbrace{\left(\frac{e}{p}\right)^{-\eta}}_{\text{Income effect}}.$$

Where p^A and p^{NA} are prices of agricultural and non-agricultural sector, p is the price of the composite consumption good $p \equiv (p^A)^\phi (p^{NA})^{1-\phi}$, e is total expenditure, and η, γ, and ν are exogenous parameters.

CES aggregation between goods. The agricultural good is, in turn, a CES aggregate of agricultural goods $k \in \{\text{beef, temporary crops, permanent crops}\}$ with elasticity of substitution θ, so that the agricultural price index p^A is equal to $p^A \equiv \left(\sum_{k=1}^{K}(p^{Ak})^{1-\theta}\right)^{\frac{1}{1-\theta}}$ and the share of agricultural good k in the overall agricultural expenditure is given by $\left(\frac{p^{Ak}}{p^A}\right)^{1-\theta}$.

CES aggregation between origins. Each of the $3+1$ final goods (the three agricultural goods and the non-agricultural good) is, in turn, a CES composite of differentiated regional varieties produced in region r, with a constant elasticity of substitution σ. Trade is taken to have symmetric iceberg trade costs $\tau_d^o \geq 1$ that do not vary by good so that a consumer from region d pays the origin price of a good from region o scaled by the bilateral iceberg trade cost. Thus, the share of goods g (which can be beef, temporary crops, perennials, or non-agricultural goods) consumed in region d, that is produced in region o, is given by[19]

$$\pi_d^{go} = \frac{(\tau_d^o)^{1-\sigma}(p^{gr})^{1-\sigma}}{\sum_{r=1}^R (\tau_d^r)^{1-\sigma}(p^{gr})^{1-\sigma}}. \tag{6}$$

We take σ from Eckert and Peters (2022) to be 9.[20] Having a value for σ, we use inter-state trade flow data from 1999 and 2017 to estimate a gravity equation derived from 6 along with the assumption that the iceberg trade costs depend on distance according to $\tau_d^o = (\text{distance}+1)^\kappa$. See section 6.5 in the appendix to see this estimation.

3.4 Market clearing: closing the model

In order to close the model, we equalise the revenues from producers of good g from origin o to the sum of the expenditures of consumers in that good from across all regions d of Brazil. These expenditures equal the trade share of d to o times the total consumption of region d in good g,

$$p^{go}Y_o^g = \sum_{d=1}^R \pi_d^{go} X_d^g. \tag{7}$$

Aggregate consumption in a region. The overall expenditure of a region in each good can be calculated in three steps. First, we estimate the aggregate share of consumer expenditure in agriculture ϑ_r^A, which will depend on the PIGL preference parameters, the relative price of agricultural goods, the aggregate expenditure, and the distribution of incomes as dictated by wages and the joint distribution of productivity shocks[21]. Then, given the fact that the investment good is a Cobb-Douglas aggregate of the agricultural and non-agricultural goods with share ϕ, the total agricultural and non-agricultural expenditures equal

$$X_r^A = \vartheta_r^A E_r + \phi p_r I_r^D, \quad X_r^{NA} = (1-\vartheta_r^A)E_r + (1-\phi)p_r I_r^D \tag{8}$$

[19] To differentiate between origin and destination prices we shall indicate the region of origin with a superscript and the region of destination with a subscript. Then the farm-gate price of temporary crops from o is $p^{A,temp,o}$, the price of temporary crops from o faced by consumers in d is $p_d^{A,temp,o} = p^{A,temp,o}\tau_d^o$, and the consumer price index of temporary crops in destination d equals $p_d^{A,temp} = \sum_{o=1}^R (\tau_d^o)^{1-\sigma}(p^{A,temp,o})^{1-\sigma}$.

[20] Given that our context is closer to Domiguez-Iino (2021), we will also calibrate the model and simulate counterfactuals with values up to $\sigma = 15$, given his estimated elasticity of substitution between "counties" in Brazil and Argentina of around 13.

[21] Given the non-homothetic nature of preferences for agricultural goods and the fact that there is within-region inequality due to the productivity shocks introduced to generate a wage gap, the formula for the region-wide agricultural share is not very concise and can be found in the appendix in section 6.1. Intuitively, in the face of higher inequality, the share of expenditure in agriculture will be lower because a higher share of earnings will accrue to those who spend proportionally less in agriculture.

where E_r is the total consumer expenditure in a region and I_r^D the total investment in deforestation. Because of the CES preferences between agricultural goods, the expenditure in agricultural good k equals $X_r^{Ak} = X_r^A \left(\frac{p_r^{Ak}}{p_r^A}\right)^{1-\theta}$. Finally, the total consumer expenditure equals the total income, labour income plus land rental rate payments, net of deforestation investments,

$$E_r = \overline{y}_r^L L_r + v_r T_r^A - p_r I_r^D.$$

where $\overline{y}_r^L$ is the average wage income in region r across all worker, equal to $s_r^A \overline{y}_r^{LA} + (1 - s_r^A)\overline{y}_r^{LNA}$.

Equilibrium definition. Consider the economy described above. Let the initial agricultural land area in each region $\{T_{r0}^A\}_r$, the initial level of area under terrestrial natural ecosystems $\{T_{r0}^N\}_r$, and the distribution of workers across space $\{N_r\}_r$ be given as exogenous. A competitive equilibrium is a set of prices $\{p_r^s\}_{r,s}$, wages $\{w_r^s\}_r$, land rental rates $\{v_r\}_r$, occupational choices $\{N_r^s\}_{r,s}$, regional deforestation levels $\{T_r^D\}_r$, and regional expenditure shares $\{\vartheta_r^A\}_r$ such that:

1. consumers' choices maximize utility (equation (12));
2. the demand for regional varieties follows equations (6);
3. firms' factor demands maximize firms' profits;
4. marginal product of labour is equalised across sectors
5. local markets clear and there is trade balance (7).

3.5 Migration

As in Eckert and Peters (2022), individuals born in a location of origin o can choose to live in destination d according to the migration costs μ_{od}, the utility at the destination $V(e_d, p_d)$, the famenities at destination d, B_d, and an idiosyncratic shock $\nu_d(i)$, drawn from a Frechet distribution with parameter ϵ. The origin-destination specific migration utility $\mathcal{U}_{od}$ is given by:

$$\mathcal{U}_{od}(i) = V(e_d, p_d) B_d \mu_{od} \nu_d(i).$$

Therefore, from the Frechet nature of the shock, the share of people who move from o to d is given by the following expression[22]

$$\rho_{od} = \frac{(V(e_d, p_d)\mu_{od}B_d)^{\epsilon}}{\sum_{r=1}^{R}(V(e_r, p_r)\mu_{or}B_r)^{\epsilon}}.$$

Hence the law of motion of population will be given by the following equations

$$N_{dt} = \sum_{o=1}^{R} \rho_{rd} N_{ot}, \quad N_{ot} = g^N_{ot-1} N_{ot-1}$$

where g^N_{ot-1} is the growth rate of population in origin region o at time t.

3.6 Data and model inversion

Data for the empirical analysis of our study comes from a variety of publicly available sources for Brazil, comprising information on land use, agricultural activity, local economies, internal trade, migration, and household expenditure. All the sources listed here are publicly available.

Model inversion. In order to invert the model and back out the total factor productivities of each sector and the deforestation productivity (Z^{Ak}_{rt}, Z^{NA}_{rt}, Z^{D}_{rt}) we assume that each year in the period 2003-2019 the economy is in equilibrium as described above and we use data on some observed features. First, we use data on the endowments that are treated as exogenous within each equilibrium. Namely, population, already agricultural land, and land in natural ecosystems. Second, we use data on the following equilibrium quantities in each region for that year: (i) the total amount of deforestation, T^D_{rt}, (ii) the agricultural labour share (s^A_{rt}), (iii) the share of agricultural land in each land use (T^{Ak}_{rt}), and (iv) agricultural revenues ($p^{Akr}_t Y^{Ak}_{rt}$).

