Agricultural Trade and Deforestation: the Role of New Roads

Douglas Gollin Julien Wolfersberger*

Abstract

In this paper, we study how new roads affect the spatial patterns of agricultural production and consequently impact deforestation and development outcomes, focusing on the historical experience of Brazil. We find that the expansion of Brazil's road network since the 1990s can account for nearly a tenth of the total amount of deforestation that the country has experienced, with significant variation across regions. Perhaps surprisingly, our results suggest that the increase in agricultural income attributable to changes in transport costs has been more limited. Focusing on complementarities with technical change, we examine how improved market access combined with new agricultural technologies impacted land conversion.

Keywords: Trade Frictions ; Natural Resources ; Spatial Economics ; Transportation Infrastructure ; Economic Development

JEL codes: F18 ; O13 ; Q10 ; Q56 ; R12

---

*Douglas Gollin: Tufts University and University of Oxford. Julien Wolfersberger: Paris-Saclay University, AgroParisTech and Climate Economics Chair of Paris-Dauphine.
We thank Christophe Gouel for helpful discussions at different stages of this work. We would like also to thank Romero Rocha, Léa Crépin, Valentin Guye, Francois Bareille, Alex Pfaff, Veronica Salazar-Restrepo and the participants of the following conferences where this work was presented: LSE Environment Week, BIOECON, Center for Studies of African Economies, Oxford Economics Jamboree, and the 2021 EAERE thematic session "Spatial Analysis of Development and Conservation Policies related to Deforestation". We also thank the participants of the following seminars: Institute of Economics of the Federal University of Rio de Janeiro, Climate Economics Chair of Paris-Dauphine, BETA Université de Lorraine, International Center for Research on Environment and Development.

1 Introduction

The destruction in recent decades of forests in the world's tropical areas stands as one of the most devastating and consequential human impacts on the biosphere. Tropical forests represent a key ecosystem, and their loss causes major environmental damages in the form of greenhouse gas emissions, biodiversity loss, and the degradation of ecosystem services. Rainforest destruction also has the perverse effect of undermining the economic development that drives the process, by putting at risk agricultural productivity and undermining the livelihood strategies of those people and communities who are exposed to climate change (IPBES, 2019).

In this paper, we study how new roads affect patterns of agricultural production and trade within Brazil, with implications for the extent and location of deforestation. Road construction, along with other improvements in transportation infrastructure, has typically been viewed by economists and policy makers as a key instrument to promote economic growth and development – a view that can be traced at least to Adam Smith (1776).[1] But as economists have long realised, investments in road construction, by connecting previously isolated areas to the market, may also be a key driver of forest clearing. The implied trade-off between economic growth and the environment is a critical one for developing countries – and arguably for the planet as a whole.

We build and calibrate a quantitative spatial general equilibrium model of agricultural trade that we apply to Brazil's 137 meso-regions. We use this model to assess the effect of new roads on the spatial equilibrium of agriculture and, in turn, on pristine forests in Brazil. We do so by using unique data sources, based on the latest Brazilian agricultural censuses as well as high-resolution satellite data on forest cover and transport networks.

In contrast to many previous studies on the relationship between road construction and deforestation, we use a general equilibrium framework that allows us to consider spatially diffuse spillovers from road construction. For instance, a road built into one agricultural region may stimulate production in that location to serve distant urban markets. That may, in turn, alter land use patterns not only in adjacent locations but in other parts of the country that produce similar outputs. Those land use changes may, in turn, drive further impacts in other locations. In this context, a general equilibrium framework is key to understanding the overall impacts on land use and deforestation. Our model uses data

[1] In his *Inquiry into the Nature and Causes of the Wealth of Nations*, Smith (1776) writes: "*Good roads, canals, and navigable rivers, by diminishing the expense of carriage, put the remote parts of the country more nearly upon a level with those in the neighbourhood of the town. They are upon that account the greatest of all improvements.*"

on the agro-ecological properties of different meso-regions to guide our understanding of the market-mediated connections between different locations, by building in an understanding of which crops are best suited to particular meso-regions.

Our results indicate that reductions in transport costs since 1995 account for nearly a tenth of total forest loss in Brazil, with substantial differences across regions. For instance, in a state like *Pará*, which has become a major locus of deforestation since the 2000s, we find that these reductions explain up to 70% of the deforestation observed in the data.

To guide our modelling choices, we start by documenting four motivating facts about internal trade, deforestation and agriculture in Brazil. First, proximity to roads varies greatly across regions, suggesting important differences in market access. Second, forest cover is positively correlated to travel time to the closest international port, illustrating the links between transport costs, market access and deforestation. Third, there is substantial heterogeneity in crop suitability across space, potentially influencing comparative advantages and specialization. Fourth, agricultural crops vary in their land intensity, with implications for the commercial pressures on standing forests in different locations.

Equipped with those facts, our model is structured as follows. The economy includes a large number of locations (corresponding to Brazilian meso-regions in our empirical application), one trade hub, and two sectors: agriculture and non-agriculture. We view the non-agricultural good as a composite good encompassing both manufacturing and service activities; for simplicity, we use the term "manufacturing" to refer to this sector. In the agricultural sector, farmers produce output using labour, imported intermediate inputs, and land. They can grow an array of crops over a continuum of plots that vary in their crop-specific land productivity. Given that crops have different land intensities, pressure on forests can vary depending on the pattern of specialization of a given region. In our setting, rural locations sell agricultural goods to the trade hub, following an Eaton and Kortum (2002) structure where transportation is subject to iceberg costs.

To open a plot for "productive use", farmers have to pay a fixed cost to convert the land from forest to cropland.[2] This cost, which is plot-specific, ultimately determines the total amount of land devoted to agriculture in a region, as in Costinot et al. (2016) or Farrokhi and Pellegrina (2023). The logic of land conversion in the model is as follows: as transport costs

[2]Farmers in our model do not view the ecosystem services provided by forests as a productive use; any environmental benefits of forests are externalities that farmers ignore. In our model, farmers view land only as an input into agricultural production. We recognize that some farmers take seriously the environmental impacts of their actions and their responsibilities as "stewards of the land." We also recognize that some market segments and market mechanisms have emerged that encourage agricultural producers to internalize the environmental benefits of standing forests. In our analysis, we treat this as a negligible feature of the overall market.

decrease, market prices of agricultural commodities rise within previously isolated areas. The increase in local prices boosts the profitability of agriculture relative to the fixed cost of land conversion, thus creating incentives for farmers to further deforest. This is the effect of improving market access. Because crops have different land intensities, the extent of forest clearing can vary depending on the production pattern of a given region; this is the effect of crop specialization.

In our model, deforestation is further stimulated because the decline in transport costs makes it cheaper for farmers in remote areas to "import" intermediate inputs for agricultural production. Together, the increase in output prices and the decline in input costs make agriculture more profitable and thus stimulate land clearing. These effects are spatially heterogeneous, however. The local patterns will depend on the types of crops that are most suitable in particular regions and the shape of the demand curves for those crops. For those crops with high foreign demand (such as soybeans), expansion of production faces few limits. But where there is little foreign demand, domestic markets may become saturated. Transport cost reductions may lead to greater market integration, allowing for the intensification of production in the most favourable areas and a reduction in land pressure on the extensive margin. There is potential, in other words, for improvements in transportation infrastructure to enhance specialization and thus to reduce pressure on Brazil's land resources. Our analysis suggests that changes in the transportation network will lead to a spatial reallocation of the production of specific crops – and potentially to a reduction of the land devoted to some crops, so the aggregate effects of forest cover are unclear *ex-ante*.

Taking the model to the data, we estimate trade elasticities by exploiting price differences in soybeans across space, taking advantage of the fact that soybeans are overwhelmingly exported, so that prices are determined on world markets. We then calibrate the model using a combination of parameters taken from the literature and those calibrated to match moments of the data. In particular, we calibrate the model to reproduce the observed spatial equilibrium of the economy for 2017 (the latest year for which agricultural census data are available). We then proceed to conduct two main counterfactual exercises.

In our first counterfactual exercise, we examine what would be the spatial equilibrium if transport costs had remained at the level implied by the road network of the early 1990s. This allows us to quantify the impact of the roads constructed since this period. We examine the impacts on deforestation but also consider other development outcomes related to agriculture. We contextualize our results by comparing the predicted changes with the *actual* changes by exploiting data from the 1995 and 2017 agricultural censuses.

In our second counterfactual, we use our model to evaluate the potential impacts of the road network in interaction with another major disruption that took place during the same period in Brazil: the introduction of new soybean varieties. To do this, we exploit the estimates of potential crop yields from the FAO-GAEZ database, which provides the maximum attainable yields under traditional and modern agricultural technologies. Following Bustos et al. (2016), we exploit the differential in predicted yields to examine whether the effect of roads was amplified or mitigated by the new agricultural technology.

Our findings suggest that, as a result of changes in transport costs, forested area in Brazil was reduced by 0.59%. Examining other outcomes, we find that Brazil's aggregate exports and income in agriculture increased by 2.3% and 0.4%, respectively, as a result of the decline in transport costs. Our analysis also shows the importance of spatial heterogeneity as a mediator of price shocks experienced by the different Brazilian meso-regions. We document that some key locations in the Amazon biome experienced declines of over 15% in estimated iceberg transport costs since the 1990s. We show that, absent these changes in transport connectivity, deforestation would have been up to 70% lower in those regions, highlighting the sizeable effects of roads. While we find that greater market integration increased agricultural income, we notice that economic gains remain relatively low in comparison to the environmental costs of the new roads.

Looking at the interactions with the introduction of soybean seeds, we find that the combined effect of reductions in transport costs and new agricultural technology explains 40% of total forest loss observed in the data. We also find that it leads to a 6.5% increase in agricultural income and a 14.5% increase in exports. Focusing on the complementarities between infrastructure expansion and technical change, we find that the addition of the new agricultural technology exacerbated the negative impacts of transportation infrastructure in many regions of the Amazon. For instance, in some meso-regions of the state of Mato Grosso, we estimate that the new soybean seeds amplified deforestation attributed to roads by an additional 20,000 hectares of forest loss. In other words, had soy productivity remained at the level of 1995, forest conversion due to roads would have been much lower. In this case, the new technology reinforced the role of comparative advantages shaped by new roads, leading to greater land clearing. Interestingly, our results also highlight that the new technology mitigated the negative impact of roads in some other locations. In these regions, due to productivity growth in soy, the impact of roads was less harmful than it would have been in the absence of innovation.

Our work offers new evidence on the trade-offs between improving market access and preserving the forest. The topic is not a new one, but relative to previous research, our approach offers several advantages. Since locations within a country are linked by trade and markets, it is important to consider general equilibrium linkages between locations and across goods. This is particularly relevant for thinking about large-scale changes in transportation infrastructure, which necessarily alter the entire spatial pattern of production. Although many empirical studies look for impacts of roads within narrowly defined geographic areas, it is difficult to find contexts where "treatment" locations can be compared with untouched "control" locations. Estimates that fail to account for the spillover effects of roads may accordingly be biased (Redding and Rossi-Hansberg, 2017).

We contribute to the existing literature in several ways, as we extensively discuss in section 2 below. Most obviously, we depart from a literature that has relied heavily on reduced-form empirical estimates of the impact of roads on deforestation. We also contribute to a large and growing literature making use of quantitative spatial models for thinking about issues related to agriculture; relative to this body of research, our contribution is to tackle the critical issue of forest clearing. By doing so, we also connect to the larger body of research on the effect of trade on the environment. Finally, this work relates to recent research on the importance of agriculture in development, by emphasizing the potential environmental costs of increased production.

In section 2, we discuss in detail the strands of literature related to our paper. Section 3 presents empirical motivations that structure the model we present in section 4. After a calibration of the model in section 5, we report the results of our two counterfactual experiments in section 6. Section 7 concludes.

2 Related literature

A substantial body of literature in development and environmental economics has specifically studied the links between roads and deforestation, generally by employing reduced-form econometric models. Among these, Pfaff (1999) uses panel data on the Brazilian Amazon over the 1978-1988 period and finds that both own- and neighbouring-county paved roads increase deforestation. Similarly, proximity to major national markets has been found to be positively correlated with forest clearing. Andersen et al. (2002) investigate the impact of roads on forest clearing, along with other development outcomes, for the Brazilian Amazon. Likewise, Pfaff et al. (2007) find that road investments increase deforestation in census tracts without roads and located at less than 100km from those census tracts that received investments. The authors interpret this result in terms of local spillovers. Asher et al. (2020) study the causal impact of transportation infrastructure on forest losses in India. Relying on different identification strategies, they exploit two large-scale transportation projects implemented by the Indian government in the 2000s. They find that the construction of new (last-mile) rural roads did not impact deforestation, while the upgrade of existing highways significantly increased it. They interpret this result in terms of geographical redistribution of economic activities; they argue that large highway investments lead to this kind of spatial impact, unlike last-mile roads. On rural roads, Kaczan (2020) finds that these could actually facilitate a long-term expansion of forest cover in the context of India.