Land use. Data on land use comes from the MapBiomas project, a multi-institution collaborative initiative that processes satellite images into publicly available datasets on land use for the Brazilian territory. The dataset we use contains pixel-level (30m resolution) of land use from Brazil between 1985 and 2020, which we aggregate to the micro-region level for the model calibration. It contains many land use categories which we aggregate into 5

[22]Note that this includes ρ_{oo}, i.e. the share of people from region o who choose to stay in region o. Typically, this will be a larger share than ρ_{od} for $o \neq d$

categories: natural ecosystems (comprised mostly of forests and to a lesser extent natural savannahs), pastures, temporary crops, perennials, and other (mostly urban, mining, water bodies, ice, and missing data).

Agricultural data. We need agricultural data for two main purposes. First, the estimation of some parameters of the model. Specifically α_k and δ. Second, for the construction of the micro-region level agricultural revenue data that is used to invert the model.

In the estimation of α_k, we rely on the 2006 Agricultural Census, a census of the universe of farming establishments. Using data on land prices, land in each agricultural activity, labour in each agricultural activity, and agricultural wages, we estimate how the ration of land income over rental income varies between pastures for cattle grazing, temporary crops, and perennials. For the estimation of δ, we rely on a fixed-effects log-log regression of deforestation rates on yearly municipality-level estimates of the land rental rates. Given the absence of panel on agricultural rental rates, we construct them in a way that is consistent with our model from our yearly estimates of agricultural revenues.

The estimation of the yearly agricultural revenues is done in two different ways. The Municipal Survey on Agricultural Production (PAM), provides yearly estimates of the revenues of 71 different crops, which we aggregate into two categories: temporary and perennial, in a way that is consistent with their classification in the agricultural census. These estimates yield very similar values to those in the agricultural census in 2006. Instead of relying on the PAM alone, we construct an region-level index of the PAM revenues that equals one in 2006 and multiply them by the 2006 Agricultural Census data.[23] Since there is no equivalent data on yearly revenues from cattle ranching at the municipality data, we leverage data on the main dimensions that could be making these revenues change from year to year: the number of cows, the weight of each cow, and the price of beef per kg. From the Municipal Survey on Livestock Production (PPM) we obtain yearly estimates of the number of cows in each municipality. From the XXXX we obtain data on national variation in the price of beef in each year. And lastly, from the XXX we get yearly state-level data on the average weight of cows being slaughtered. Given that the average age of a cow at the time of slaughter is 3 years, we use indices on the changes of these variables lagged by 3 years multiplied by the 2006 Agricultural Census data on revenues from cattle ranching in order to estimate yearly revenues.

[23] $\widehat{Reveue}_{rt}^{AK} = Revenue_{r2006}^{AK,AgCensus} \times \frac{Revenue_{rt}^{AK,PAM}}{Revenue_{r2006}^{AK,PAM}}$

Labour and Migration. Data on micro-region level labour markets comes from 10% samples of the 2000 and 2010 National Census, which is representative at the municipal level (smaller than the micro-region), made publicly available by IBGE. For each municipality, we compute the share of individuals who report working in the agricultural sector, as well as the average municipal earnings for agricultural and non-agricultural activities. We get the number of workers in agricultural and non agriculture for both of these years and then, for each micro-region, we linearly interpolate and extrapolate these values to approximate their time series from 2003 to 2019. Given the static nature of the equilibrium concept, it is not too important to accurately measure relative yearly changes in population and labour shares, but to get the levels right. Therefore, we think that the errors introduced by this approximation are unlikely to change our main results. We also use the 2010 census to obtain information on internal migration, as respondents are also asked to report in which municipality they lived 5 years prior to data collection.

Trade. Data on interstate trade flows comes from (de Vasconcelos, 2001), as cited by Morten and Oliveira (2018). We use trade flows in 1999 and 2017 and get similar estimates for the distance elasticity of trade flows, which is what we use to calculate iceberg costs.

Preferences. Finally, we use the 2017-2018 Household Expenditure Survey (POF) to estimate the preference parameter that governs the non-homotheticity in individual utility. We construct household expenditure shares on food and non-food items, as a proxy for consumption of agricultural goods.

3.7 Calibrated productivities

In this theoretical framework, a high level of deforestation is rationalised by either the demand for agricultural land or the supply of deforestation. The supply-side regional fundamental governing deforestation is Z_r^D. The demand for agricultural land is given, in turn, by the equilibrium in the market for agricultural goods. From the agricultural goods supply side, we expect to see more deforestation in regions which have higher agricultural TFPs, especially for the land-intensive agricultural activities (i.e. cattle grazing). From the demand side, the regions that have higher market access should experience greater demand for agricultural goods, and hence for agricultural land, as discussed in Donaldson and Hornbeck (2016). In summary, there are three sets regional fundamentals driving the spatial allocation of deforestation, each of which have different implications for the level of leakage: (1) the deforestation TFP, (2) the TFP of agricultural activities, especially the most land-intensive ones, and (3) market access. Table 4 presentes the estimated coefficients of a simple regression of the baseline levels of (log) deforestation on the (log) productivities of various sectors and the log of a measure of market access approximated by $MA_r \approx \sum_d (\tau_d^r)^{\sigma-1} L_d$ as in Donaldson and Hornbeck (2016). Column 1 shows that deforestation is well explained ($R^2 = 0.923$) by the combination of those factors but that the largest share of the explanatory power comes from the TFP of deforestation, the immobile factor that we are shocking through the counterfactual policy simulation.

	Outcome: log-observed annual deforestation			
	(1)	(2)	(3)	(4)
(log) temp. crops TFP	-0.54***	-0.21***		0.10***
	[0.04]	[0.03]		[0.00]
(log) perm. crops TFP	-0.39***	0.03		-0.00
	[0.03]	[0.03]		[0.00]
(log) pasture TFP	-0.16***	-0.61***		1.14***
	[0.05]	[0.04]		[0.01]
(log) non-ag TFP	1.08***	0.67***		-0.38***
	[0.03]	[0.03]		[0.00]
(log) Market acces		-1.19***		0.26***
		[0.02]		[0.00]
(log) Deforestation TFP			1.39***	1.90***
			[0.01]	[0.00]
R^2	0.122	0.306	0.811	0.987
Dep. Var. Mean	2.69	2.69	2.69	2.69
Dep. Var. SD	1.95	1.95	1.95	1.95
Observations	8587	8587	8587	8587

Table 4: Correlates of observed deforestation

4 Counterfactual Analysis

The theoretical framework developed in the previous section allows us to simulate counterfactual scenarios of local deforestation policy, and understand how deforestation would have evolved in general equilibrium in these alternative scenarios. In particular, we can apply our model to analyse the national impact of the two types of local anti-deforestation policies implemented by the Brazilian government over the past decades (Prority List and Protected Areas), taking into account general equilibrium effects driven by leakage.

For each of the two policies, we simulate two counterfactual scenarios for the evolution of nation-wide deforestation between the years 2003 and 2018.

Counterfactual A: No policy – a scenario in which, between 2003 and 2018, the policy was never enacted.

Counterfactual B: Policy, no leakage – a scenario in which, between 2003 and 2018, the policy was implemented, but leakage to other areas is not possible

The difference between counterfactual A and the observed evolution of deforestation measures the overall effect of the policy, including both the direct effect via banned deforestation in targeted areas, and the indirect effect via potential leakage to non-targeted areas. The difference between the observed evolution of deforestation and counterfactual B measures the amount of leakage caused by the localised policies. The difference between counterfactuals A and B measures the total effect of the localised policies under no deforestation leakage, i.e. only considering the direct effect via banned deforestation in targeted areas.

For each of the two policies, we construct the counterfactuals as follows: First, we calibrate the baseline regional fundamentals $\{Z^s_{rt},\, B_{rt}\}$ to fit the observed deforestation between the years of 2003 and 2018. We construct $(Z^D_{rt})^{nopolicy}$, the deforestation productivity of each region in a counterfactual scenario wihtout the policy was not implemented as:

$$(Z^D_{rt})^{nopolicy} = \overline{Z}^D_{rt}(T^F_{rt} + T^P T_{r,t})^{\psi} \tag{9}$$

Where T^F_{rt} is the remaining observed area of unprotected forest in region r at time t and $T^P T_{r,t}$ is the area belonging to region r that is observed to be under a local conservation policy at time t. In a counterfactual with no policy, this area is considered to be unprotected forest.

For counterfactual A, we simulate total land cleared for agriculture in each year between 2003 and 2018, solving for the optimal deforestation level $(T^D_{rt})^{nopolicy}$ according to Equation

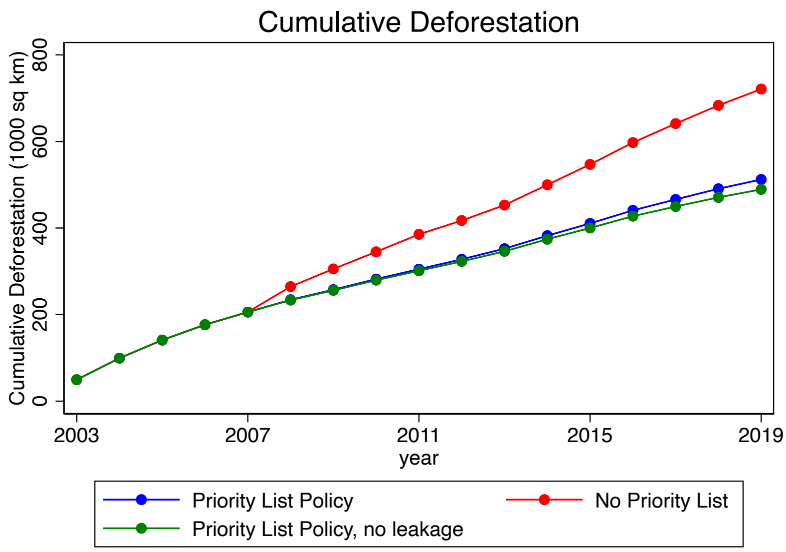
5 using the calculated regional deforestation productivity $(Z_{rt}^D)^{nopolicy}$. Total agricultural land at each period is then given by equation 10.

$$(T_{rt}^A)^{nopolicy} = (1-\rho)(T_{rt-1}^A)^{nopolicy} + (T_{rt}^D)^{nopolicy} \tag{10}$$

For counterfactual B, we calculate a "no leakage" scenario that has the implied prices coming from counterfactual A (i.e. no-policy prices), and baseline region deforestation productivities calibrated from the model Z_r^D (with policy). Fixing the prices under the no-policy scenario effectively shuts down the channel through which deforestation could leak to other areas, as both labors market and goods market relocation mechanisms operate through price channel.

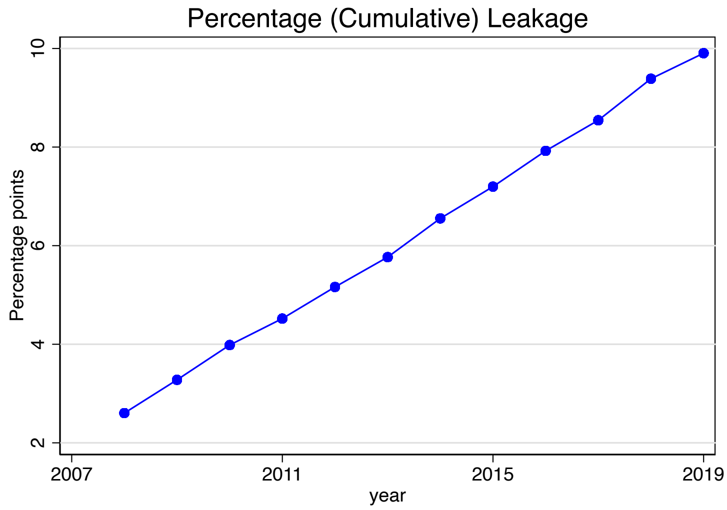
$$(T_r^D)^{noleakage} = \left(\delta \frac{(q_r)^{nopolicy}}{(p_r)^{nopolicy}}\right)^{\frac{\delta}{1-\delta}} (Z_r^D)^{\frac{1}{1-\delta}} \tag{11}$$

Figure 12: Counterfactual deforestation: Priority List policy



Note: This figure illustrates the evolution of total cumulative deforestation across the entire country in thousands of squared kilometers between the years of 2003 and 2019 under counterfactual simulations based on the Priority List policy. The blue curve represents observed deforestation, the red curve represents a counterfactual scenario in which the policy was never implemented, and the green curve represents a counterfactual in which the policy was implemented and spatial leakage is shut down by considering the prices under the no-policy scenario.

Figure 13: Cumulative leakage from Priority List policy

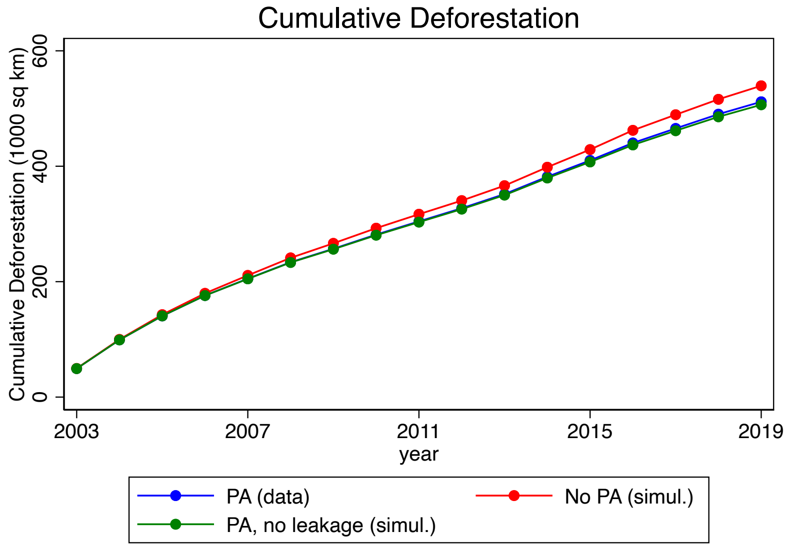


Note: This figure illustrates the evolution of total estimated deforestation leakage under the Priority List policy between the years of 2007 and 2019. Leakage as calculated as the ratio between: total increase in deforestation in non-targeted areas (differece between observed deforestation and the simulated deforestation in a counterfactual where prices do not adjust to policy), and total avoided deforestation in targeted areas (difference between simulated deforestation in a counterfactual where the policy was never implemented and the simulated deforestation in a counterfactual where the leakage channel is shut down). A leakage of 100% would mean that the entire deforestation reduction caused by the localised policies is outdone by increases in deforestation in other regions.

Figure 12 shows the evolution in total observed deforestation across Brazil, as well as the results from simulating counterfactuals A and B for the Priority List Policy. We can see that, in a counterfactual where the priority list polciy is never implemented, total deforestation between 2003 and 2009 would rise from around 450 thousand km^2 to around 700 thousand km^2. Figure 13 illustrates the evolution of the cumulative leakage to non-treated areas expressed as a share of the avoided deforestation. Leakage at the start of the considered period is low, below 3%, but increases gradually over time. By 2019, approximately 10% of the deforestation avoided by the policy in the targeted areas is outdone by spillovers elsewhere.