We contribute to this body of literature by employing a complementary method, based on recent developments in the quantitative spatial and trade literature. Specifically, we use a model that allows us to take into account both spillovers and spatial reorganization of economic activities. Furthermore, by allowing prices to adjust both in space and between goods, we can avoid some of the biases that arise in reduced-form work.

We note that the question of where roads should be built, given their potential environmental costs, has also been studied outside economics – and particularly in the environmental sciences. Laurance et al. (2014) compute agricultural gains and the associated environmental costs of new roads across the world, and they map the locations where new roads would provide the largest benefits at the lowest cost for biodiversity and other environmental indicators. Vilela et al. (2020) present a similar exercise for the whole Amazon and quantify the effects of 75 planned projects over 5 years. This question of optimal road location differs from our quantification exercise, and it does not discuss how prices, output and economic activities change across space.

A fairly recent literature uses structural and quantitative spatial models to study the impact of different shocks (e.g., reductions in transport costs) on economic development and welfare (Redding and Rossi-Hansberg, 2017). Some of this literature includes models of land use choices, often in multi-country settings. For example, Costinot et al. (2016) and Gouel and Laborde (2021) study the impact of climate change on agriculture. This is also the case of Farrokhi and Pellegrina (2023) who incorporate endogenous technology choices into a model of agricultural trade.

With these models, other authors have studied questions related to land use within countries. This is the case of Sotelo (2020), who uses a model with heterogeneous land quality in Peru to analyse a set of counterfactuals involving changes in trade costs and commodity prices. Using a similar setting, Pellegrina (2022) studies the consequences, in Brazil, of the introduction of new soybean seeds in the 1970s as well as of a rise in Chinese demand. Fajgelbaum and Redding (2022) also build a model featuring land-use competition within a country to study trade and structural transformation in Argentina. None of the papers from this quantitative literature has, to the best of our knowledge, studied the issues of deforestation and environmental harm that can accompany agricultural expansion.

A set of recent papers also tackle issues related to deforestation and ask different questions. Hsiao (2021) develops a quantitative model to evaluate trade policy as a substitute for domestic regulation, examining the palm oil market. Dominguez-Iino (2021) examines the effects of environmental policies on agricultural value chains in Argentina and Brazil and their impact on the forest. Araujo et al. (2022) study land-use change in the Amazon when farmers internalize the cost of carbon. Farrokhi et al. (2023) study cross-country deforestation given different scenarios of trade policy. We depart from these works by focusing on transportation infrastructures in the historical context of Brazil. Our work is closest to Araujo et al. (2023), but a key difference is that we treat agriculture as a sector producing heterogeneous outputs. This is important since the patterns of specialization across locations will depend on both the production requirements of different crops and the differing structures of demand. Indeed, depending on land productivity and the land intensity of a given crop, the amount of land required to produce a given output can greatly differ, as we discuss in section 3.

Our paper also touches on a macro and growth literature in which a number of papers pose questions about agricultural trade, but without directly accounting for the endogeneity of land clearing. For instance, Allen (2014), Tombe (2015), Porteous (2019) and Adamopoulos (2025) all use quantitative models with agricultural trade. In these models, there are costs associated with shipping goods, but there is not any way to evaluate the environmental impacts of different policy regimes. On issues related to the environment, Desmet and Rossi-Hansberg (2015) use a multi-region dynamic model featuring two sectors to study the spatial impact of climate change. The authors quantify the importance of migration and trade restrictions, as Conte et al. (2021) do, by focusing on the patterns of specialization resulting from climate change dynamics. Desmet et al. (2018) concentrate on coastal flooding and evaluate the costs of sea-level rise worldwide while Balboni (2019) uses a dynamic spatial equilibrium model and focuses on the case of Vietnam. Given current climate change scenarios and future inundation risks, Balboni (2019) finds that coastal favouritism in investment has significant costs. For instance, she finds that avoiding the most vulnerable Vietnamese regions could have led to a 72% increase in welfare gains under a central sea level rise scenario. We contribute to this strand of the literature by studying a different environmental issue, namely deforestation. By doing so, our work is related to Farrokhi et al. (2023), who study this question at the global level.

Our work is also related to an important literature on policy interventions and deforestation in Brazil. Recent works include Souza-Rodrigues (2019) who offers a framework for measuring the cost-effectiveness of alternative policies in the Amazon, or Assunção et al. (2023) who use a novel instrumental approach to study enforcement's impact on deforestation. Similar to our paper, Restrepo and Mariante (2024) use a spatial general equilibrium model to look at land use policies in Brazil and their effectiveness in reducing deforestation.

Finally, this work also speaks to longstanding debates on the effect of transportation infrastructures on economic growth. Recent work on the topic includes Donaldson and Hornbeck (2016) who study the historical impact of railroads on US in 1890. Storeygard (2016) focuses on the role of intercity transport costs for urban income in Sub-Saharan Africa. In the context of India, Donaldson (2018) studies the benefits of historical railroad construction and finds that it has promoted trade and increased real income. Asher and Novosad (2020) assess the impact of a national rural road construction program and identify mixed results. Our contribution is to evaluate the economic benefits together with the ecological costs from a shock in transport costs in Brazil, which our quantitative model allows us to do.

3 Empirical background

This section describes the stylized facts that motivate the structure of our model. Data sources are detailed in Appendix D.

Road density varies across Brazil. Figure 1 shows road and port networks in Brazilian meso-regions.

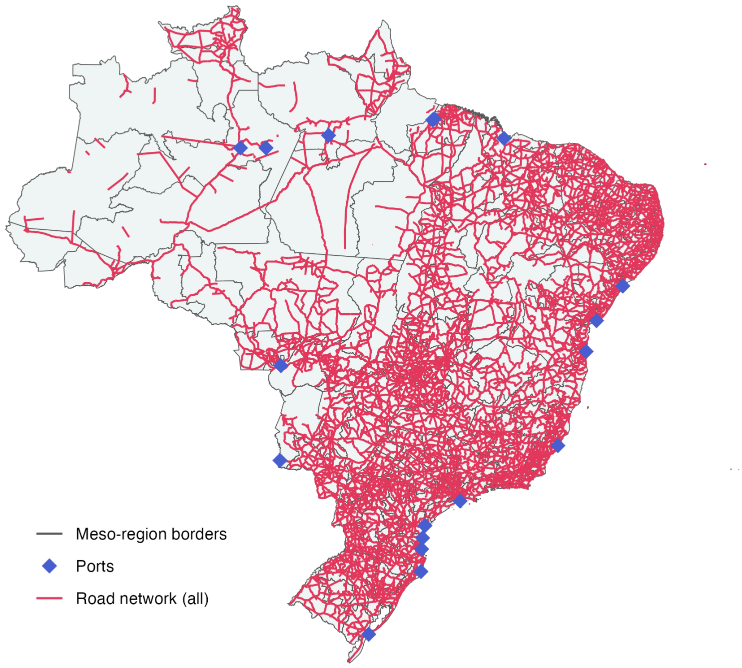


Figure 1: Access to roads and ports across meso-regions. Note: 2014 road network (IBGE) including all types of road but excluding planned roads.

A considerable disparity holds across Brazil in the density of roads. In general, meso-regions in the Central-East, North, and Northeast regions are characterised by a much lower density of roads than those located in the South and along the Atlantic coast. Since the capacity to transport goods at low cost influences local crop production, this unequal proximity to infrastructure also shapes land use and the pressure for agricultural expansion at the expense of standing forests.

Global market access is positively correlated to deforestation. To provide a sense of how market access can influence deforestation, we regress the share of land under forests on the distance between each meso-region's centroid and the closest international port:

$$y_i = \beta_0 + \beta_1 \log \{\text{distance}\}_i + \epsilon_i.$$

Our dependent variable y_i is computed at the meso-region level, and it is the forest area divided by total area, in 2017. We use two different measures of distance to the port. One is distance via the shortest path, using the existing road network. The second is the time duration of travel, using this same network. The two measures are derived by exploiting the

Open Source Routing Machine from OpenStreetMap (Huber and Rust, 2016). Results are shown in Table 3 of Appendix A.

The results show that the share of land under forests in a meso-region is positively correlated to both measures of distance to the closest international port. A decrease of 10% in the travel time to the closest exporting port is associated with a 1.1% decrease in forest cover. This elasticity points to the importance of global markets here as a source of the demand that is driving forest clearing. While *domestic* demand curves for agricultural products might be steeply downward sloping, the size of the international market is such that it exerts a strong effect on land use. In short, land in Brazil is not being cleared for agriculture only to feed domestic consumers; it is being cleared in part to supply an international market, with the result that connectivity to ports matters at least as much as connectivity to domestic urban centres. A caveat here is that Brazil's ports *are* some of its largest cities, so the two sources of demand are not separately identified. Nevertheless, the data point strongly towards the role of international markets in driving deforestation.

Heterogeneity in land productivity matters for impacts. Another factor mediating the impact of road construction is the heterogeneity of land quality for agriculture – and specifically the suitability of different plots for different crops. As an example, Figure 2 shows heterogeneity in land productivity for soybeans relative to corn.

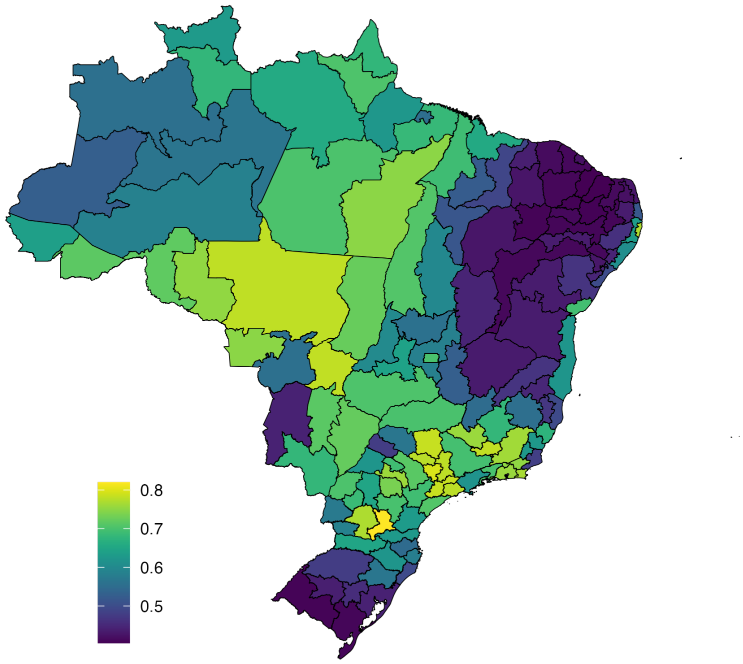


Figure 2: Soybeans Potential Yields Relative to Corn. Source: GAEZv4 (IIASA, FAO). Period: 1981-2010. High inputs and rainfed irrigation.

As the figure shows, in southern Brazil and along the Atlantic coast, some locations can obtain less than 0.5 tons of soybean output per hectare for one ton of corn. Conversely, in some regions located in the states of Mato Grosso or Pará, the relative soybean yields rise to 0.8 tons per hectare.

This heterogeneity in land productivity, illustrated here with relative soybean yields, shapes comparative advantages across regions in the production of crops. As we discuss below, this crop-specific heterogeneity creates varied patterns of pressure on forests. Deforestation in a given region will depend on the suitability of its land for particular crops and on the changes in demand for those crops.

Factor intensity varies across crops. Table 1 shows summary statistics on land intensity for four major crops at the meso-region level. For each crop, the amount of land area (in hectares) per ton produced is displayed.

Each ton of corn output requires, on average, 0.52 hectares of land. For soybeans and wheat, the land requirement is lower, at 0.38 ha for soybeans and wheat, on average. To draw a comparison, according to the data, producing one ton of sugarcane requires 0.04

Table 1: Land-area per unit of output for 4 selected crops (ha/ton). Source: 2017 IBGE agricultural census.

	N	Mean	Std Dev.	Min	Max
Corn	137	0.52	0.51	0.12	2.34
Soy	88	0.38	0.35	0.25	3.13
Sugarcane	137	0.04	0.04	0.01	0.23
Wheat	38	0.38	0.10	0.24	0.70

hectares of land. Notice that corn and sugarcane were produced in every meso-region in 2017; this was not the case for soybeans and wheat.

These land-intensity differences across crops predictably impact land-clearing patterns. When a region specializes in the production of a crop, for any fixed amount of output, pressure on standing forests will vary according to the land-intensity of this crop. This justifies the importance of using a multi-crop model to understand the effect of specialization on deforestation through general equilibrium adjustments.

The empirical motivations presented in this section will guide our model in the following section.