Figures 14 and 15 illustrate the same results for the Protected Areas policy. From Figure 14, we can see that, in counterfactual scenario without the Protected Areas, deforestation increases by much less than in counterfactual without the Priority List policy; from around 450 thousand km^2 to around 500 thousand km^2 between 2003 and 2019. From Figure 15, we can see that leakage in the Protected Areas policy follows a similar pattern to leakage under the Priority Lists policy: it starts at a low value of about 5%, and increases gradually over time to just over 15% at the end of our considered period.

Figure 14: Counterfactual deforestation: Protected Areas policy



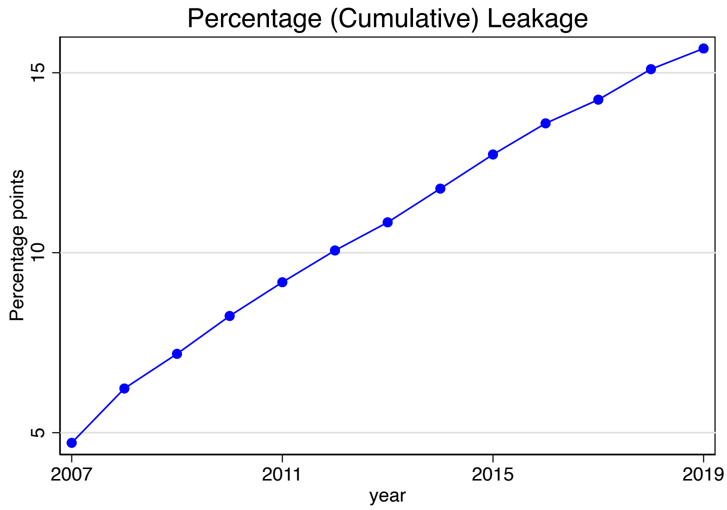
Note: This figure illustrates the evolution of total cumulative deforestation across the entire country in thousands of squared kilometers between the years of 2003 and 2019 under counterfactual simulations based on the Protected Areas policy. The blue curve represents observed deforestation, the red curve represents a counterfactual scenario in which the policy was never implemented, and the green curve represents a counterfactual in which the policy was implemented and spatial leakage is shut down by considering the prices under the no-policy scenario.

Overall, the conclusion from our simulation exercise is that, even though there is a detectable amount of leakage to non-targeted areas in localised anti-deforestation policies, it does not mean that such policies are ineffective in reducing global deforestation. Both localized policies implemented by the Brazilian government over the past decades were not only effective in reducing deforestation locally, but were able to retain more than 80% of this effect when considering national deforestation levels. This finding suggests that concerns about leakage outdoing the majority of gains in localised anti-deforestation policies might be unwarranted, and that these policies may be effective in reducing global forest loss.

5 Conclusion

To what extent do spatially targeted policies are globally effective when considering the possible geographical displacement of environmental damage as a response to the policy? The answer to this question has deep implications for policy design, given the global public good nature of ecosystem conservation. We address this issue in the context of tropical deforestation - an activity that has been responsible for one-fifth of global CO2 emissions in

Figure 15: Cumulative leakage from the Protected Areas policy



Note: This figure illustrates the evolution of total estimated deforestation leakage under the Protected Areas policy between the years of 2007 and 2019. Leakage as calculated as the ratio between: total increase in deforestation in non-targeted areas (differece between observed deforestation and the simulated deforestation in a counterfactual where prices do not adjust to policy), and total avoided deforestation in targeted areas (difference between simulated deforestation in a counterfactual where the policy was never implemented and the simulated deforestation in a counterfactual where the leakage channel is shut down). A leakage of 100% would mean that the entire deforestation reduction caused by the localised policies is outdone by increases in deforestation in other regions.

the past two decades - in Brazil - home to a third of the world's remaining rainforests.

The issue of conservation leakage is intrinsically a general equilibrium problem in space: a policy crackdown on deforestation in a spatially delimited region changes economic incentives through a shock to the price of deforested land, generating both sectoral and spatial reallocation. This reallocation depends on the costs of deforestation, the productivity of industries with varying land-intensity, consumers' substitution elasticities between goods and between origins, trade costs, regional amenities and migration frictions. Consequently, reduced-form evaluations conflate deforestation reduction in targeted areas with leakage to non-targeted ones. To separate these two effects, we develop a multi-sector spatial economic model of the Brazilian economy. Given the tight link between deforestation and agriculture, we model agricultural land as the endogenous output of a deforestation sector intermediate to the production of agricultural goods.

We apply our framework to quantify the global effects on national deforestation of two spatially targeted policies implemented by the Brazilian government over the past decades: the establishment of priority regions with high levels of illegal deforestation that receive extra resources for command and control, and the delimitation of protected areas with high forest coverage where deforestation is entirely banned. We find that both policies are highly effective in reducing global deforestation. Over a period of 12 years, leakage outside of targeted areas undoes between 10% and 15% of the policy impact in targeted areas.

Overall, we show that localised policies can be effective in reducing overall deforestation. Our theoretical framework provides a useful starting point to answer other questions of interest that require a spatial general equilibrium model of the relationship between deforestation and agricultural goods. Future research can build on our framework to study, for example, other indirect consequences of deforestation such as its impact on rainfall in nearby areas (Leite-Filho et al., 2021), or the role of technological change in agriculture on environmental conservation.

References

L. J. Alston, G. D. Libecap, and B. Mueller. *Titles, Conflict, and Land Use: The Development of Property Rights and Land Reform on the Brazilian Amazon Frontier*. University of Michigan Press, 1999. ISBN 9780472110063. URL http://www.jstor.org/stable/10.3998/mpub.16208.

J. A. Alvarez. The agricultural wage gap: Evidence from brazilian micro-data. *American Economic Journal: Macroeconomics*, 12(1):153–73, January 2020. doi: 10.1257/mac.20170436. URL https://www.aeaweb.org/articles?id=10.1257/mac.20170436.

J. E. Anderson. A theoretical foundation for the gravity equation. *The American Economic Review*, 69(1):106–116, 1979. ISSN 00028282. URL http://www.jstor.org/stable/1802501.

P. S. Armington. A theory of demand for products distinguished by place of production (une théorie de la demande de produits différenciés d'après leur origine) (una teoría de la demanda de productos distinguiéndolos según el lugar de producción). *Staff Papers (International Monetary Fund)*, 16(1):159–178, 1969. ISSN 00208027. URL http://www.jstor.org/stable/3866403.

J. Assunção and R. Rocha. Getting greener by going black: the effect of blacklisting municipalities on amazon deforestation. *Environment and Development Economics*, 24(2): 115–137, 2019. doi: 10.1017/S1355770X18000499.

J. Assunção, C. Gandour, and R. Rocha. Deforestation slowdown in the brazilian amazon: prices or policies? *Environment and Development Economics*, 20(6):697–722, 2015. doi: 10.1017/S1355770X15000078.

T. Boppart. Structural change and the kaldor facts in a growth model with relative price effects and non-gorman preferences. *Econometrica*, 82(6):2167–2196, 2014. doi: https://doi.org/10.3982/ECTA11354. URL https://onlinelibrary.wiley.com/doi/abs/10.3982/ECTA11354.

J. Brner and S. Wunder. Paying for avoided deforestation in the brazilian amazon: from cost assessment to scheme design. *International Forestry Review*, 10(3):496–511, 2008.

R. Burgess, M. Hansen, B. A. Olken, P. Potapov, and S. Sieber. The political economy of deforestation in the tropics. *The Quarterly Journal of Economics*, 127(4):1707–1754, 2012. ISSN 00335533, 15314650. URL http://www.jstor.org/stable/41812147.

R. Burgess, F. Costa, and B. A. Olken. The brazilian amazon's double reversal of fortune. 2019.

P. Bustos, B. Caprettini, and J. Ponticelli. Agricultural productivity and structural transformation: Evidence from brazil. *American Economic Review*, 106(6):1320–65, June 2016. doi: 10.1257/aer.20131061. URL `https://www.aeaweb.org/articles?id=10.1257/aer.20131061`.

S. Calonico, M. D. Cattaneo, and R. Titiunik. Robust nonparametric confidence intervals for regression-discontinuity designs. *Econometrica*, 82(6):2295–2326, 2014.

CNA and CEPEA. Pib do agronegócio. 2023.

B. R. Copeland and M. S. Taylor. Trade, growth, and the environment. *Journal of Economic Literature*, 42(1):7–71, March 2004. doi: 10.1257/002205104773558047. URL `https://www.aeaweb.org/articles?id=10.1257/002205104773558047`.

X. Cui. Climate change and adaptation in agriculture: Evidence from us cropping patterns. *Journal of Environmental Economics and Management*, 101:102306, 2020. ISSN 0095-0696. doi: https://doi.org/10.1016/j.jeem.2020.102306. URL `https://www.sciencedirect.com/science/article/pii/S0095069620300292`.

J. M. C. da Silva, L. C. F. Barbosa, J. Topf, I. C. G. Vieira, and F. R. Scarano. Minimum costs to conserve 80% of the brazilian amazon. *Perspectives in Ecology and Conservation*, 20(3):216–222, 2022.

J. R. de Vasconcelos. Matriz do fluxo de comércio interestadual de bens e serviços no brasil – 1999. Job market paper, August 2001.

R. Dix-Carneiro and B. K. Kovak. Trade liberalization and regional dynamics. *American Economic Review*, 107(10):2908–2946, 2017.

T. Domiguez-Iino. Efficiency and redistribution in environmental policy: An equilibrium analysis of agricultural supply chains. 2021.

D. Donaldson and R. Hornbeck. Railroads and american economic growth: A "market access" approach. *The Quarterly Journal of Economics*, 131(2):799–858, 02 2016. ISSN 0033-5533. doi: 10.1093/qje/qjw002. URL `https://doi.org/10.1093/qje/qjw002`.

DOU. Lei nº 6.321, de 21 de dezembro de 2007. *Diário Oficial [da] República Federativa do Brasil*, 2007.

DOU. Lei nº 12.651, de 25 de maio de 2012. *Diário Oficial [da] República Federativa do Brasil*, 2012.

F. Eckert and M. Peters. Spatial structural change. Working Paper 30489, National Bureau of Economic Research, September 2022. URL http://www.nber.org/papers/w30489.

F. Farrokhi and H. S. Pellegrina. Global trade and margins of productivity in agriculture. Working Paper 27350, National Bureau of Economic Research, June 2020. URL http://www.nber.org/papers/w27350.

B. Herrendorf, R. Rogerson, and Ákos Valentinyi. Chapter 6 - growth and structural transformation. In P. Aghion and S. N. Durlauf, editors, *Handbook of Economic Growth*, volume 2 of *Handbook of Economic Growth*, pages 855–941. Elsevier, 2014. doi: https://doi.org/10.1016/B978-0-444-53540-5.00006-9. URL https://www.sciencedirect.com/science/article/pii/B9780444535405000069.

A. Hsiao. Coordination and commitment in international climate action: Evidence from palm oil. 2021.

IPCC. Ipcc, 2019: Climate change and land: an ipcc special report on climate change, desertification, land degradation, sustainable land management, food security, and greenhouse gas fluxes in terrestrial ecosystems. 2019.

S. Jayachandran, J. de Laat, E. F. Lambin, C. Y. Stanton, R. Audy, and N. E. Thomas. Cash for carbon: A randomized trial of payments for ecosystem services to reduce deforestation. *Science*, 357(6348):267–273, 2017. doi: 10.1126/science.aan0568. URL https://www.science.org/doi/abs/10.1126/science.aan0568.

D. Lagakos and M. E. Waugh. Selection, agriculture, and cross-country productivity differences. *American Economic Review*, 103(2):948–80, April 2013. doi: 10.1257/aer.103.2.948. URL https://www.aeaweb.org/articles?id=10.1257/aer.103.2.948.

A. T. Leite-Filho, B. S. Soares-Filho, J. L. Davis, G. M. Abrahão, and J. Börner. Deforestation reduces rainfall and agricultural revenues in the brazilian amazon. *Nature Communications*, 12(1):2591, 2021. doi: 10.1038/s41467-021-22840-7. URL https://doi.org/10.1038/s41467-021-22840-7.

M. Morten and J. Oliveira. The effects of roads on trade and migration: Evidence from a planned capital city. 2018.

H. S. Pellegrina and S. Sotelo. Migration, specialization, and trade: Evidence from brazil's march to the west. Working Paper 28421, National Bureau of Economic Research, January 2021. URL http://www.nber.org/papers/w28421.

A. Pfaff and J. Robalino. Spillovers from conservation programs. *Annual Review of Resource Economics*, 9(1):299–315, 2017. doi: 10.1146/annurev-resource-100516-053543. URL https://doi.org/10.1146/annurev-resource-100516-053543.

E. Souza-Rodrigues. Deforestation in the Amazon: A Unified Framework for Estimation and Policy Analysis. *The Review of Economic Studies*, 86(6):2713–2744, 12 2018. ISSN 0034-6527. doi: 10.1093/restud/rdy070. URL https://doi.org/10.1093/restud/rdy070.

D. Szerman, J. Assunção, M. Lipscomb, and A. M. Mobarak. Agricultural productivity and deforestation: Evidence from brazil. Working Paper 1091, Yale Economic Growth Center, January 2022. URL https://elischolar.library.yale.edu/egcenter-discussion-paper-series/1091.

6 Appendix

6.1 Deriving local expenditure by sector

Wages in equilibrium assuming Frechet Frank-copula. Let the joint distribution of pairs of individual productivities (z_{ir}^A, z_{ir}^{NA}) is given by the Frank copula as in LW with parameters $(\chi^A, \chi^{NA}, \rho)$. Then there are no simple closed-form expressions for the share of employment in agriculture and the labour income in each sector. Instead,

$$s_r^A = Pr\big[z_{ir}^{NA}/z_{ir}^A \leq w_r^A/w_r^{NA_r}\big] = \int_0^\infty \int_0^{z^A w_r^A/w_r^{NA}} f(z^A, z^{NA})\, dz^{NA} dz^A$$