4 The model

This section presents our spatial model of trade and land allocation between agriculture and forest.[3]

4.1 Environment

Consider a spatial economy comprising a trade hub and a set of rural locations $i \in \mathcal{L} = \{1, ...I\}$. The trade hub in the model is a central market where all rural locations exchange goods with local consumers and with the rest of the world. All rural areas produce (only) agricultural goods. The trade hub hosts a non-agricultural sector and serves as an embarkation point for all exports. In our empirical work, the trade hub will correspond to an aggregate of the eighteen largest exporting ports in Brazil.[4]

[3] Appendix C details the derivations of the model.

[4] In 2023, 88% of the Brazilian population lived in urban areas, according to the World Development Indicators, making rural non-agricultural consumption a relatively small fraction of the total. Our assumption – that all consumption occurs at the trade hub – allows us to simplify the consumption side of our model and to avoid modelling the migration frictions that would arise if consumers faced different prices in different locations. Other models using a similar trade hub setting include Adamopoulos (2025) who studies

Each rural location has a total land endowment $\bar{L}_i$ and is composed by a continuum of plots ω of size one, whose set is denoted Ω_i. Those plots can be left under undisturbed forest $F_i(\omega)$ or can possibly be used for agriculture $L_i(\omega)$ to produce several crops $k \in \mathcal{K}$. Total land in region i is the sum of cultivated plots plus those left under forest:

$$\bar{L}_i = \int_{\Omega_i} \left(\sum_{k \in \mathcal{K}} L_{ik}(\omega) + F_i(\omega) \right) d\omega. \tag{1}$$

There are three types of economic agent in this economy. First, a representative farmer chooses the profit-maximizing factors for agricultural production in each rural location. This output is then shipped towards the trade hub, where it can be either locally consumed or exported. Second, the economy features a representative consumer who owns and rents out land and who supplies labour inelastically. Third, there are non-agricultural firms that hire labour to produce a composite of manufactured goods and services.

Consumers and non-agricultural firms are located in the trade hub. Trade of agricultural products from a region i to the trade hub is subject to iceberg transport costs denoted by τ: for one unit of crop k to arrive from region i to the trade hub, $\tau_i \geq 1$ units must be shipped.

4.2 Production

4.2.1 Agricultural sector

In a region $i \in \mathcal{L}$, let production of a crop k over plot ω be given by:

$$Q_{ik}(\omega) = (N_{ik}(\omega))^{\alpha_k} (h_{ik}(\omega))^{\beta_k} (A_{ik}(\omega) L_{ik}(\omega))^{\gamma_k}, \tag{2}$$

where output depends on labour $N_{ik}(\omega)$, intermediate input $h_{ik}(\omega)$, land productivity $A_{ik}(\omega)$, and $L_{ik}(\omega)$, which is the share of plot ω allocated to crop k. Our parameters of factor-intensity, α_k, β_k and γ_k, vary across crops and are constrained such that $\alpha_k + \beta_k + \gamma_k = 1$ for any $k \in \mathcal{K} = 1, \ldots, K$. The parameter γ_k is equal to $(1 - \alpha_k - \beta_k)$ and thus defines each crop's land intensity. Labour is paid a wage w_i, and intermediate inputs are entirely imported from abroad, with their local price denoted v_i. This implies that trade frictions inside the country are affecting the economy in two distinct ways: through the sale of outputs (to cities and the rest of the world) and through importing inputs.

agricultural trade in Ethiopia, or Fajgelbaum and Redding (2022) who focus on structural transformation in Argentina.

To "open" a plot for productive use, a representative farmer must pay a fixed-cost $A_{i0}(\omega)$ expressed in units of non-agricultural good.[5] We view $A_{i0}(\omega)$ as an investment, such as the supplementary labour cost required to cut down standing forest (e.g., slash and burning activities).

In each rural location, the representative farmer chooses a crop mix and inputs (including land) to maximize profits. This gives rise to the necessary condition that the price of a commodity k equals its marginal cost of production, implying that the *net* plot-specific rental rate can be written as:

$$r_i(\omega) = \psi_{ik} A_{ik}(\omega) - p_0 A_{i0}(\omega), \tag{3}$$

where:

$$\psi_{ik} = \left(\tilde{\alpha} \frac{p_{ik}}{w_i^{\alpha_k} \upsilon_i^{\beta_k}} \right)^{\frac{1}{\gamma_k}},$$

is the profitability index of crop k and where $\tilde{\alpha} = \alpha_k^{\alpha_k} \beta_k^{\beta_k} \gamma_k^{\gamma_k}$.

In this formulation, equation (3) expresses the rental rate $r_i(\omega)$ as a function of land productivity $A_{ik}(\omega)$ and crop profitability ψ_{ik}, which is itself increasing in commodity prices p_{ik} and decreasing in the wage rate w_i and the intermediate input cost υ_i, for any region $i \in \mathcal{L}$. The last term on the RHS of (3) is the value of the land-clearing investment, or fixed conversion cost, $p_0 A_{i0}(\omega)$. Here, p_0 designates the price of the outside good, which is also our numéraire. Whenever the cost $p_0 A_{i0}(\omega)$ is higher than the *gross* rental rate for all crops k, $A_{ik}(\omega)\psi_{ik}$, a given plot ω is left under forest.

Similar to Farrokhi and Pellegrina (2023), the vector of land productivities and investment parameter, $A_i(\omega) \equiv \{A_{ik}(\omega)\, \forall k \in \mathcal{K}, A_{i0}(\omega)\}$, is randomly distributed across plots $\omega \in i$ from a nested Fréchet with parameters $\{Z_{ik}, Z_{i0}, \theta_1, \theta_2\}$, that is:

$$\Pr\left(A_i(\omega) < A_i\right) = \exp\left\{ -\tilde{\gamma} \left[\left(\frac{A_{i0}}{Z_{i0}} \right)^{-\theta_1} + \left(\left(\sum_{k \in \mathcal{K}} \left(\frac{A_{ik}}{Z_{ik}} \right)^{-\theta_2} \right)^{-\frac{1}{\theta_2}} \right)^{-\theta_1} \right] \right\}, \tag{4}$$

[5] The model of Costinot et al. (2016) displays a similar feature with its "labor intensity" requirement. Fajgelbaum and Redding (2022) also propose a fixed-cost approach in an extension of their spatial model. In the paper of Farrokhi and Pellegrina (2023), it is an "investment requirement".

where $\tilde{\gamma} \equiv \left[\Gamma\left(1 - \frac{1}{\theta_1}\right)\right]^{-\theta_1}$ is a normalization of the Gamma function.[6] The parameter $Z_{i0} > 0$ is the region-specific investment intensity required to open the land for agriculture, while Z_{ik} measures land productivity for a given crop k.

The parameter θ_1 drives the substitution between agricultural and forest land use. In the lower nest, θ_2 controls the dispersion of productivity draws across crops within each region. When θ_2 is large, land is more homogeneous. The case where $\theta_1/\theta_2 = 1$ corresponds to the standard one-nest Fréchet distribution, where the substitution elasticities among crops and between forest and crops are the same.

Land Use. The representative farmer chooses the crop k which maximizes the land rents from (3). Deforestation in plot ω occurs whenever this rent is positive. The solution to this problem yields land shares S_{ik} for any crop k that are the product of two components. The first component is the share of land under agriculture in location i:

$$\pi_i^A = \frac{\Phi_i^{\theta_1}}{(p_0 Z_{i0})^{\theta_1} + \Phi_i^{\theta_1}}, \tag{5}$$

where we have:

$$\Phi_i = \left(\sum_{k' \in \mathcal{K}} (Z_{ik'}\psi_{ik'})^{\theta_2}\right)^{\frac{1}{\theta_2}}. \tag{6}$$

The second component is the share of agricultural land allocated to crop k in region i:

$$\pi_{ik} = \frac{(Z_{ik}\psi_{ik})^{\theta_2}}{\Phi_i^{\theta_2}}. \tag{7}$$

Equation (5) shows that the share of land allocated to agriculture in region i decreases with the fixed conversion cost, $p_0 Z_{i0}$, and increases with aggregate agricultural profitability, measured by Φ_i. As detailed by equation (6), for each crop k the latter has two components: land productivity Z_{ik}, which is given and exogenous, and a crop profitability index ψ_{ik} that depends on the price of crop k and input costs w_i and υ_i. The conversion cost, $p_0 Z_{i0}$, matters for the extensive margin: the higher the cost of converting forest to cropland, the lower is the share of land attributed to any crop $k \in \mathcal{K}$.

Equation (7) shows that the share of land allocated to crop k within open plots increases with the crop profitability index, ψ_{ik}, and also with aggregate land-productivity, Z_{ik}, relative

[6] Our formulation differs from that of Farrokhi and Pellegrina (2023) as in their paper, land productivities vary across another dimension: agricultural technologies. Here, we adopt this approach to capture the differences in land-use substitution between forest and crops or between crops. In that sense, we follow Dominguez-Iino (2021).

to all crops $k' \in \mathcal{K}$. This implies that land is generally used where its *comparative* advantage is highest, rather than where its *absolute* advantage is strongest.

The parameters θ_1 and θ_2 are key to understanding changes in land use as a response to new prices. Given the definition of Φ_i and equation (5), the ratio θ_1/θ_2 governs the impact of changes in crop prices $p_{ik}\,\forall k \in \mathcal{K}$ on the allocation of land between forest and agriculture. The parameter θ_2 governs the distribution of productivity across plots, such that a larger θ_2 implies less heterogeneity in land productivity within region i. It is also important for the elasticity of substitution between crops. A higher θ_2 will make the model more sensitive to changes in crop prices, and will lead to a more pronounced reallocation of land across crops. This is the intensive margin of land use.

Finally, note that the amount of land covered by forests at the regional level is the total area multiplied by the share of land not allocated to any crop k:

$$F_i = \left(1 - \pi_i^A\right) \bar{L}_i. \tag{8}$$

To anticipate the role played by transport costs, it is worth examining the numerator of equation (5). The term Φ_i can be viewed here as the relative profitability of agriculture in region i. It is equal to the sum, for each crop k, of the following term:

$$\left[Z_{ik} \left(\tilde{\alpha} \frac{p_{ik}}{w_i^{\alpha_k} \upsilon_i^{\beta_k}} \right)^{\frac{1}{\gamma_k}} \right]^{\theta_2}. \tag{9}$$

Let us focus on two important components directly linked to trade frictions: p_{ik} and υ_i. The spatial equilibrium in the economy is such that the price of a crop k at the trade hub is equal to the farm-gate price adjusted by transport costs; i.e., $p_k = \tau_i p_{ik}$. When a region is poorly connected to the trade hub, thus facing a high iceberg cost, the farm-gate price p_{ik} will be lower. It follows that, all things being equal, the amount of land dedicated to any crop k in that region will be lower.

Another way to appreciate the effect of iceberg costs is to examine the role of υ_i, the cost of the intermediate input. Since this input is entirely imported from the trade hub, υ_i is equal to the world price of the intermediate input times the cost τ_i of shipping it to region i. Again, keeping all other parameters constant, when a region i is isolated and τ_i is high, the cost υ_i of input use $h_{ik}\,\forall k$ is higher. As a consequence, cultivating any crop k becomes less profitable.

4.2.2 Non-Agricultural Sector

Production in the non-agricultural sector takes place in the trade hub and is based on labour only. Technology is given by $Q^M = \Lambda \left(N^M\right)^{\alpha^M}$, with Λ representing a productivity parameter and $0 < \alpha^M < 1$ is labour intensity. Firms maximize profits denoted Π^M, and labour is paid at the marginal value of its output, such that we have:

$$p_0 \alpha^M \frac{Q^M}{N^M} = w, \tag{10}$$

where the wage rate w also pins down the labour cost w_i for all regions i.

4.3 Consumer's problem

Total expenditures in the economy are given by the sum of wages in both sectors, rental income from the land, and manufacturing profits. A constant share $b^A \in (0, 1)$ of income is spent on a composite agricultural good. From the consumer's perspective, individual crops are imperfect substitutes, and the composite is aggregated from the consumption of each crop according to the expression:

$$C^A = \left(\sum_{k=1}^{K} C_k^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}.$$

In this formulation, $\sigma > 0$ represents the elasticity of substitution between crops. The remaining share of the consumer's income is spent on the non-agricultural good and is denoted C^M.

Provided that $p_k = \tau_i p_{ik}$, the value of optimal consumption of crop k is given by:

$$p_k C_k = \left(\frac{p_k}{P}\right)^{1-\sigma} b^A E, \tag{11}$$

with $P = \left(\sum_{k \in \mathcal{K}} p_k^{1-\sigma}\right)^{\frac{1}{1-\sigma}}$ the price index in the trade hub.[7] By definition we have $b^A E = \sum_{k \in \mathcal{K}} p_k C_k$ and the consumption of manufacturing goods is simply $C^M = (1 - b^A)E/p_0$.

4.4 Competitive equilibrium

A competitive equilibrium consists of:

[7] In our framework, the total demand for crop k consists of both domestic and foreign demand, which we denote X_k, at the trade hub. In our quantitative analysis, we take X_k from the data as we detail in section 5.