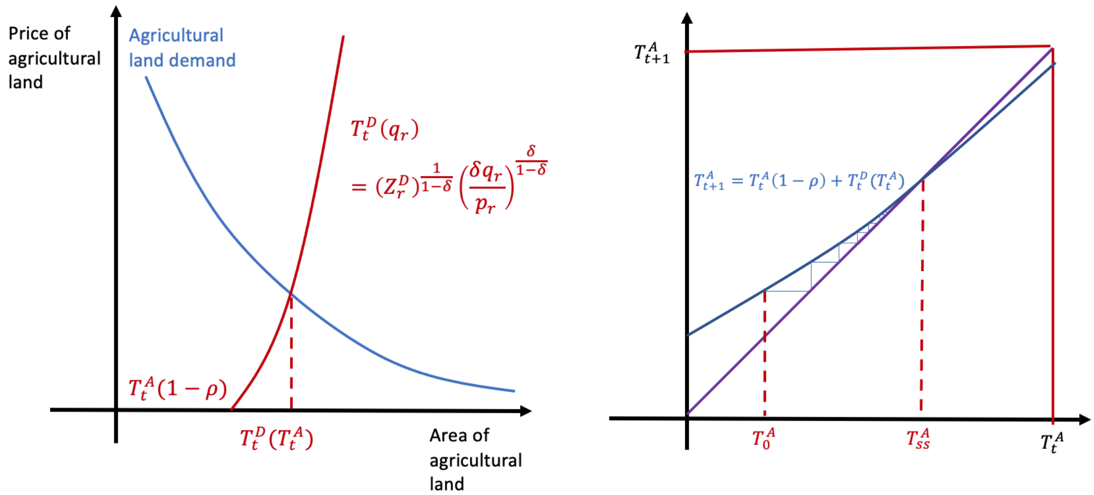
$$\overline{y}_r^{LA} = w_r^A \mathbb{E}\big[z_{ir}^A \big| z_{ir}^{NA}/z_{ir}^A \leq w_r^A/w_r^{NA_r}\big] = w_r^A \frac{1}{s_r^A} \int_0^\infty \int_0^{z^A w_r^A/w_r^{NA}} z^A f(z^A, z^{NA})\, dz^{NA} dz^A$$

$$\overline{y}_r^{LNA} = w_r^{NA} \mathbb{E}\big[z_{ir}^{NA} \big| z_{ir}^{NA}/z_{ir}^A > w_r^A/w_r^{NA_r}\big] = w_r^{NA} \frac{1}{s_r^{NA}} \int_0^\infty \int_{z^A w_r^A/w_r^{NA}}^\infty z^{NA} f(z^{NA}, z^{NA})\, dz^{NA} dz^A$$

Aggregate agricultural expenditure share in each region The share of consumer expenditure in agriculture of a household with expenditure e_i equals

$$\vartheta_{ir}^A = \phi + \nu \Big(\frac{p_r^A}{p_r^{NA}}\Big) \Big(e_{ir}\Big)^{-\eta} p_r^\eta. \tag{12}$$

Figure 16: Agricultural Land Accumulation with Forest Regrowth



Note: This figure illustrates the transition dynamics in the market for agricultural land under a constant forest regrowth rate. The graph on the left shows supply and demand curves for land, and the equation for the equilibrium in the market for the equilibrium quantity. The graph on the right shows the transition curve for agricultural land: with forest regrowth, deforestation reaches a steady state, where the blue curve intercepts the 45-degree line.

"

Thus, the total agricultural expenditure in region r equals

$$X_r^A = \phi p_r I_r^D + \int \vartheta_{ir}^A e_{ir}\, dG(i) = \phi(E_r + p_r I_r^D) + \nu\Big(\frac{p_r^A}{p_r^{NA}}\Big) p_r^\eta \int \big(e_{ir}\big)^{1-\eta}\, dG(i),$$

Assume that people with higher labour incomes get proportionately higher land rents and also spend proportionately more on deforestation. Define the ratio between land rents and labour income from the aggregates, a_r, and the ratio of the total income spent on deforestation investments, b_r.

$$a_r \equiv \frac{v_r T_r^A}{\int_i y_{ir}^L dG(i)}, \quad b_r \equiv \frac{p_r I_r^D}{\int_i y_{ir}^L dG(i) + v_r T_r^A}.$$

Then the assumption of proportional land rents and deforestation expenditures can be expressed formally as:

$$(1)\, y_{ir}^T = a_r y_{ir}^L, \quad (2)\, x_{ir}^d = b_r y_{ir} \implies (3)\, \frac{e_{ir}}{y_{ir}^L} = (1+a_r)(1-b_r) \equiv m_r$$

$$e_{ir} = \underbrace{\max\{w_r^{NA} z_{ir}^{NA}, w_r^A z_{ir}^A\}}_{y_{ir}^L} m_r,$$

where

$$m_r = \frac{E_r}{L_r \overline{y}_r^L}, \quad \overline{y}_r^L = s_r^A y_r^{LA} + s_r^{NA} y_r^{LNA} = \mathbb{E}(y_{ir}^L)$$

For the second term of the total agricultural expenditure equation, we compute the following integral

$$\begin{aligned}
\int (e_{ir})^{1-\eta} dG(i) =& (m_r)^{1-\eta} L_r \mathbb{E}\big((y_{ir}^L)^{1-\eta}\big) \\
\mathbb{E}\big((y_{ir}^L)^{1-\eta}\big) =& \Bigg(\int_0^\infty \int_0^{z^A w_r^A / w_r^{NA}} (w_r^A z^A)^{1-\eta} f(z^A, z^{NA})\, dz^{NA} dz^A \\
& + \int_0^\infty \int_{z^A w_r^A / w_r^{NA}}^\infty (w_r^{NA} z^{NA})^{1-\eta} f(z^A, z^{NA})\, dz^{NA} dz^A \Bigg)
\end{aligned}$$

Thus, the aggregate local share of consumer expenditure in agriculture equals

$$\vartheta_r^A = \phi + \nu \Big(\frac{p_r^A}{p_r^{NA}}\Big) \Big(\frac{E_r}{p_r L_r}\Big)^{-\eta} \frac{\mathbb{E}\big((y_{ir})^{1-\eta}\big)}{\big(\mathbb{E}(y_{ir})\big)^{1-\eta}}. \tag{13}$$

6.2 Deforestation parameters estimation

- Recall optimal deforestation level condition

$$T_{rt}^D = (T_{rt}^F)^{\frac{\psi}{1-\delta}} (\overline{Z}_{rt}^D)^{\frac{1}{1-\delta}} \Big(\frac{\delta q_{rt}}{p_{rt}}\Big)^{\frac{\delta}{1-\delta}}$$

- Take logs

$$\begin{aligned}
\log(T_{rt}^D) =& \frac{1}{1-\delta} \log(\overline{Z}_{rt}^D \delta^\delta) + \frac{\psi}{1-\delta} \log(T_{rt}^F) \\
& + \frac{\delta}{1-\delta} \log q_{rt} - \frac{\delta}{1-\delta} \log p_{rt}
\end{aligned}$$

- Regression specification

$$\boxed{\log(T_{rt}^D) = \alpha_r^1 + \alpha_t^2 + \frac{\delta}{1-\delta} \log v_{rt} + \epsilon_{rt}}$$

- Instrument for $\log v_{rt}$ with $\log v_{rt-1}$ and $\log v_{rt-2}$ (Anderson-Hsiao)

- Implied δ: 0.28-0.52

Table 5: Estimating δ

	Outcome: log(Deforestation)			
	(1) OLS	(2) OLS	(3) IV	(4) IV
IHS(v_r)	0.391***	0.253***	0.817***	1.065***
	[0.025]	[0.038]	[0.165]	[0.283]
Mun. FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Instrument	-	-	First Difference	First Difference
Weight	None	Agri. Area	None	Agri. Area
R2	0.240	0.346	0.002	-0.001
Observations	89072	89072	77938	77938

Table 6: Estimation of deforestation elasticity

Now let us derive the equation that we use to estimate ψ.

- Steady state condition

$$T^D_{r,ss} = \rho T^A_{r,ss}$$

$$\iff \frac{T^F_{r,ss}}{T^A_{r,ss}} = \rho(\overline{Z}^D_{r,ss})^{\frac{1}{\delta-1}} \delta^{\frac{\delta}{\delta-1}} \left(\frac{q_{r,ss}}{p_{r,ss}}\right)^{\frac{\delta}{\delta-1}} (T^F_r)^{1-\frac{\psi}{1-\delta}}$$

- Take logs

$$\begin{aligned} \log\left(\frac{T^F_{r,ss}}{T^A_{r,ss}}\right) &= \log\left(\rho(\overline{Z}^D_{r,ss})^{\frac{1}{\delta-1}} \delta^{\frac{\delta}{\delta-1}}\right) \\ &\quad - \frac{\delta}{1-\delta}\log\left(\frac{q_{r,ss}}{p_{r,ss}}\right) \\ &\quad + \left(1 - \frac{\psi}{1-\delta}\right)\log(T^F_{r,ss}) \end{aligned}$$

- Regression specification

$$\boxed{\log\left(\frac{T^F_{rt}}{T^A_{rt}}\right) = \alpha_{s(r)t} - \frac{\delta}{1-\delta}\log(v_{rt}) + \left(1 - \frac{\psi}{1-\delta}\right)\log(T^F_{rt}) + \epsilon_{rt}}$$

- Implied δ: 0.26-0.29
- Implied ψ: 0.30-0.32

Table 7: Estimating ψ

	Outcome: log((Forest Area)/(Agri. Area)) (1)	(2)	(3)
log(Unprotected Forest Area)	0.668***	0.589***	0.571***
	[0.004]	[0.004]	[0.004]
$\log(v_r)$		-0.419***	-0.357***
		[0.005]	[0.007]
State X Year FE	Yes	Yes	Yes
Weight	Agri. Area	Agri. Area	Agri. Area
Sample			Low deforest.
R2	0.735	0.793	0.805
Observations	28244.000	27875.000	16899.000

6.3 Preference parameters estimation

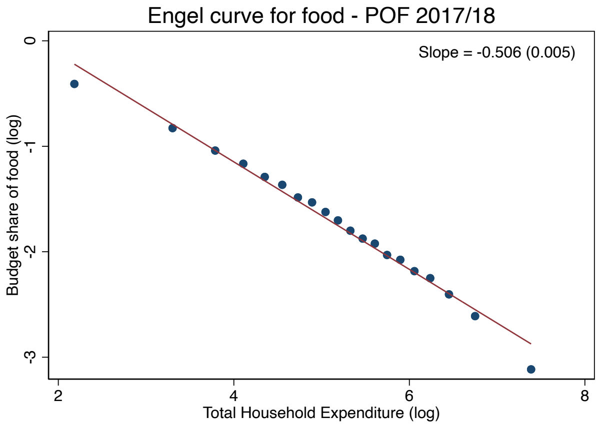
The parameter governing the non-homotheticity of preferences, η, captures the rate at which higher-income consumers shift their budget shares away from agricultural goods (i.e. food items). We estimated by regressing the logarithm of the share of expenditure in food on the logarithm of total expenditure using the household-level data from the 2017/2018 consumer expenditure survey (POF). The table below shows the coefficients from that regression. For robustness, we consider both food expenditure as a share of total non-durables expenditure (column 1 - our preferred specification) and as a share of total household income (column 2).

	Outcome: log-budget share on food items (1)	(2)
Log(Non-durable Expenditure)	-0.506***	
	[0.005]	
Log(Income)		-0.575***
		[0.008]
Constant	0.914***	0.632***
	[0.025]	[0.056]
R^2	0.309	0.195
Dep. Var. Mean	-1.730	-3.229
Observations	45322	45322

Note: *p<0.1, **p<0.05, *** p<0.01.

Table 8: Estimates of Engel curves for food consumption

Figure 17: Engel curve for food consumption



Note: This figure displays the estimated Engel curve for food: the relationship between income and budghet share of food, which illustrates the non-homotheticity in consumer preferences. The dots represent the average share of total household expenditure dedicated to food items for 20 bins of total househodl expenditure.

6.4 Migration parameters

- Assume the utility of migrant i moving from o to d depends on:
  - consumption utiliy V_d
  - residential amenities B_d
  - an origin-destination migration cost μ_{od}
  - an idiosyncratic preference shock with Frechet distribution and shape parameter ϵ

$$\text{Share of Migrants}_d^o = \frac{\left(V_d B_d \mu_d^o\right)^\epsilon}{\sum_j \left(V_j B_j \mu_j^o\right)^\epsilon}$$

- Taking logs it becomes the following equarion with origin and destination fixed effects

$$\log(\text{Share of Migrants}_d^o) = \delta_o + \delta_d + \epsilon \log\left(\mu_d^o\right)$$

- Equilibrium conditions, up to approximation, yield

$$\log(\text{Income}_d^o) = -\frac{1}{\epsilon}\log(\text{Share of Migrants}_d^o) + \xi_o + \xi_d + u_d^o$$

- Column (2) below show results of PPML regression

$\Rightarrow\ \epsilon \approx 14.08$

	(1) Share of Migrants	(2) Monthly income
log(Distance+1)	-1.550***	
	[0.002]	
log(Share of Migrants)		-0.071***
		[0.004]
Origin FE	Yes	Yes
Dest. FE	Yes	Yes
Pseudo R2	0.805	0.221
Observations	3.11e+05	57227

Table 9: Estimation of migration parameters

6.5 Trade parameters estimation

Iceberg trade costs are calibrated to fit trade costs in 1999 and 2017 from the residuals in the regression equation of (log) trade flows with origin and destination fixed effects. Taking logarithms from equation 6, and sending the volume of expenditure of consumers in region d to the right hand side so that the data on trade flow rather than shares can be used, yields

$$\log X_d^o = \log p^o - \log p_d X_d + (1 - \sigma)\log(\tau_d^o)$$

Notice that this data does not include the observations for which the origin and the destination are the same, as it is data from inter-state customs. Thus, if we assume that $\tau_d^o = (\text{dist}_{od} + 1)^\kappa$, then the coefficient of the log-log regression of trade flows on distance (+1) with origin and destination fixed effects equals $\kappa(1 - \sigma)$. Given the results from the regressions reported in the table below, $\kappa \approx \frac{-0.9}{1-9} \approx 0.11$.

6.6 Model fit

	Outcome: log-trade flows	
	(1)	(2)
log(Distance+1)	-0.871***	-0.916***
	[0.053]	[0.036]
Year	1999	2017
Origin FE	Yes	Yes
Dest. FE	Yes	Yes
Pseudo R2	0.966	0.957
Observations	702	702

Note: *p<0.1, **p<0.05, *** p<0.01.