(a) Prices p_k in the trade hub and p_{ik} in all regions $i \in \mathcal{L}$ for all crops k, and price p_0 for the non-agricultural good ;

(b) Wage rate $w = w_i$ for labour employed in both sectors ; price v_i for the use of intermediate inputs imported from abroad in all regions $i \in \mathcal{L}$;

(c) Expenditure, $E = \sum_i w_i N_i^A + wN^M + \sum_i R_i + \Pi^M$, in the trade hub ;

(d) Demands for labour N_{ik} and for intermediate input use h_{ik} for all crops $k \in \mathcal{K}$ produced in region $i \in \mathcal{L}$, and demand N^M for labour in manufacturing ;

(e) Trade flows for all crops $k \in \mathcal{K}$, with imports M_k and exports X_k, and imports of intermediate inputs h_{ik};

(f) Markets clear for labour and goods for all crops $k \in \mathcal{K}$ and locations $i \in \mathcal{L}$, such that:

$$\begin{aligned} \bar{N} &= \textstyle\sum_i \sum_{k \in K} N_{ik} + N^M; \\ C_k &= \textstyle\sum_i Q_{ik}/\tau_i + M_k - X_k; \end{aligned} \tag{12}$$

Conditions (a) to (f) define the spatial equilibrium in the economy. The market for the manufacturing good clears by Walras' Law. The equilibrium features a land allocation in each location between crops and forests. The equilibrium also implies that consumers buy agricultural goods from the cheapest supplier and that trade with the rest of the world is balanced by the definition of E_i.

5 Taking the model to the data

In this section, we detail our approach to bringing the model to the data. Our model includes pasture together with the following eight major crops: beans, cassava, coffee, corn, rice, soybeans, sugarcane, and wheat. Below, we describe the calibration of the model's parameters, the estimation of the transport cost elasticity and the model fit.

5.1 Calibration

To calibrate our factor cost shares α_k, β_k and γ_k for every crop k, we use the parameters from Pellegrina (2022). The author uses data on payments to labour and revenues for more than ten crops in Brazil. To give one example, we obtain $\alpha_{soy} = 0.06$ and $\beta_{soy} = 0.44$ for soybeans.[8]

In our numerical exercise, we calibrate θ_1 and θ_2 to values used in recent literature. We set our parameter of land supply elasticity, θ_1, to 1.4 (Dominguez-Iino, 2021). From Sotelo

[8]To obtain values for α_k, we combine variable and non-variable labour in Pellegrina's estimates.

(2020), we set θ_2, the elasticity of substitution between crops within agricultural land, to 1.6.

To obtain the fixed cost of land conversion, Z_{i0}, we invert the model by exploiting equation (5) and Mapbiomas data on forest cover. This allows us to match perfectly forest endowments by meso-region at baseline.

To calibrate the productivity parameters Z_{ik} for all crops and meso-regions in our model, we use the FAO-GAEZ data on potential yields under high inputs and rain-fed irrigation. Following Porteous (2024), we account for potential measurement errors or multiple harvests by considering a region-level scaling factor where we have $Z_{ik} = \bar{z}_{ik}\tilde{Z}_{ik}^{\{\text{GAEZ}\}}$. The parameter $\tilde{Z}_{ik}^{\{\text{GAEZ}\}}$ represents the average value of potential yields for crop k in region i, and to obtain Z_{ik} we calibrate $\bar{z}_{ik}$ by inverting the equation of crop supply given by:

$$Q_{ik} = \frac{1}{\gamma_k}\left(\frac{\psi_{ik}}{p_{ik}}\right)\bar{L}_i Z_{ik}\left(\pi_i^A\right)^{\frac{\theta_1-1}{\theta_1}}\left(\pi_{ik}\right)^{\frac{\theta_2-1}{\theta_2}}. \tag{13}$$

To invert (13), we use data on crop output, total agricultural land share π_i^A and crop shares π_{ik} from the 2017 agricultural census. By doing this, we match exactly the quantity produced at baseline that year.

Productivity in the outside sector, Λ, is calibrated by exploiting the labour demand condition in that sector as well as data on labour employment in 2017. We set the income share α^M to 0.7, following Gollin (2002).

The demand elasticity σ, reflecting substitutability between crops, is calibrated according to the estimates of Sotelo (2020) and takes the value 2.4. For international trade, we use national data from FAOSTAT and calibrate imports M_k and exports X_k of crop k to match the imported and exported shares of national output in 2017. To capture the reaction of foreign demand to shocks like new roads, as well as to account for the fact that Brazil is a major exporter of certain crops on the world market, we assume a constant price elasticity. This way, demand coming from foreign countries is downward-sloping. Specifically, we choose the following functional form:

$$X_k = X_{k_0}\left(\frac{p_k}{p_{k_0}}\right)^{-\sigma},$$

where p_{k_0} and X_{k_0} are respectively the trade-hub price and export value of a crop k in the baseline economy. The same approach is adopted for imports, and we have $M_k = M_{k_0}\left(p_k/p_{k_0}\right)^{\sigma}$.[9]

[9]Note that imports also exhibit a constant price elasticity with downward-sloping demand. The reason is that in our trade hub setting, the price of imports is the same as the country/world price. Consumers are thus indifferent between domestic and imported goods.

5.2 Transport costs estimation

As we do not observe all trade flows within Brazil, we estimate the transport cost wedges by using the respective costs of shipping soybeans from each meso-region to the closest international port. This commodity was chosen since it is largely exported, following the method of Donaldson (2018) or Sotelo (2020). Formally, we estimate the following model:

$$\log\left(\frac{p_P^{\{\text{soy}\}}}{p_i^{\{\text{soy}\}}} - 1\right) = \beta_0 + \beta_1 \{\text{TC}\}_{iP} + \beta_2 X_i + \epsilon_i, \tag{14}$$

where the dependent variable is the log of the soybean price difference between the closest exporting port P (p_P^{soy}) and location i (p_i^{soy}). Prices are computed from the 2017 agricultural census (IBGE).[10] Our variable of interest on the RHS is the transportation cost $\{\text{TC}\}$ between location i and the closest international port P. The variable X_i is a set of controls which includes temperature, precipitation, and land productivity. The error term is given by ϵ_i.

Equation (14) is estimated at the district level (*municipios*), rather than meso-regions, for two reasons. First, it allows us to increase the size of our sample and thus increase the precision of our estimates. Indeed, estimating equation (14) at the meso-region level leaves us with 81 observations, compared to over 2,150 at the district level. Second, since we are working on the distance from centroids to ports, the smaller the size of our administrative level, the less likely our empirical analysis is to face measurement error.

Different indexes can be used for our variable of transportation cost between a district's centroid and the closest international port. One could use shortest-path distances, measured in kilometres, or travel time, based on real road infrastructure. To bring additional precision, we instead use data provided by de Castro Victoria et al. (2021) on the cumulative cost of moving from each grid cell to the closest exporting port. This index of "Relative Distance" is also computed by using the shortest path through the existing road network, but it accounts for different road surface types. These types include absence of road, unpaved or paved roads, which can be themselves divided into several sub-types. For instance, among paved roads, the data allow us to distinguish between single- and double-lane roads, as well as roads with a second lane under construction. Furthermore, this variable takes into account

[10] Precisely, we extract it from PAM (*Producao Agricola Municipal*) data, available at https://sidra.ibge.gov.br/pesquisa/pam. To compute p_P^{soy}, we average p_i^{soy} across municipalities i that both have a port and produce soy.

the activity of each port, *i.e.* whether agricultural commodities are exported or not over the period of study.[11]

Estimating equation (14) using OLS may pose a threat to identification. It is possible – indeed, almost certain – that road placement is not random. Road locations may be highly endogenous to patterns of agricultural land use, along with other factors, such as political considerations (Souza-Rodrigues, 2019; Morten and Oliveira, 2024) or policy interventions (Pfaff and Robalino, 2017). To address the potential endogeneity of our index of relative distance, we adopt an approach similar to that of Souza-Rodrigues (2019): we construct a network linking municipality centroids to the nearest international port and state capital using straight-line distances. This network serves as an instrument for our cumulative cost variable, and the model specified in equation (14) is then estimated using 2SLS.

The underlying assumptions are as follows. First, this straight-line network is expected to be correlated with transportation costs between municipalities and ports. Second, the delimitation and location of municipalities were determined long before the large wave of road expansion implemented in Brazil in the 1970s. For these reasons, using straight-line distances to ports and capitals provides us with a reasonable instrument for transportation costs.

Table 4 in Appendix A presents the results of the second stage of our 2SLS estimation. Our estimation yields a trade cost elasticity of 0.28 using OLS, significant at the 1% level. In the 2SLS model, the magnitude of the elasticity increases to approximately 0.34 (significant at the 1% level). This is within the range of estimates from the literature, as Pellegrina (2022) finds 0.13 for perishable agricultural products in Brazil, and Sotelo (2020) reports 0.47 for coffee in Peru. Our estimate suggests that a 10% increase in the cumulative cost of shipping goods from a district i centroid to the nearest export port increases the soybean price gap by 3.43%.[12] Equipped with the estimated elasticity, we then compute iceberg transport costs τ_i for each meso-region using infrastructure network data from 1995 and 2017.[13]

[11] This information is provided by Agrostat and the Ministry of Agriculture.

[12] Since many municipalities do not produce soybeans, we have just over 2,150 observations for this estimation (out of more than 5,000 Brazilian municipalities). As a robustness check, we also estimate the elasticity using other crops, both in the full sample and within the soy-producing subsample, and find qualitatively similar results. Given that our identification strategy, inspired by Donaldson (2018), relies on the commodity being largely exported, we consider soybeans the most appropriate choice for our baseline analysis.

[13] We use data on federal roads, accounting for access to active railroads and ports in both time periods to ensure that, in our counterfactuals, export flows are not attributed to infrastructure that did not yet exist.

5.3 Model Validation

In this section, we discuss the fit of the estimated model with different aspects of the data important for our study. Specifically, in Figure 7 of Appendix B, we display land uses and output per hectare in our baseline economy and in the data for the year 2017.

By construction of the fixed cost parameter Z_{i0}, the forest land area exactly matches the data. Regarding the land area of each crop, shown in the left panel of Figure 7, we obtain model fit with a R-squared of 0.88.[14]

The right panel presents the output per hectare for our nine agricultural goods. It is done for each meso-region in the data and for the model's predictions. We can see that those are reasonably well predicted across meso-regions. Regressing the log of the value in the data over that predicted by the model yields a R-squared of 0.36, with a slope of 0.64. Other moments we do not display here to save space include the dispersion in the production values of each crop, which our model explains at 91% at baseline.

6 Results

This section presents the results of our counterfactual exercises. The first scenario examines the impact of roads built since 1990. In the second scenario, we explore possible complementarities between infrastructure expansion and technical change. To do this, we analyse the impact of roads built since 1990 together with another major disruption that occurred more or less contemporaneously: the introduction of new soybean varieties.

6.1 Assessing historical reductions in transport costs

As a first counterfactual exercise, we are interested in evaluating the impact of past infrastructure improvements. To do this, we focus on roads that were built from the 1990s up to our baseline year (2017). Using historical road maps, we first calculate the iceberg transport costs that prevailed in 1995 and examine the model's predictions for both land use and a set of development outcomes. The year 1995 was chosen since it was an agricultural census year, for which we have rich data provided by IBGE. This allows us to compare the model's predictions to contemporaneous data.

To start, Figure 3 shows the change in iceberg transport costs across locations at the meso-region level when roads built since the 1990s are removed.

[14]In Figure 8 of Appendix B, we further display the model fit in terms of land share to exclude scale differences. For crops, we obtained an R-squared value of 0.75 with a slope of 0.78. Forest shares match the data perfectly, by construction.

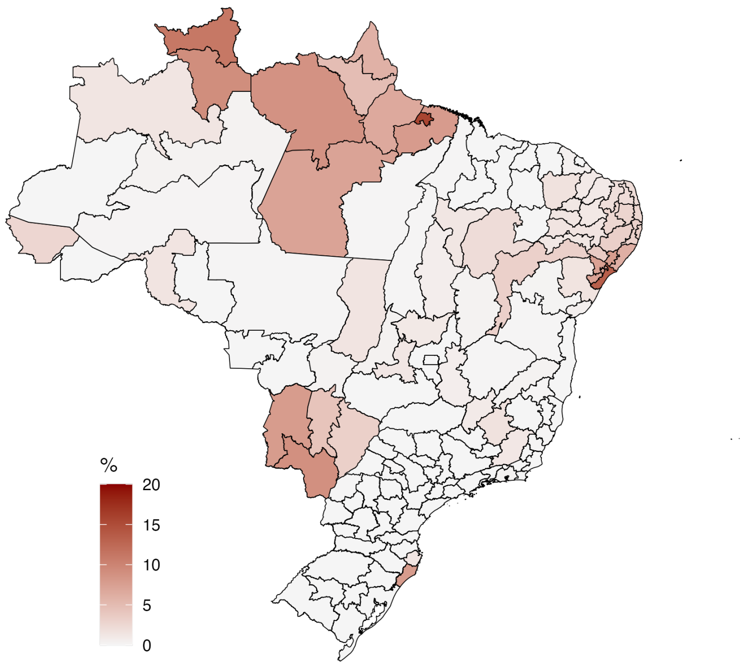


Figure 3: Variation in iceberg transport costs between 1995 and 2017 across the Brazilian meso-regions

The map shows that the meso-regions that experienced the largest reductions in transport costs over that period were mainly located in the North and Central-West regions. For instance, according to our estimates, in the meso-region of *Nordeste Paraense*, in the state of *Pará*, iceberg transport costs dropped by 8% over the period as a result of road construction. By contrast, in some other locations, we observe no change.

Let us now assess the effect of these shocks by introducing 1995 transport costs to our baseline economy as a first counterfactual exercise. We first summarize the aggregated outcomes and then turn to the spatially detailed findings.

Aggregate results. Table 2 displays the results when transport cost variations affect both inputs and output (i.e., τ_i and υ_i).

Compared with the baseline economy, we find that the total amount of forest cover would be 0.59% higher than its actual 2017 level if transport costs had remained at their 1995 level. As we detail below, this aggregate level conceals very large heterogeneity across meso-regions. For instance, some meso-regions would have as much as 70% higher forest cover in 2017 if they had remained as remote as they were in 1995. On average, across

Table 2: The economy with past transport costs: predicted rates of change

		Counterfactual
Forest Cover	$\Delta\left(\sum_i F_i\right)$	0.59
Ag. Income	$\Delta\left(\sum_i \sum_k p_{ik} Q_{ik}\right)$	-0.43
Q1 income	-	0.10
Q4 income	-	-8.08
Ag. exports	$\Delta\left(\sum_k X_k\right)$	-2.28
Input adoption	$\Delta\left(\sum_i h_i\right)$	-1.16

Notes: This table reports the predicted changes in several outcomes relative to the baseline economy. Q1 and Q4 designate changes in income in the richest and poorest (respectively) quartiles of meso-regions.

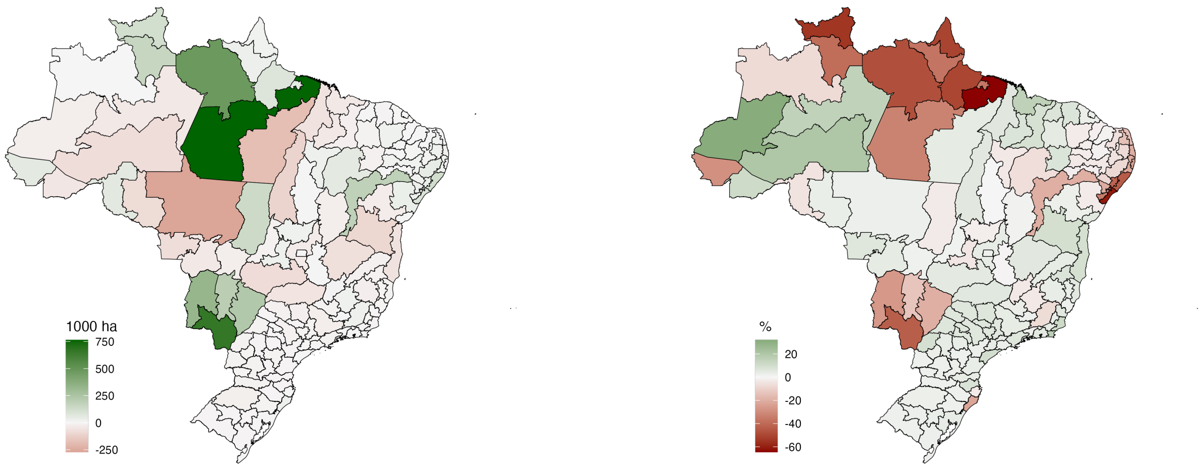
meso-regions, we estimate that the reduction in transport costs led to a 1.8% deforestation rate, which accounts for a substantial share of the actual average variation in the data.

In the counterfactual economy, agricultural exports would have been 2.3% lower. This decline in exports reflects the negative price shock faced by the economy as a whole, and particularly by more remote meso-regions. In comparison with the baseline, in the counterfactual economy, some locations would have become less attractive for agriculture, as both τ_i and υ_i are affected by transport costs. Locations remote from transportation hubs would have experienced substantial increases in their production costs, alongside decreases in farm-gate prices. Put differently, the actual road investments during this period had a major impact on increasing agricultural output – but did so at a significant cost in terms of forest loss.

The model predicts that the decline in transport costs during this period caused nearly one tenth of the total loss observed in the data over 1995-2017. This assumes that the observed allocation of land in 2017 fully incorporates all the land use changes associated with the expansion of the transport network and thus reflects the new spatial equilibrium. To the extent that land conversion may take some time, it is possible that transport expansion from the past few years may not yet be fully embedded in observed land use. In that sense, our model gives us a lower-bound prediction of the deforestation impacts. Let us now examine other economic outcomes in the counterfactual economy.

Had transport costs remained at their 1995 levels, aggregate agricultural income would be 0.43 percent lower than the 2017 baseline. This average result hides important disparities. Indeed, the meso-regions with the highest agricultural income (the top quartile) would be 0.10% better off in terms of income, while the quartile with the lowest income would be worse off by 8%. This suggests that road construction since the 1990s has reduced spatial inequalities in terms of agricultural income.[15] We now go further and disentangle the spatial disparities of these results.

Regional results. Figure 4 presents the changes in forests and in agricultural income predicted by the model when we return the model economy to 1995 iceberg costs.



(a) Predicted Change in Forest Cover. (b) Predicted Change in agricultural income.

Figure 4: Spatial impacts of reversion to 1995 levels of transport costs. Note: Figure 4a plots predicted change in hectares of forest per meso-region with 1995 transport costs compared to the baseline equilibrium. Figure 4b for the percentage change in income.

Figure 4a shows the spatially heterogeneous distribution of deforestation across Brazilian meso-regions. It is important to keep in mind that the figure shows what would happen relative to the baseline 2017 economy when historical transport costs (from 1995) are re-introduced. Some locations exhibit a high rate of forest gain, while some others actually lose forest stock. In the latter case, this means that new roads from 1995 to 2017 appear to have led to reforestation in some regions.

Consider first a location where road construction led to forest clearing. In the meso-region of *Nordeste Mato-grossense*, in the Central-West part of the country, we find that an extra 138,000 hectares of forest would have been preserved in 2017 had Brazil's transport network stayed as it was in 1995. In the state of Pará, in the meso-region of *Nordeste Paraense*, the gains reach 760,000 hectares and explain almost 70% of what we observe in the data:

[15]The income quartiles are computed using data on the value of agricultural production at the meso-region level in 1995, before road expansion.

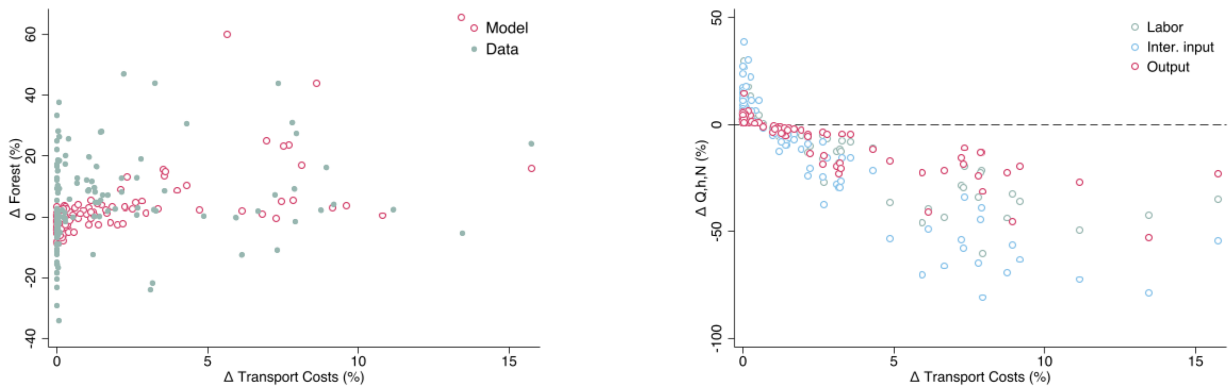
the model tells us that changes in the road network accounted for the majority of the forest clearing that took place in this meso-region.

Figure 4b allows us to discuss further our results about the impacts on the top and bottom quartiles of meso-regions, in terms of agricultural income distribution.[16] As shown by the figure, compared to the baseline, many locations in the economy would have experienced lower agricultural incomes in 2017 if iceberg costs had remained at their 1995 levels. Put differently, this means that new roads have contributed to an increase in agricultural income in some meso-regions, where a change in agricultural activities has taken place due to spatial changes in prices. On average across all meso-regions, we find that agricultural income increased by 3.8% as a result of the decline in transport costs.

At the economy-wide level, however, the changes in agricultural income due to roads appear to have been relatively modest. Some regions gained and others lost, but we find that gross agricultural revenues, $\sum_i \sum_k p_{ik} Q_{ik}$, only grew by 0.43% as a result of infrastructure expansion. The gains in remote regions were offset to some degree by losses in other areas that were undercut by the expansion of agricultural production into previously forested areas. The net gains in agricultural income are thus substantially smaller than the gains in the regions that were most favoured. And the welfare gains would still be smaller, especially if measured against the local and global environmental costs.

To better grasp the consequences of improvements in market access, Figure 5 plots the predicted change in forest cover in the model and in the data, as well as the reallocation of output and factors.

[16]Recall that each region in the model has a single representative consumer. Thus, the only distributional issue we can address is the distribution across meso-regions, i.e., the spatial inequality of our economy. Recall also that all consumption takes place in the (urban) transportation hub, so all consumers face the same prices, regardless of the region in which they produce output. Our measure is thus simultaneously a measure of inequality in agricultural income and a measure of inequality in consumption.



(a) Changes in transport costs and the forest

(b) Output and resources reallocation

Figure 5: Changes in forest cover and production patterns across meso-regions. Note: Figure 5a plots the predicted rate of change in forests per meso-region with 1995 transport costs compared to the baseline equilibrium, and Figure 5b the reallocation of output and inputs use.

Variations in land use and in transport costs seem to be correlated, both in the data and in the model. Since there are many other factors taking place and potentially affecting the spatial patterns of forest changes (e.g., logging, mining and other forms of extraction mediated through varying degrees of local regulatory effectiveness), we do not expect our model to explain all the variation observed in the data. Keeping this in mind, Figure 5a shows the linear relationship between the removal of trade frictions and land expansion from 1995 to 2017. The broad pattern is clear: the more frictions diminish, the more land is converted into cropland.

In the model, changes in forest cover are mainly linked to soy and pasture production, as shown by Figure 9 in Appendix B. In some meso-regions, the changes in iceberg transport costs have led to increases of over 30% in the land area devoted to soybeans. Some locations in Northern Brazil have nearly 25% more land under pasture with 2017 transport costs than would have been the case had transport costs remained at their 1995 levels. While the changing spatial patterns of agricultural production do highlight the “positive” impacts of specialization, the clear cost, in this case, is that livestock production has been moving to previously forested regions in Northern parts of Brazil.

Note that these changes do not affect only those regions that experienced direct changes in their own transport costs. Because of spatial spillovers and general equilibrium effects, there appear to have been substantial changes even in locations that had little change in transport costs.

New roads thus threaten forests also in locations indirectly affected by road construction – and in some cases locations that are quite distant from new roads. The impact of roads is felt not only in close proximity, but in locations that produce similar commodities and that supply them to the same urban markets. Note that we would not be able to take into account these effects using a standard empirical analytical framework. Among the advantages of our framework and approach is that we can assess these general equilibrium effects at the crop-specific level. This proves to be important for studying the effects of a policy like infrastructure improvements, where outcomes are likely to spill into non-treated locations.

Another way to examine this is through Figure 5b, which shows the relationship between changes in iceberg transport costs and the reallocation of output and inputs. The figure suggests that the historical decrease in transport costs led to a reallocation of resources – labour and intermediate inputs – across meso-regions towards those that were newly connected. This is consistent with the idea that new roads have driven a spatial reorganization of production in Brazil, shifting from regions with high transport costs to regions with lower transport costs.[17] Interestingly, Figure 5b also shows that some regions experienced little to no change in transport costs but still saw variations in agricultural output and inputs use. These effects arise from general equilibrium adjustments, where roads redefine comparative advantages across space, leading to a different allocation of resources.

6.2 Complementarities: new soybean varieties and market access

In this section, we turn to our second counterfactual exercise, where we explore whether the adoption of new soybean varieties has amplified or mitigated the impact of roads on deforestation. The key innovation came in 2003, when Brazil allowed the introduction of new genetically modified herbicide-resistant soybean seeds, which facilitated no-tillage and low-tillage planting techniques (Bustos et al., 2016).