Table 10: Estimation of trade parameters

Figure 18: Model fit: land prices

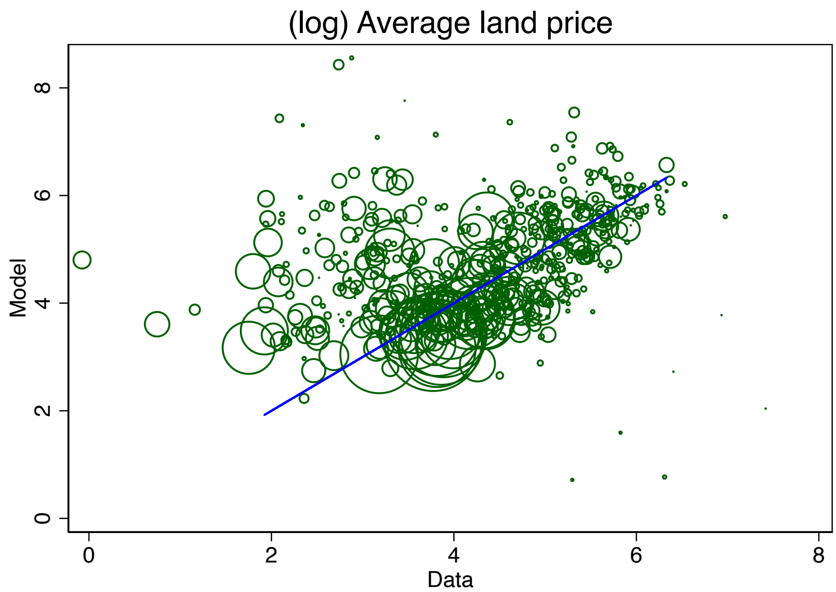
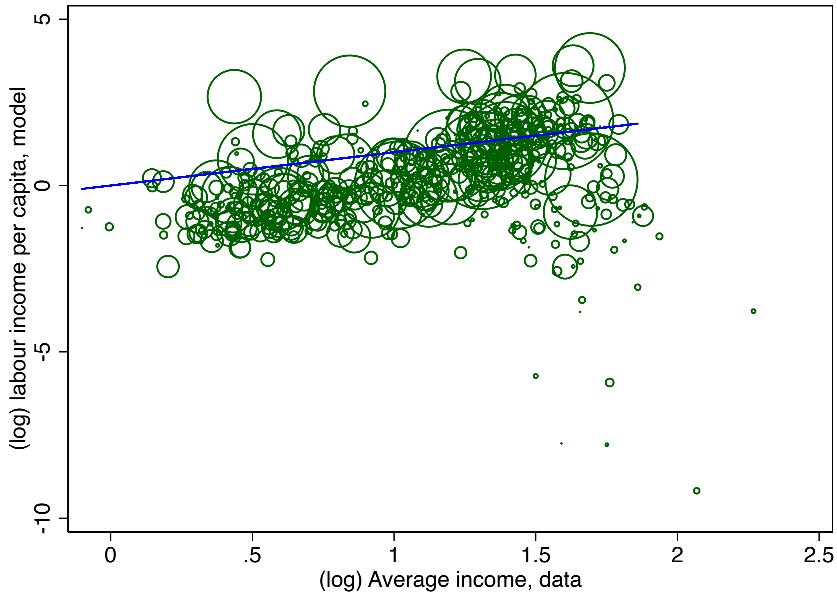


Figure 19: Model fit: comparing predicted and observed land prices

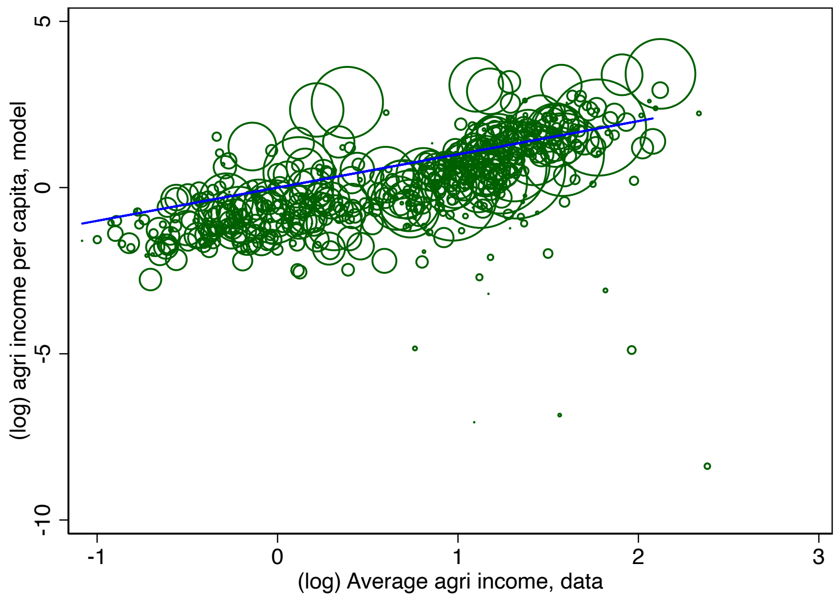
Note: This compares the predicted land prices from model calibration q_r with observed land value. Each circle is a micro-region, and the size of the circle corresponds to the total agricultural area

Figure 20: Model fit: labour income



Note: This compares the predicted average labour income from model calibration y_r^L with observed average labour income. Each circle is a micro-region, and the size of the circle corresponds to the total agricultural area

Figure 21: Model fit: labour income in agriculture



Note: This compares the predicted average labour income from model calibration y_r^{LA} with observed average labour income from the agricultural sector. Each circle is a micro-region, and the size of the circle corresponds to the total agricultural area

Figure 22: Model fit: value added in agriculture

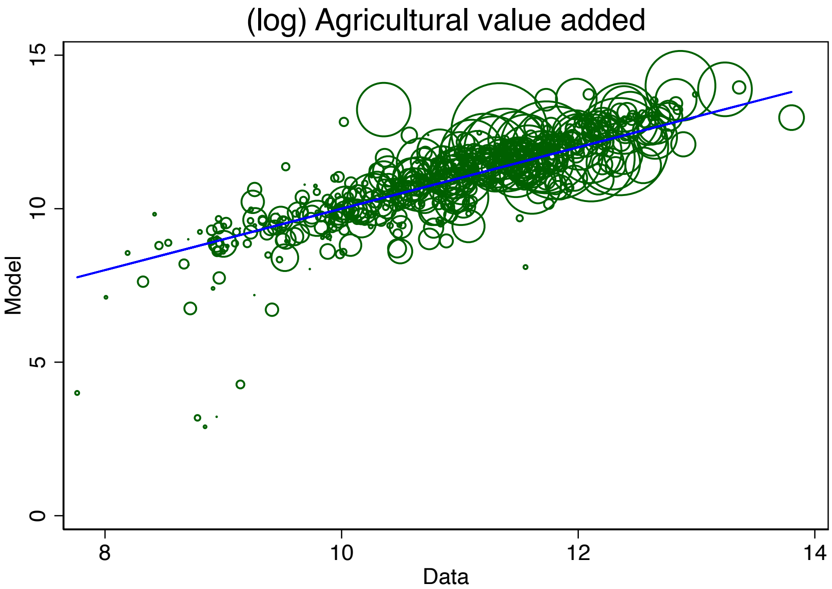
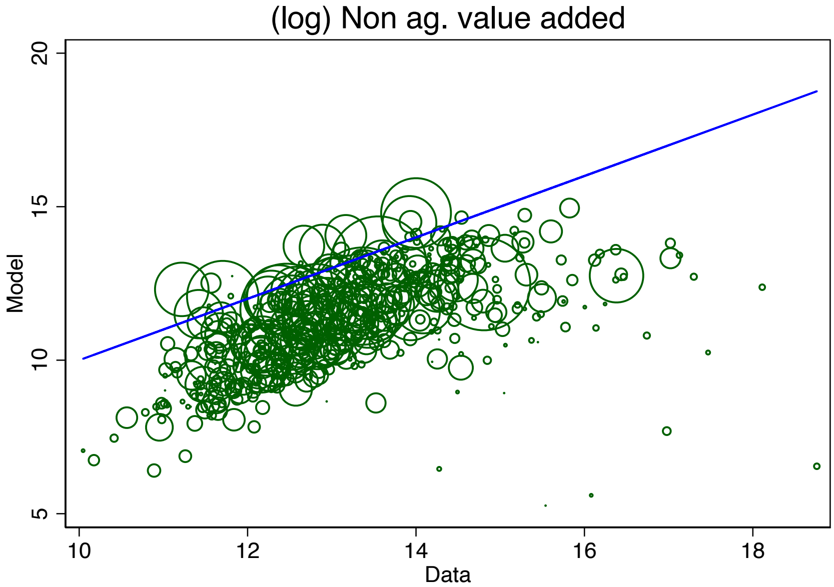


Figure 23: Model fit: comparing predicted and observed agricultural value added

Note: This compares the predicted total GDP in the agricultural sector from model calibration with observed value added from the agricultural sector. Each circle is a micro-region, and the size of the circle corresponds to the total GDP

Figure 24: Model fit: value added in non-agriculture



Note: This compares the predicted total GDP in the non-agricultural sector from model calibration with observed value added from the non-agricultural sector. Each circle is a micro-region, and the size of the circle corresponds to the total GDP