In our counterfactual analysis, we use the FAO-GAEZ data on potential yields, specifically the variation in predicted yields under low and high inputs scenarios, as in Bustos et al. (2016). In this counterfactual, we not only remove roads but also assume that soy in our model is produced with low-input productivity. This is achieved by maintaining the same scale parameter, $\bar{z}_{i\{\text{soy}\}}$, as calibrated in section 5, but multiplying it by the potential yields under low inputs.

[17] Note that in this framework, labour is mobile across regions, in order to capture the full equilibrium effects of new roads. In Table 6 of Appendix E, we present results with a version of the model where agricultural labour is immobile. The impact of roads on forest cover is more limited in this case due to the absence of labour reallocation towards newly connected regions. This can be viewed as a lower bound for the effect of new transportation infrastructures.

To save space, the aggregate results of this counterfactual exercise are presented in Table 5 of Appendix A. Two scenarios are explored. First, we consider the impact of soy productivity changes by themselves. Second, we look at the joint effect of changes in soy productivity and transport costs by setting both to their 1995 levels. The difference between these two scenarios allows us to isolate the complementary effect of roads and improved soy productivity.

We note that the counterfactual suggests that the combined effect of changes in iceberg costs and soy productivity has been a 3.48% reduction in total forest cover; i.e., less forest cover in 2017 than would have been the case if iceberg costs and soy technology had remained at their 1995 levels. According to the model's results, the two changes together account for 40% of the observed deforestation in the data (second column of Table 5). Agricultural income and exports are lowered by 6.48% and 14.51%, respectively, when the economy is set to 1995 values for transport costs and soy productivity. This reflects the importance of soybean production for the country's exports.

Let us now focus on isolating our variable of interest: the impact of new roads. Figure 6 displays the difference between *the impact of roads only* in our first and second counterfactuals. For clarity in the exposition, all values were multiplied by minus one.

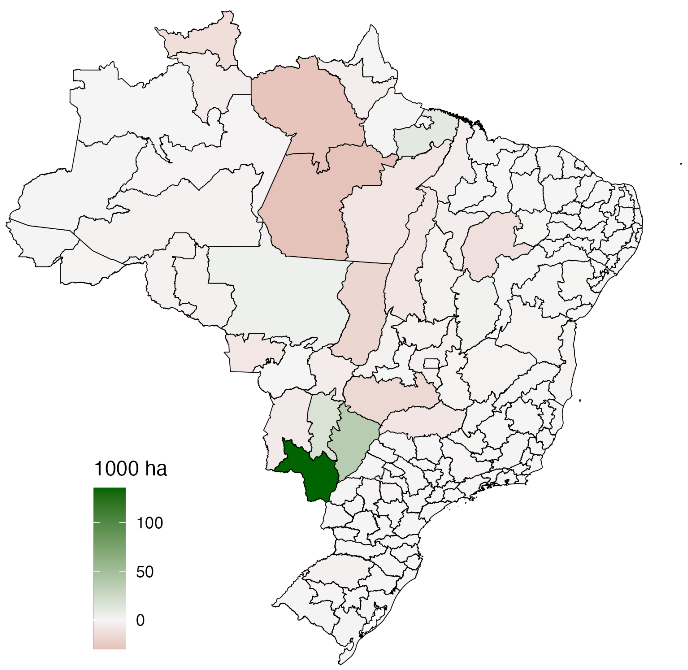


Figure 6: Variation in the net effect of roads with and without soybeans innovation

Negative values denote locations where new roads combined with new soybean seeds to bring about more deforestation than new roads alone – or, equivalently, where reforestation was less extensive with soybean innovation. In locations with positive values, the forest loss due to new roads was lower with high-input soybeans than in their absence. Although both outcomes are possible, the map shows that the introduction of new soybean seeds has tended to amplify the impact of roads on deforestation in most cases.

In the meso-regions of the Center-West and some in the North of the country, particularly in the states of Mato Grosso and Pará, modern soybean seeds amplified the effect of roads on deforestation. In the meso-region of *Nordeste Mato-Grossense*, reductions in transport costs combined with soybeans innovation led to an additional 20,000 hectares of forest clearing, relative to what would have happened with new transport infrastructure only. This represents almost a 15% increase in that location relative to the effect of roads estimated under low-input soybeans. This is because the new seeds made it more profitable to convert land into agriculture, especially in regions with a comparative advantage in soybeans (see Figure 2).[18]

Interestingly, in some regions, the innovation in soybean seeds reduced the impact of roads on deforestation. In *Sudoeste de Mato Grosso do Sul*, in the southwestern part of the country, the model predicts that while the total effect of new transportation infrastructure was negative (Figure 4), the introduction of new soybean seeds lowered the total amount of forest clearing attributed to roads by approximatively 135,000 hectares.

Overall, this exercise highlights the importance of considering the complementarities between different factors when assessing the impact of new roads on deforestation. The introduction of new soybean seeds has amplified the impact of roads on deforestation in many regions and has mitigated it in some others. This suggests that infrastructure investments and technical change can interact in ways that affect conservation differently.

Estimated carbon cost. Having discussed how new roads impact land use and economic outcomes in sections 6.1 and 6.2, we now briefly extend our analysis to estimate the carbon cost of these changes. A simple calculation can compare the income gains with the carbon price of lost CO_2 sequestration in Brazil's forests.

[18] Also, in our current model, we assume perfect competition. Intuitively, if local oligopsonists were present, as in the work of Dominguez-Iino (2021), they might exert downward pressure on prices paid to local suppliers, which could reduce incentives for agricultural expansion and thus lower deforestation rates in some regions. However, the overall effect would likely depend on the spatial concentration of oligopsonistic firms and their degree of market power. In regions where oligopsonies dominate, one might expect lower input costs but also lower income redistribution effects due to concentrated market power.

Consider the first counterfactual exercise. According to the model, reductions in transport costs since the 1990s caused slightly more than 3.6 million hectares of forest losses. In 2022, the carbon price in a geographically close country, such as Colombia, was $5.01 per ton, according to the Carbon Pricing dashboard from the World Bank. Even on the basis of this *particularly* low carbon price, forest losses represent a huge carbon cost. For example, Paniagua-Ramirez et al. (2021) recently estimated that carbon storage in tropical forests varies from 204 tonnes per hectare (secondary forest) to 447 tonnes per hectare (mature forest). This implies that, with a carbon price as low as the one provided by the World Bank for Colombia, deforestation in Brazil caused by new transport infrastructure since the 1990s incurred a total cost between 3.7bn and 8.1bn of US$.[19]

When one considers the carbon price used by the previous US administration of US $51 (Rennert et al., 2022), this yields an upper bound to the deforestation cost of new roads since the 1990s of over 82.7bn of US$. We also note that our calculation does not account for other costs, such as biodiversity losses, existential threats, or the displacement of local communities – costs that are arguably even larger than those we have currently estimated.

This quantification is simply meant to provide a sense of the total costs associated with new roads. Yet, it is important to discuss two points regarding the costs and benefits presented here. The first one is on the temporality. Here, gains are measured using a static framework, which possibly misses the cumulative long-term effects of new roads. The second one is on geography. The carbon cost associated with deforestation is not confined to the boundaries where the deforestation occurs. Carbon emissions from lost CO_2 sequestration have global consequences, which affect climate systems worldwide. The implications of this are substantial. By contrast, the local income gains from road expansion and subsequent land use changes are confined to Brazil. This illustrates the disparity between local economic incentives and global environmental costs originating from externalities.

7 Conclusion

In this paper, we have studied how new roads may disrupt the spatial equilibrium of an agricultural economy and affect both environmental and development outcomes. To do so, we developed a quantitative spatial model of agricultural trade that incorporates multiple

[19]Our back-of-the-envelope calculation ignores potentially important spatial variation in the quantities of carbon that are sequestered in forests within Brazil. Since the term "forest" can be used to characterize a number of different biomes, these spatial differences can be significant. Our model does not give sufficiently fine-grained predictions to identify which locations within meso-regions are subject to forest clearing, so we prefer to use a rough average figure for carbon sequestration in a unit of forest. Taken across the whole country, this should offer a reasonable quantification, subject to the obvious limitations.

crops, trade frictions, and fixed costs for converting forests into cropland. We then took our model to the data to quantify the effect of infrastructure improvements in Brazil.

In a first counterfactual exercise, we found that reductions in transport costs between 1995 and 2017 contributed to nearly a tenth of total deforestation in the country during this period. An important finding is that the impacts were highly heterogeneous across Brazilian meso-regions. By increasing market access and reducing the adoption costs of intermediate inputs, new roads raised the profitability of agriculture in many previously isolated areas. In locations benefiting from a comparative advantage due to high land productivity, the increase in market access pushed prices up to a level high enough that converting more land into agriculture was now profitable. Our model, for example, captured this reasonably well in states like Mato-Grosso or Pará, which have become major deforestation hotspots since the 2000s.

In a second counterfactual exercise, we explored the complementarities between transportation infrastructure and technical change. Specifically, we studied the combined effects of the decrease in transport costs and the introduction of new soybean varieties in the early 2000s. We found that, in some cases, technical change amplified the net effect of roads, while in others, it mitigated it. Our results highlight that, conditional on access to more productive soybean seeds, improved market access in certain meso-regions exacerbated negative impacts on forests. In many parts of the Amazon biome, such as Mato Grosso, where soy production is prominent, our findings confirm that soybean innovation magnified deforestation linked to roads. In other meso-regions, the opposite occurred: the negative impact of market access on the environment was reduced by the presence of highly productive soybean seeds, indicating a pattern of land intensification and again emphasizing the importance of general equilibrium effects.

Our findings depend on assumptions regarding foreign demand in our baseline economy. In this analysis, we have assumed a downward-sloping global demand curve for soybeans and other agricultural commodities. A cynical view might be that global demand for soybeans is almost perfectly elastic, from the vantage point of Brazil. In this case, the expansion of soybean cultivation will not be limited by declining prices.

Our counterfactual exercises shed light on the impact of new roads on deforestation in the historical context of Brazil. The project also emphasizes the importance of general equilibrium effects in assessing how transportation infrastructures affect the environment. Roads do not only alter production in the places they traverse; they also alter incentives in locations much farther away. Our results highlight the need to think carefully about the land use impacts of investment in infrastructure. While roads are widely viewed in economic and policy work as a key tool for growth, their environmental impacts, especially in ecologically sensitive locales, call for caution. Regions with the largest ecological reserves could potentially be negatively affected, sometimes for rather modest economic benefits.

References

Adamopoulos, T. (2025). Spatial integration and agricultural productivity: Quantifying the impact of new roads. *American Economic Journal: Macroeconomics*, 17(1):343–378.

Allen, T. (2014). Information frictions in trade. *Econometrica*, 82(6):2041–2083.

Andersen, L. E., Granger, C. W., Reis, E. J., Weinhold, D., Wunder, S., et al. (2002). *The dynamics of deforestation and economic growth in the Brazilian Amazon*. Cambridge University Press.

Araujo, R., Assunção, J., and Bragança, A. A. (2023). The effects of transportation infrastructure on deforestation in the amazon. *Policy Research Working Paper*, (10415).

Araujo, R., Costa, F., and Sant'Anna, M. (2022). Efficient forestation in the brazilian amazon: Evidence from a dynamic model. *Working Paper*.

Asher, S., Garg, T., and Novosad, P. (2020). The Ecological Impact of Transportation Infrastructure. *The Economic Journal*, 130(629):1173–1199.

Asher, S. and Novosad, P. (2020). Rural roads and local economic development. *American Economic Review*, 110(3):797–823.

Assunção, J., Gandour, C., and Rocha, R. (2023). Deter-ing deforestation in the amazon: Environmental monitoring and law enforcement. *American Economic Journal: Applied Economics*, 15(2):125–156.

Balboni, C. A. (2019). *In harm's way? infrastructure investments and the persistence of coastal cities*. PhD thesis, The London School of Economics and Political Science (LSE).

Bustos, P., Caprettini, B., and Ponticelli, J. (2016). Agricultural productivity and structural transformation: Evidence from brazil. *American Economic Review*, 106(6):1320–1365.

Conte, B., Desmet, K., Nagy, D. K., and Rossi-Hansberg, E. (2021). Local sectoral specialization in a warming world. *Journal of Economic Geography*, 21(4):493–530.

Costinot, A., Donaldson, D., and Smith, C. (2016). Evolving comparative advantage and the impact of climate change in agricultural markets: Evidence from 1.7 million fields around the world. *Journal of Political Economy*, 124(1):205–248.

de Castro Victoria, D., da Silva, R. F. B., Millington, J. D., Katerinchuk, V., and Batistella, M. (2021). Transport cost to port though brazilian federal roads network: Dataset for years 2000, 2005, 2010 and 2017. *Data in brief*, 36:107070.

de Vries, G., Arfelt, L., Drees, D., Godemann, M., Hamilton, C., Jessen-Thiesen, B., Kaya, A. I., Kruse, H., Mensah, E., and Woltjer, P. (2021). The economic transformation database (etd): content, sources, and methods. Technical report, WIDER Technical Note 2/2021.

Desmet, K., Kopp, R. E., Kulp, S. A., Nagy, D. K., Oppenheimer, M., Rossi-Hansberg, E., and Strauss, B. H. (2018). Evaluating the economic cost of coastal flooding. Technical report, National Bureau of Economic Research.

Desmet, K. and Rossi-Hansberg, E. (2015). On the spatial economic impact of global warming. *Journal of Urban Economics*, 88:16–37.

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

Donaldson, D. (2018). Railroads of the raj: Estimating the impact of transportation infrastructure. *American Economic Review*, 108(4-5):899–934.

Donaldson, D. and Hornbeck, R. (2016). Railroads and american economic growth: A market access approach. *The Quarterly Journal of Economics*, 131(2):799–858.

Eaton, J. and Kortum, S. (2002). Technology, geography, and trade. *Econometrica*, 70(5):1741–1779.

Fajgelbaum, P. and Redding, S. (2022). Trade, structural transformation and development: Evidence from argentina 1869-1914. *Journal of Political Economy*.

FAO and IIASA (n.d.). Global agro ecological zones version 4 (Gaez v4). Technical report. Accessed on September 2021.

FAOSTAT (n.d.). Trade database. Crops and livestock products. Accessed on September 2021.

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

Farrokhi, F. and Pellegrina, H. S. (2023). Trade, technology, and agricultural productivity. *Journal of Political Economy*, 131(9):2509–2555.

Gollin, D. (2002). Getting income shares right. *Journal of Political Economy*, 110(2):458–474.

Gouel, C. and Laborde, D. (2021). The crucial role of domestic and international market-mediated adaptation to climate change. *Journal of Environmental Economics and Management*, 106:102408.

Hsiao, A. (2021). Coordination and commitment in international climate action: evidence from palm oil. *Unpublished, Department of Economics, MIT*.

Huber, S. and Rust, C. (2016). Calculate travel time and distance with openstreetmap data using the open source routing machine (osrm). *The Stata Journal*, 16(2):416–423.

IPBES (2019). Summary for policymakers of the global assessment report on biodiversity and ecosystem services of the intergovernmental science-policy platform on biodiversity and ecosystem services.

Kaczan, D. J. (2020). Can roads contribute to forest transitions? *World Development*, 129:104898.

Laurance, W. F., Clements, G. R., Sloan, S., O'connell, C. S., Mueller, N. D., Goosem, M., Venter, O., Edwards, D. P., Phalan, B., Balmford, A., et al. (2014). A global strategy for road building. *Nature*, 513(7517):229–232.

MapBiomas Project (n.d.). Collection 6 of the annual land use land cover maps of brazil. Accessed on September 2021.

Morten, M. and Oliveira, J. (2024). The effects of roads on trade and migration: Evidence from a planned capital city. *American Economic Journal: Applied Economics*, 16(2):389–421.

Paniagua-Ramirez, A., Krupinska, O., Jagdeo, V., and Cooper, W. J. (2021). Carbon storage estimation in a secondary tropical forest at CIEE sustainability center, monteverde, costa rica. *Scientific reports*, 11(1):23464.

Pellegrina, H. S. (2022). Trade, productivity, and the spatial organization of agriculture: Evidence from brazil. *Journal of Development Economics*, 156:102816.

Pellegrina, H. S. and Sotelo, S. (forthcoming). Migration, specialization, and trade: Evidence from brazil's march to the west. *Journal of Political Economy*.

Pereira, R. H. and Goncalves, C. N. (2024). geobr: download official spatial data sets of brazil. *R package version*, 1(0):18.

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

Pfaff, A. and Robalino, J. (2017). Spillovers from conservation programs. *Annual Review of Resource Economics*, 9(1):299–315.

Pfaff, A., Robalino, J., Walker, R., Aldrich, S., Caldas, M., Reis, E., Perz, S., Bohrer, C., Arima, E., Laurance, W., et al. (2007). Road investments, spatial spillovers, and deforestation in the brazilian amazon. *Journal of Regional Science*, 47(1):109–123.

Porteous, O. (2019). High trade costs and their consequences: An estimated dynamic model of african agricultural storage and trade. *American Economic Journal: Applied Economics*, 11(4):327–66.

Porteous, O. (2024). Agricultural trade and adaptation to climate change in Sub-Saharan Africa. *SSRN Electronic Journal*.

Redding, S. J. and Rossi-Hansberg, E. (2017). Quantitative spatial economics. *Annual Review of Economics*, 9(1):21–58.

Rennert, K., Errickson, F., Prest, B. C., Rennels, L., Newell, R. G., Pizer, W., Kingdon, C., Wingenroth, J., Cooke, R., Parthum, B., et al. (2022). Comprehensive evidence implies a higher social cost of CO2. *Nature*, 610(7933):687–692.

Restrepo, V. S. and Mariante, G. L. (2024). Does conservation work in general equilibrium? *Unpublished working paper, Department of Economics, London School of Economics*.

Smith, A. (1776). An inquiry into the nature and causes of the wealth of nations.

Sotelo, S. (2020). Domestic trade frictions and agriculture. *Journal of Political Economy*, 128(7):2690–2738.

Souza, C. M., Z Shimbo, J., Rosa, M. R., Parente, L. L., A Alencar, A., Rudorff, B. F., Hasenack, H., Matsumoto, M., G Ferreira, L., Souza-Filho, P. W., et al. (2020). Reconstructing three decades of land use and land cover changes in brazilian biomes with landsat archive and earth engine. *Remote Sensing*, 12(17):2735.

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

Storeygard, A. (2016). Farther on down the road: transport costs, trade and urban growth in sub-saharan africa. *The Review of Economic Studies*, 83(3):1263–1295.

Tombe, T. (2015). The missing food problem: Trade, agriculture, and international productivity differences. *American Economic Journal: Macroeconomics*, 7(3):226–58.

Verdin, A., Funk, C., Peterson, P., Landsfeld, M., Tuholske, C., and Grace, K. (2020). Development and validation of the chirts-daily quasi-global high-resolution daily temperature data set. *Scientific Data*, 7(1):303.

Vilela, T., Malky Harb, A., Bruner, A., Laísa da Silva Arruda, V., Ribeiro, V., Auxiliadora Costa Alencar, A., Julissa Escobedo Grandez, A., Rojas, A., Laina, A., and Botero, R. (2020). A better amazon road network for people and the environment. *Proceedings of the National Academy of Sciences*, 117(13):7095–7102.

Zu Ermgassen, E. K., Godar, J., Lathuillière, M. J., Löfgren, P., Gardner, T., Vasconcelos, A., and Meyfroidt, P. (2020). The origin, supply chain, and deforestation risk of Brazil's beef exports. *Proceedings of the National Academy of Sciences*, 117(50):31770–31779.

Appendices for "Agricultural Trade and Deforestation: The Role of New Roads"

A Additional Tables

Table 3: Correlation between forest cover and distance to the closest international port

	Share of land under forests	
Travel Distance	0.0696**	
	(0.0286)	
Travel time		0.1117***
		(0.0223)
Constant	-0.5005	-0.2776*
	(0.3792)	(0.1437)
Observations	137	137

Notes: ***, **, * significant at the 1, 5 and 10% level, respectively. Robust standard errors in parentheses. Geographical unit: meso-regions. Year: 2017.

Table 4: Transport costs estimation

Dep. var:	*Soybeans price differences*	
	OLS	2SLS
Log Transport Costs	0.2810***	0.3434***
	(0.0500)	(0.0591)
Constant	-4.2399***	-3.4686**
	(1.3899)	(1.4459)
Observations	2,153	2,153
F-stat		1901.13

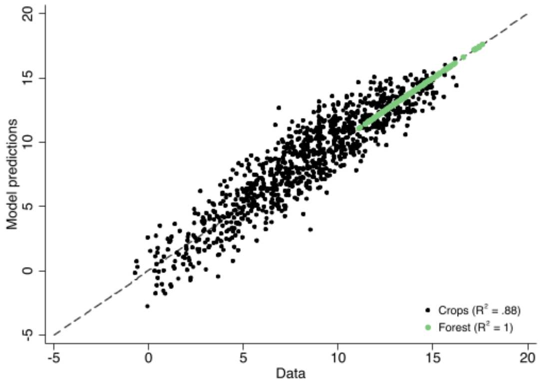
Notes: ***, **, * significant at the 1, 5 and 10% level, respectively. Municipality level. Each regression controls for land productivity, temperature, and precipitation.

Table 5: The economy with past transport costs and low inputs in soybeans

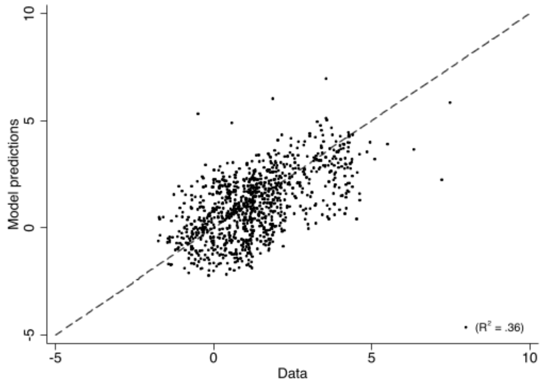
		$\Delta Z_{i,\text{soy}}$	$\Delta Z_{i,\text{soy}} + \Delta\tau_i$
Forest Cover	$\Delta\left(\sum_i F_i\right)$	2.89	3.48
Ag. Income	$\Delta\left(\sum_i \sum_k p_{ik} Q_{ik}\right)$	-6.10	-6.48
Q1 income	-	-6.52	-6.38
Q4 income	-	-0.42	-8.54
Ag. exports	$\Delta\left(\sum_k X_k\right)$	-12.42	-14.51
Input adoption	$\Delta\left(\sum_i h_i\right)$	-5.21	-6.31

Notes: This table reports the predicted rates of change in several outcomes relative to the baseline economy. The first column reports the case where soybean seeds are at low input levels and the transport costs are those of 2017. The second column reports the case where soybean seeds are low inputs but the transport costs are those of 1995.

B Additional Figures



(a) Land-Use



(b) Output per hectare

Figure 7: Model Fit. Notes: All variables are in logs. The dotted line in each figure shows the 45-degree line.

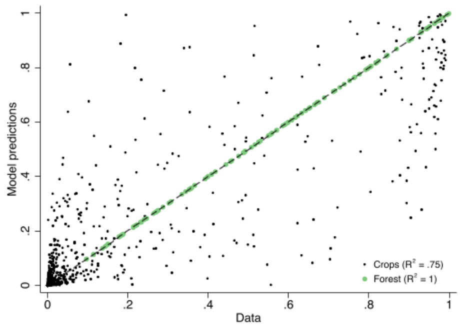
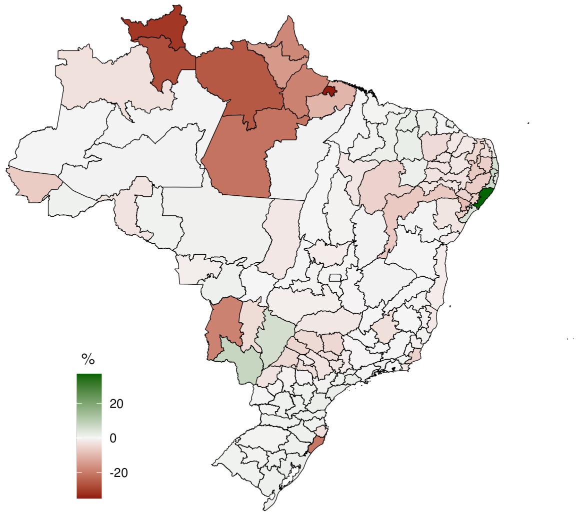
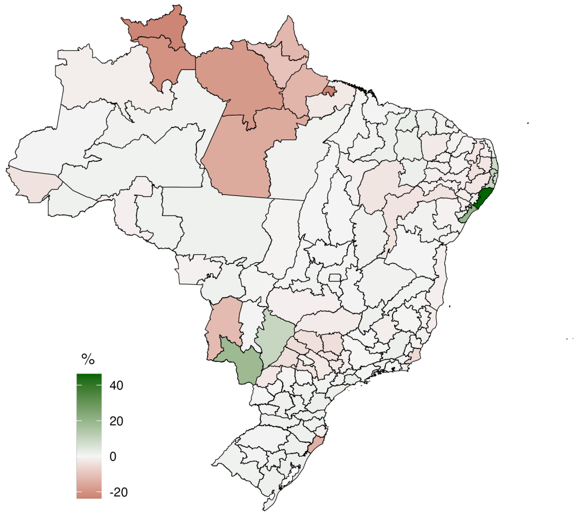


Figure 8: Land Shares Fit. Notes: All crop variables are in shares of total agricultural area. The dotted line shows the 45-degree line.



(a) Predicted Change in Soybeans area



(b) Predicted Change in Pasture area

Figure 9: Soybeans and pasture expansion. Note: this figure plots the predicted rate of change in hectares allocated to soybeans and pasture activities per meso-region with 1995 transport costs compared to the baseline economy.

C Solving the Model

We detail the steps to solve the model.

Agricultural output. Over a given plot ω in region i, the representative farmer minimizes the following total cost:

$$\min_{\{N_{ik},h_{ik},L_{ik}\}} \left[w_i N_{ik}(\omega) + \upsilon_i h_{ik}(\omega) + r_i(\omega) L_{ik}(\omega)\right], \tag{15}$$

subject to:

$$(N_{ik}(\omega))^{\alpha_k} (h_{ik}(\omega))^{\beta_k} (A_{ik}(\omega) L_{ik}(\omega))^{\gamma_k} \geq \bar{Q}_k.$$

In the objective function (15), w_i is wage, υ_i is the price of the intermediate input, and $r_i(\omega)$ is the plot-specific rental rate.

We obtain the following labour- and intermediate input-to-land ratios:

$$\frac{N_{ik}(\omega)}{L_{ik}(\omega)} = \frac{\alpha_k}{\gamma_k} \frac{r_i(\omega)}{w_i},$$
$$\frac{h_{ik}(\omega)}{L_{ik}(\omega)} = \frac{\beta_k}{\gamma_k} \frac{r_i(\omega)}{\upsilon_i}.$$

Note that given the distributional assumptions regarding $A_{ik}(\omega)$, when a plot is open for productive use, the representative farmer uses all land area available in that plot for the most productive crop k, such that $L_{ik}(\omega) = 1$. Substituting each of the above first-order-conditions in the technology constraint, provided that in equilibrium $Q_{ik}(\omega) = \bar{Q}_k$ holds, we can write the marginal cost $MC_{ik}(\omega)$ of producing k in plot ω in region i as:

$$MC_{ik}(\omega) = \frac{w_i^{\alpha_k} \upsilon_i^{\beta_k} r_i(\omega)^{\gamma_k}}{(A_{ik}(\omega))^{\gamma_k} \alpha_k^{\alpha_k} \beta_k^{\beta_k} \gamma_k^{\gamma_k}}.$$

As profit maximization over each plot imposes that the price p_{ik} of a commodity k in region i equals its marginal cost of production, $MC_{ik}(\omega)$, we obtain the rental rate per plot net of the fixed conversion cost given by equation (3).

Combining this result with the ratios of labor-to-land and intermediate input-to-land we obtain the optimal factors' demand per plot:

$$N_{ik}(\omega) = A_{ik}(\omega)\psi_{ik}\frac{\alpha_k}{\gamma_k}w_i^{-1}, \tag{16}$$

$$h_{ik}(\omega) = A_{ik}(\omega)\psi_{ik}\frac{\beta_k}{\gamma_k}\upsilon_i^{-1}. \tag{17}$$

The land shares are then derived by making use of the nested Fréchet assumption regarding investment $A_{i0}(\omega)$ and productivity $A_{ik}(\omega)$, with $A_i(\omega) \equiv \{A_{ik}(\omega)\,\forall k \in K, A_{i0}(\omega)\}$, as in Farrokhi and Pellegrina (2023). We then obtain equations (5) and (7).

Finally, aggregating (16) and (17) over all plots ω in region i gives the regional demand for labour and intermediate inputs, respectively:

$$N_{ik} = \frac{\alpha_k}{\gamma_k}\left(\frac{\psi_{ik}}{w_i}\right)\bar{L}_i Z_{ik}\left(\pi_i^A\right)^{\frac{\theta_1-1}{\theta_1}}\left(\pi_{ik}\right)^{\frac{\theta_2-1}{\theta_2}}, \tag{18}$$

$$h_{ik} = \frac{\beta_k}{\gamma_k}\left(\frac{\psi_{ik}}{\upsilon_i}\right)\bar{L}_i Z_{ik}\left(\pi_i^A\right)^{\frac{\theta_1-1}{\theta_1}}\left(\pi_{ik}\right)^{\frac{\theta_2-1}{\theta_2}}. \tag{19}$$

The regional amount of output per crop is given by:

$$Q_{ik} = \frac{1}{\gamma_k}\left(\frac{\psi_{ik}}{p_{ik}}\right)\bar{L}_i Z_{ik}\left(\pi_i^A\right)^{\frac{\theta_1-1}{\theta_1}}\left(\pi_{ik}\right)^{\frac{\theta_2-1}{\theta_2}}. \tag{20}$$

Land rent in region i is:

$$R_i = \sum_{k\in\mathcal{K}}\gamma_k p_{ik} Q_{ik} - D_{i0}, \tag{21}$$

where we denote the quantity of investment required for deforestation in region i by $D_{i0} = p_0 Z_{i0}\bar{L}_i\left[1-\left(1-\pi_i^A\right)^{\frac{\theta_1-1}{\theta_1}}\right]$.

Non-agricultural output. In our setting, producing the composite good of manufacturing and services requires labour only. The producer's problem is to maximize $\Pi^M = p_0\Lambda(N^M)^{\alpha^M} - wN^M$, with a first-order-condition given by equation (10), with $w = w_i$.

Consumers. Total expenditures in the economy are given by:

$$E = w_i\sum_{i\in\mathcal{L}}N_i^A + wN^M + \sum_{i\in\mathcal{L}}R_i + \Pi^M. \tag{22}$$

Consumption for all crops k can be found by solving the following maximization problem:

$$\max_{\{C_k\}} \left(\sum_{k \in \mathcal{K}} (C_k)^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}} \text{ s.t. } \sum_{k \in \mathcal{K}} p_k C_k \leq b^A E. \tag{23}$$

Provided that $p_k = \tau_i p_{ik}$. The rest of the income, $(1 - b^A)E$, is spent on the outside good.

Equilibrium. The competitive equilibrium in the economy is as defined by section 4.4. As in Sotelo (2020), with the definition of E, the value of exports equals that of imports:

$$\sum_k p_k X_k + E_{\{\text{ROW}\}} = \sum_k p_k M_k + \sum_i \sum_k v_i h_{ik}, \tag{24}$$

where $E_{\{\text{ROW}\}}$ denotes the net consumption of the non-agricultural good by the rest of the world. It follows that labour income in this sector is equal to the sum of total consumption in manufacturing and investment in land clearing, provided factor intensity α^M.

D Data Sources

Geographic features and infrastructures. We work at the meso-region level, a formal political boundary defined by the Brazilian statistical bureau (IBGE). We collect data on road network and ports from the National Department of Transport Infrastructure (DNIT). Data on railroads were taken from the MapBiomas Project infrastructure section.

Forests. Our data on forest cover are satellite data taken from the MapBiomas project (collection v6). This project reconstructs annual land use and land cover information between 1985 and 2017 for Brazil. It provides detailed mapping for forests and other land covers (e.g., pasture) at a 30m pixel resolution based on Landsat archive (Souza et al., 2020).

Agriculture. Our main source of agricultural data is the agricultural censuses from 1995 and 2017 provided by IBGE. We obtain the amount of hectares cultivated for each crop, quantities produced and sold at different administrative levels. Data at the municipality level are taken from the *Producao Agricola Municipal*, available at sidra.ibge.gov.br/pesquisa/pam.

To inform us on estimated potential yields per crop in each location and under different input scenarios, we use the GAEZ data, provided by FAO and IIASA.

Non-agricultural sector. Data for the non-agricultural sector are taken from the Economic Transformation Database (de Vries et al., 2021). We collect information on the number of persons engaged and value-added in the non-agricultural sector.

International trade. For data on agricultural trade with the rest of the world, we use country data from FAOSTAT complemented by data from the Transparency for Sustainable Economies (Trase) initiative (Zu Ermgassen et al., 2020). We obtain quantities imported and exported for the crops included in our analysis.

E Model extensions and robustness checks

E.1 Input adoption and labour mobility

We present two additional results to highlight the role played by intermediate inputs and labour mobility. In Table 6, the column labelled "Output only" displays results from an economy where intermediate inputs still face the transport costs prevailing in 2017, but output faces the transport costs from 1995. The column "No labour mobility" presents results from our main counterfactual (transport costs from 1995 for both output and inputs), with agricultural labour being immobile between meso-regions.

Table 6: The role of intermediate inputs and labour

		Output only	No labor mobility
Forest Cover	$\Delta\left(\sum_i F_i\right)$	0.44	0.32
Ag. Income	$\Delta\left(\sum_i \sum_k p_{ik} Q_{ik}\right)$	-0.31	-0.45
Q1 income	-	0.11	-0.15
Q4 income	-	-6.52	-5.09
Ag. exports	$\Delta\left(\sum_k X_k\right)$	-1.62	-2.31
Input adoption	$\Delta\left(\sum_i h_i\right)$	-0.43	-1.24

Notes: This table reports robustness checks in two scenarios. The first one is when intermediate inputs are not impacted by the change in transport costs. The second one is when labour is immobile across regions.

E.2 The Role of Land Intensity

In this section, we examine the role of land intensity, γ_k, in driving the results of the paper. To highlight its importance, we run our model using different values for this parameter. For this purpose, we simplify the production function described by equation (2) and use a version of the model where factor cost shares are no longer crop-specific:

$$Q_{ik}(\omega) = (N_{ik}(\omega))^{\alpha}(h_{ik}(\omega))^{\beta}\left(A_{ik}(\omega)L_{ik}(\omega)\right)^{\gamma}.$$

We then proceed to run the model, specifically the first counterfactual analysis, as in Section 6.1. That is, we remove roads built in the country since the 1990s. We calibrate the parameters α, β, and γ based on the values of that of soy. Table 7 presents the results.

Table 7: The impact of roads with different land intensity factors

		$\gamma_k = \gamma_{\{\text{soy}\}}$
Forest Cover	$\Delta\left(\sum_i F_i\right)$	1.26
Ag. Income	$\Delta\left(\sum_i \sum_k p_{ik} Q_{ik}\right)$	0.00
Q1 income	-	1.61
Q4 income	-	-3.12
Ag. exports	$\Delta\left(\sum_k X_k\right)$	-2.32
Input adoption	$\Delta\left(\sum_i h_i\right)$	-1.24

Notes: This table shows the rate of change from baseline in forest cover, agricultural income, exports and intermediate inputs use when all crops have the land intensity of soybeans.

When the model is calibrated on soy factor shares, instead of having crop-specific α_k, β_k, and γ_k, the model predicts that new roads led to 1.26% of total forest clearing in Brazil since the 1990s. This is higher than what we find in Section 6.1. This highlights the importance of incorporating multiple crops into the model in order to improve the accuracy of our estimation.

F Numerical algorithm

Given endowments $\{\bar{L}_i, \bar{N}\}$, parameters $\{\alpha_k, \beta_k, \gamma_k, b^A, \theta_1, \theta_2, \Lambda, \sigma, Z_{ik}\}$, trade costs $\{\tau_i\}$, our algorithm solves for crop prices $\{p_k\}_{k\in\mathcal{K}}$ and wage rate $\{w\}$ as follows:

1. Guess crop prices $\{p_k\}_{k\in\mathcal{K}}$ and wage rate $\{w\}$.

2. Compute local price p_{ik} and profitability index ψ_{ik}.

3. Compute land fixed cost Z_{i0} and land productivity Z_{ik}.

4. Compute returns to crops Φ_i, land shares π_i^A, π_{ik}.

5. Compute output Q_{ik} and demand for inputs N_{ik}, h_{ik}, exports X_k , imports M_k.

6. Compute clearing investment per region D_{i0} and land rent R_i.

7. Compute labour demand N^M, output Q^M, profits Π^M in the manufacturing sector.

8. Compute expenditures E, consumer price index P.

9. Compute for all crops $k \in \mathcal{K}$ demand C_k and demand for manufacturing goods C^M.

10. Compute, for all crops $k \in \mathcal{K}$, aggregate demand, $D_k = p_k\left(C_k + X_k\right)$, supply, $Y_k = \sum_i p_k Q_{ik}/\tau_i + p_k M_k$; demand for the manufacturing good, C^M ; labour demand, $ND = \left(\sum_i \sum_k N_{ik} + N^M\right)$.

11. Update global crop prices:
$$p_k^{\{\text{new}\}} = p_k\left(\frac{D_k}{Y_k}\right)^{\rho}.$$

12. Update wage rate:
$$w^{\{\text{new}\}} = w\left(\frac{ND}{\bar{N}}\right)^{\rho},$$
where $\rho \in [0,1]$ is a dampening parameter. Convergence is achieved when $\max|\left(D_k - Y_k\right)/Y_k|$ and $\max|\bar{N} - ND|$ are lower than a tolerance threshold ϵ. Otherwise, let $p_k = p_k^{\{\text{new}\}}$ and $w = w^{\{\text{new}\}}$, and return to step (2).